Work Package 4 Final Report: Model inter-comparison and validation in Inner and Outer Loch Linnhe

January 2023

Scottish Government
Riaghaltas na h-Alba

Work Package 4 Final Report:

Model inter-comparison and validation in Inner and Outer Loch Linnhe.

Authors: Meadhbh Moriarty, Philip Gillibrand, Soizic Garnier, Rory O'Hara Murray, Berit Rabe, Stevie Brain, Alejandro Gallego.

Internal reviewers: Tom Adams, Sandy Murray

Report to satisfy requirements for WP4 Milestones 8

Contents

1. Executive Summary

This report describes the extensive bio-physical model calibration and validation exercise carried out within the Wider Loch Linnhe System. Three particle tracking models i) Biotracker; ii) FISCM; iii) UnPTRACK were coupled with a 2011-2013 hindcast of the Wider Loch Linnhe System (WLLS) hydrodynamic model. The WLLS is a sub model domain of Marine Scotland Science's (MSS) Scottish Shelf Model (SSM; Scottish Government 2016). Each particle tracking model underwent an iterative process of model calibration and validation. Validation is a process of comparing the model and its behaviour to the real system and its behaviour. While calibration was achieved here through a sensitivity analysis on key parameters, this is the iterative process of comparing the model with spatially and temporally relevant field data (or expert judgement where no data is available), revising the model as necessary, comparing again, until the model is accepted. This is an important task, as while common sense verification is possible, a strict verification of a model is intractable. Additionally, an assessment of model uncertainty has been integrated through an ensemble model process. This gives decision makers confidence in the model outputs, for a variety of end user needs.

The report includes; Details on the parameterisation of the three particle tracking models used; hydrodynamic hindcast validation results, conclusions and next steps; a detailed sensitivity analysis of various parameters tested in each of the particle tracking models; examples of the comparisons of the three bio-physical with biological observations collected in the field; An example of the ensemble model outputs with key next steps required to investigate ensemble model usage prior to making recommendations for management.

Major findings:

1. Hydrodynamic calibration and validation is a key component of sea lice dispersal modelling. Specific attention should be paid to ensure flow fields, which provide the forcing to advect sea lice in the particle tracking simulation, adequately represent the system.

2. Temperature and salinity are important physical parameters which influence sea lice behaviour and these modelled parameters should be compared to temperature and salinity observations within the system.

3. Particle tracking model calibration and validation is important for sea lice dispersal modelling. The sensitivity analysis carried out highlighted the importance of including appropriate sea lice swimming behaviours to improve model skill.

4. Sentinel fish lice count data provided data which could be compared with confidence to modelled infective pressure (lice days m^{-2}) over the deployment period. Data collection or model validation should be focused in late summer and Autumn when lice numbers tend to be highest.

5. Pelagic sea lice count data was not suitable for linear comparison with modelled lice density. This is due to the patchy and transient nature of lice in the environment.

6. The coupled bio physical models produced different results within the Loch Linnhe system in terms of local lice abundance, but each fitted well to the patterns of variation in lice on salmon in sentinel cages for Autumn 2011, the time period of the sensitivity analysis focus. The “best fit” models were used to create an ensemble model, which allows visualisation of how coherence of the models varied across the spatial domain.

The coupled bio physical models performed “best” in Autumn 2011 when compared to the other periods based on the Pearson Correlation Coefficient.

Major Recommendations:

1. Physical data should be collected in a systematic way in all areas relevant to aquaculture to allow hydrodynamic model validation.

2. Biological data is important for modelling: fish numbers on farms and associated lice counts are required as source data for predictions of sea lice dispersal models, and sentinel fish lice counts provided the most useful data for model validation in this work package. However, for this work, the source data was estimated for 7 of 11 farms in Loch Linnhe based on data from the other 4 farms (because the data from the 7 were not available). Further controlled field trials using sentinel fish when the farm lice numbers can be properly quantified is recommended.

3. Sensitivity analysis should be carried out in a systemic way, and results should be used to improve model performance.

4. Measures of model variability and uncertainty should be communicated to appropriate regulators and managers to help aid robust decision making.

1.1 Work Package Objectives

The overarching goal of this work package was to complete a model inter-comparison and validation exercise based on the Loch Linnhe area on the west coast of Scotland. This involved three of the project partners SAMS from academia, Mowi Scotland Ltd. from industry and Marine Scotland from government research, each applying their own particle tracking models (PTMs) coupled with the flow fields from the Wider Loch Linnhe System (WLLS) hydrodynamic model. The PTMs used were;

i) Biotracker (Adams, 2019), applied by SAMS modeller;

ii) FISCM (Liu et al. 2015, Wolf et al. 2016), applied by MSS modeller;

iii) UnPTRACK (Gillibrand, 2022), applied by Mowi modeller.

Through individual model calibration and validation, and a comparison and validation exercise involving all three models, the goals were to improve confidence in model predictions, through providing a measure of model uncertainty, and provide insights into the importance of detailing different processes affecting lice larvae which differ slightly between the PTMs.

1.2 Overview of Study Area - Loch Linnhe

Loch Linnhe, one of Scotland's largest sea lochs, is located on the west coast, spanning about 60 km from Fort William in the North to the Sound of Mull and Firth of Lorne in the South (Figure 1.1). The Inner and Outer Loch (with depth of >200 m) are separated by the narrow (200 m) and shallow (11 m) sill at the Corran Narrows. Steep hills lead to channelling of the wind along the Southwest – Northeast direction of the loch, which receives large amount of fresh water input. Wind forcing, fresh water input and tides influence the hydrodynamic conditions in the loch (Rabe and Hindson 2017).

The unique characteristics of the Loch Linnhe system has led it to be the study area for many research projects, investigating environmental conditions (Rabe and Hindson 2017 and references therein) as well as the impact of aquaculture (for example Salama et al. 2013, 2018).

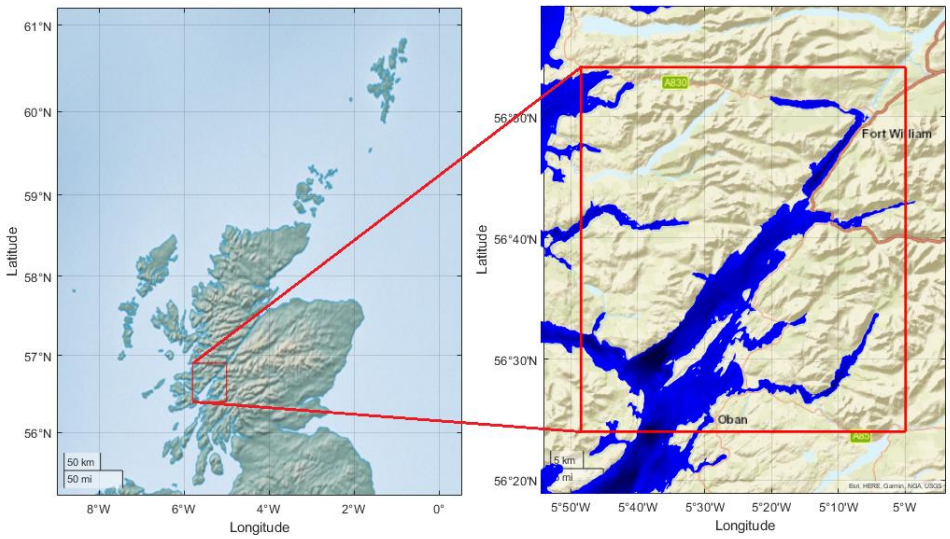


Figure 1.1. Location of Loch Linnhe in Scotland.

2. Hydrodynamic Modelling in Loch Linnhe

Ensuring the underlying flow fields from hydrodynamic models adequately represent reality is a critical component of sea lice dispersal modelling. Much effort was spent on the hindcast hydrodynamic model of the WLLS for 2011-2013 to make certain it describes parameters of importance to sea lice dispersal modelling; temperature, salinity, wind direction, current velocity. Through an iterative process of numerical modelling and validation an acceptable model hindcast was agreed.

2.1 Description of Hydrodynamic Model (WLLS Scottish Shelf Model)

The hydrodynamic model used in this study is the Wider Loch Linnhe System (Price et al. 2016, Figure 2.1(a)) which is an implementation of the

Finite Volume Community Ocean Model (FVCOM, Chen et al. 2003). The regional scale model domain covers the western coast of Scotland from the southern tip of the Mull of Kintyre up to the Isle of Skye in the North, approximately 55 - 57.5° N and from the eastern most reach of Loch Linnhe at about 5 - 7.5° W offshore. The WLLS model has a horizontal unstructured grid with a nominal node spacing of 100-150 m in outer Loch Linnhe and 20-100 m in the inner Loch Linnhe, the side lochs and Loch Sunart, with the nominal horizontal resolution going up to 15 m in some side lochs.

2.2 Hydrodynamic model hindcast development

In order to ensure the most appropriate model hindcast was available for describing the patterns of sea lice during this time period an iterative process of changing model input data, testing model parameters, and comparison with field data (predominately sea temperature) was undertaken. Table 2.1 summaries the key model iterations and the parameters and/or input data that were varied between model runs. All other forcing data and model setup remained unchanged between runs.

Table 2.1: Summary of model setup differences between hindcast versions

Parameter / input data type	Hindcast A	Hindcast B	Hindcast M
Atmospheric forcing data	WRF 1/2° resolution supplied by CEH (Vieno et al. 2014, 2016) for all atmospheric parameters	Wind velocities from WRF 1/18° resolution supplied by CEH (Vieno et al. 2014, 2016). All other parameters from ECMWF ERA5	WRF 1/18° resolution supplied by CEH (Vieno et al. 2014, 2016) for all atmospheric parameters
FVCOM atmospheric forcing methods	- WRF data supplied on regular grid - COARE algorithm applied to WRF data as pre-processing step - Heating calculated OFF - Fresh water heating n/a	- WRF data supplied on regular grid - ERA5 data interpolated to FVCOM grid - Heating calculated ON - Fresh water heating ON	- WRF data supplied on regular grid - WRF data reformatted to allow FVCOM COARE method (heating calculated) on regular grid - Heating calculated ON - Fresh water heating OFF
Horizontal mixing coefficient	0.3	0.3	0.2
Vertical Prandtl number	0.1	0.1	1.0
Convective overturning	ON	ON	OFF
Adcor ON	OFF	OFF	ON

In general, the sea temperature in Hindcast A was too high in the summer and too low in the winter. It was assumed at this stage that this was due to the atmospheric forcing data (either the underlying data or an inappropriate use of certain atmospheric fields) as the model was originally developed using ECMWF ERA-Interim data (Dee et al. 2011). There are also a number of methods that can be used to supply atmospheric forcing data to FVCOM and these are summarised to an extent in Table 2.1. Hindcast B

attempted to addressed this unrealistic summer heating, and winter cooling, by using atmospheric heating data from another model, ECMWF ERA5 (Hersbach et al. 2020). The Centre for Ecology and Hydrology (CEH) winds data (Vieno et al. 2014, 2016) were retained, at a higher resolution, as this was significantly higher resolution than ERA5 which was deemed important to adequately represent the orographically steered wind fields across the region. The sea temperature in Hindcast B did not heat up as much as Hindcast A, but was similarly cool in the winter and was too cold in the Autumn.

A number of other model parameters were investigated and eventually Hindcast M was settled on, which used the CEH forcing for all atmospheric parameters – including heating – now supplied to FVCOM on the native high resolution regular grid and using an inbuilt FVCOM COARE algorithm to calculate some parameters. The key difference was the values of the turbulent mixing parameters used, namely the horizontal mixing coefficient and the vertical Prandtl number. Hindcast M provided the most consistent comparison with field data and it is detailed below. Section 2.4 presents a validation of the model, focusing on Hindcast M but including some results from other hindcast versions (A and B) in order to summarise the differences between model versions and show the reasons for settling on Hindcast M.

The model hindcast (hindcast M) was run for three years from 16 December 2010 to 31st December 2013. The hindcast was forced at the open boundary with hourly current vectors, water elevations, temperature and salinity derived from the Atlantic Margin Model 7 km resolution (Edwards et al. 2012; O'Dea 2012) reanalysis (non-tidal) and forecast (tidal) data. The atmospheric forcing, including hourly wind speed vectors across the domain, were from a 1/18° WRF model (https://www2.mmm.ucar.edu/wrf/users/) provided by CEH (Vieno et al.

2014, 2016). Fresh water input from rivers and coastal runoff (river flux) were from a climatology derived from Grid-to-Grid model (Bell et al. 2007; Cole & Moore 2009) output provided by CEH, and a river temperature climatology was used based on a climatological analysis of data from the Scottish River Temperature Monitoring Network (Jackson et al. 2016). The model bathymetry data were derived from high resolution surveys where available, EMODNET, and Admiralty Chart data where no other data were available. In order to resolve the fresher surface layer of the region of interest, WLLS uses vertical coordinates based on a hybrid sigma layer scheme with 10 standard terrain following sigma layers in shallow water (<13 m depth) and 2 fixed 1 m thick surface layers, 2 fixed 2.5 m thick bottom layers and 6 intermediate sigma layers, in deeper water (>13 m depth). Parameters, including temperature, salinity, and current velocities were output every hour for the duration of the simulation. This hindcast model run was chosen after an extensive validation which tested various hindcast versions against observational data to best describe the temperature, salinity and currents across the domain.

2.3 Physical field data

The quality of the hindcast models was assessed by comparing model outputs to relevant physical parameters. An initial validation of Hindcast A utilised temperature, salinity, water elevation and currents data, and showed that the model temperature did not compare well with observations (as outlined in Section 2.2 above). The following validation exercise mainly focused on temperature. Once a final model run was decided upon (Hindcast M) a repeat validation of all parameters was conducted. Figure 2.1 (a) shows the location of the Conductivity Temperature Depth (CTD) profiles taken at various times (March, May, November) across the three-year time span which were used to validate model temperature and salinity. Figure 2.1(b) shows the location of two

moorings, close to Corran Narrows (inner Loch Linnhe) and the island of Tiree (further out in the model domain), which provided time series data for periods corresponding to part of the model run, and the location of tide gauge and water level recorder and profiling current meter (Aanderaa Recording Doppler Current Profiler, RDCP) time series used for model validation. Observational profiles of temperature and salinity, taken at different times over the course of the model run, were compared with model profiles at approximately the same locations, both in Loch Linnhe and further offshore. Comparisons were made with temperatures and salinities from the closest model node to observation locations. The Corran Narrows data buoy measured near surface temperature, salinity and flows and the Tiree mooring measured temperature and salinity approximately 11 m above the seabed. Four RDCP deployment locations are indicated in Figure 2.1, from which data from six deployments were analysed. The Kilmaleu RDCP was deployed in April 2021, two deployments were made near Gearradh in April and October 2012, the Corran RDCP was deployed in January 2013, and two deployments were made near Gorsten in January and March 2013 (Gorsten, and Gorsten North). The length of the data time series from these RDCP deployments varied, but they were typically longer than 1 month.

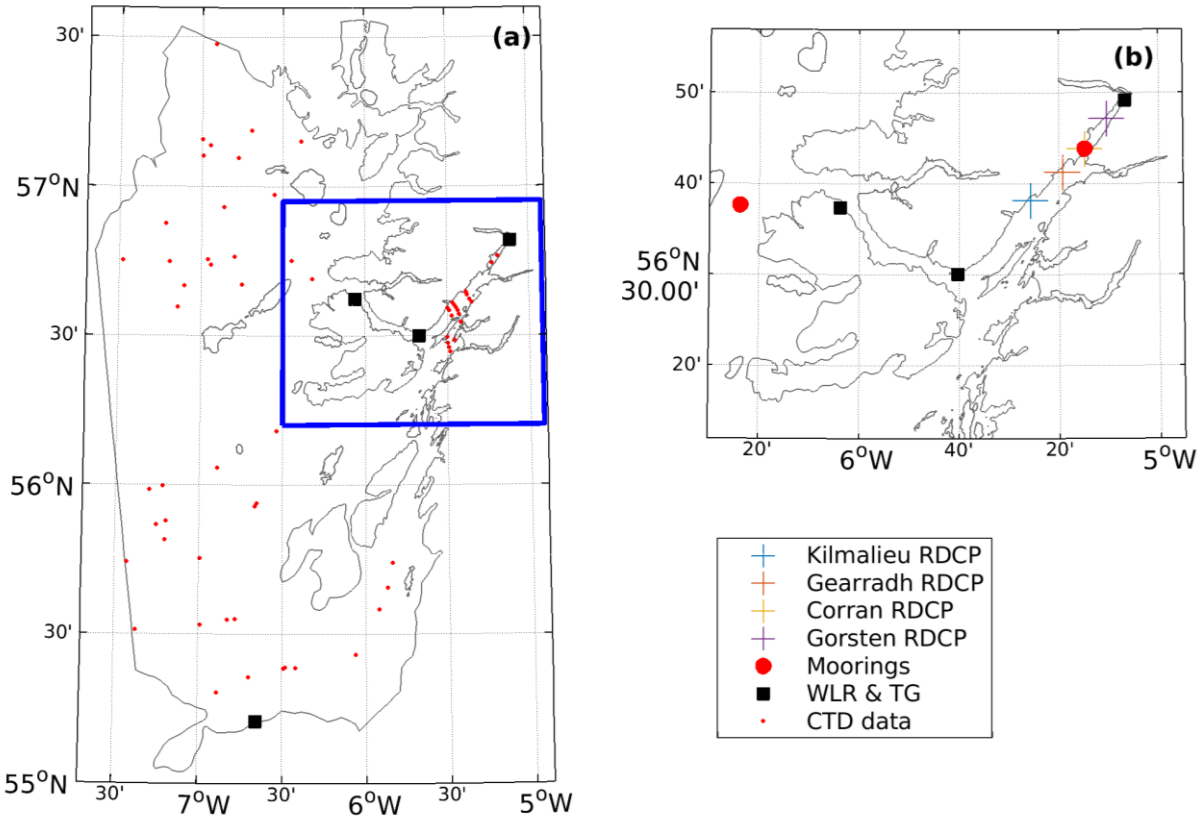


Figure 2.1 (a) WLLS model domain showing the CTD data location and (b) Loch Linnhe region showing water level recorder (WLR) and tide gauge (TG), moorings and RDCP locations.

2.4 Comparison of hydrodynamic models to physical field data

Figure 2.2 shows a temperature time series comparison between the models and the Corran Narrows and Tiree moorings. Both hindcasts A and B compared poorly, whereas hindcast M reproduced both the SST in Upper Loch Linnhe at Corran Narrows and SBT close to Tiree extremely well. CTD data were gathered within Loch Linnhe only during May in 2011-2013, and data outside Loch Linnhe were mainly available from March and November 2011-2013, with 4 profiles from Feb 2012 and 1 profile from Dec 2012. Figure 2.3 shows a scatter comparison of observed (from the CTD profiles) and modelled near surface (interpolated to 5 m depth) temperature and salinity, for the three hindcast simulations. In terms of temperature, hindcast A appears perform well in May and November, but is too cold Feb and March. Hindcast B is too cold most of the time, whereas hindcast M

performs well all the time. In terms of salinity, hindcast A is mainly a little too fresh, but hindcasts B and M perform very well. The R squared values shown in Figure 2.3 confirms this.

Figures 2.4 and 2.5 show comparisons of measured temperature and salinity profiles, respectively, with model hindcasts A, B and M. Only data from Loch Linnhe in May are shown, as this is the primary region (and time) of interest. Hindcast A reproduced the temperature profiles well within Loch Linnhe, whereas hindcast B was consistently too cold. This is not surprising, as the Corran Narrows time series (Figure 2.2) shows that hindcast A reproduced temperature reasonably well in May, but not at other times of the year. Hindcast M also performed very well at reproducing the May temperature profiles in Loch Linnhe (Figure 2.4), and we know that it also performed well at Corran Narrows throughout the year (Figure 2.2). As for the salinity profiles in May (Figure 2.5), all three models performed reasonably well, generally reproducing the shape of the profiles. All models were a little fresher than observed. Close to the bed, hindcast A was consistently around 2 smaller than observed, hindcast B was around 1 less than observed, but hindcast M was extremely close, no more than 0.25 than observed.

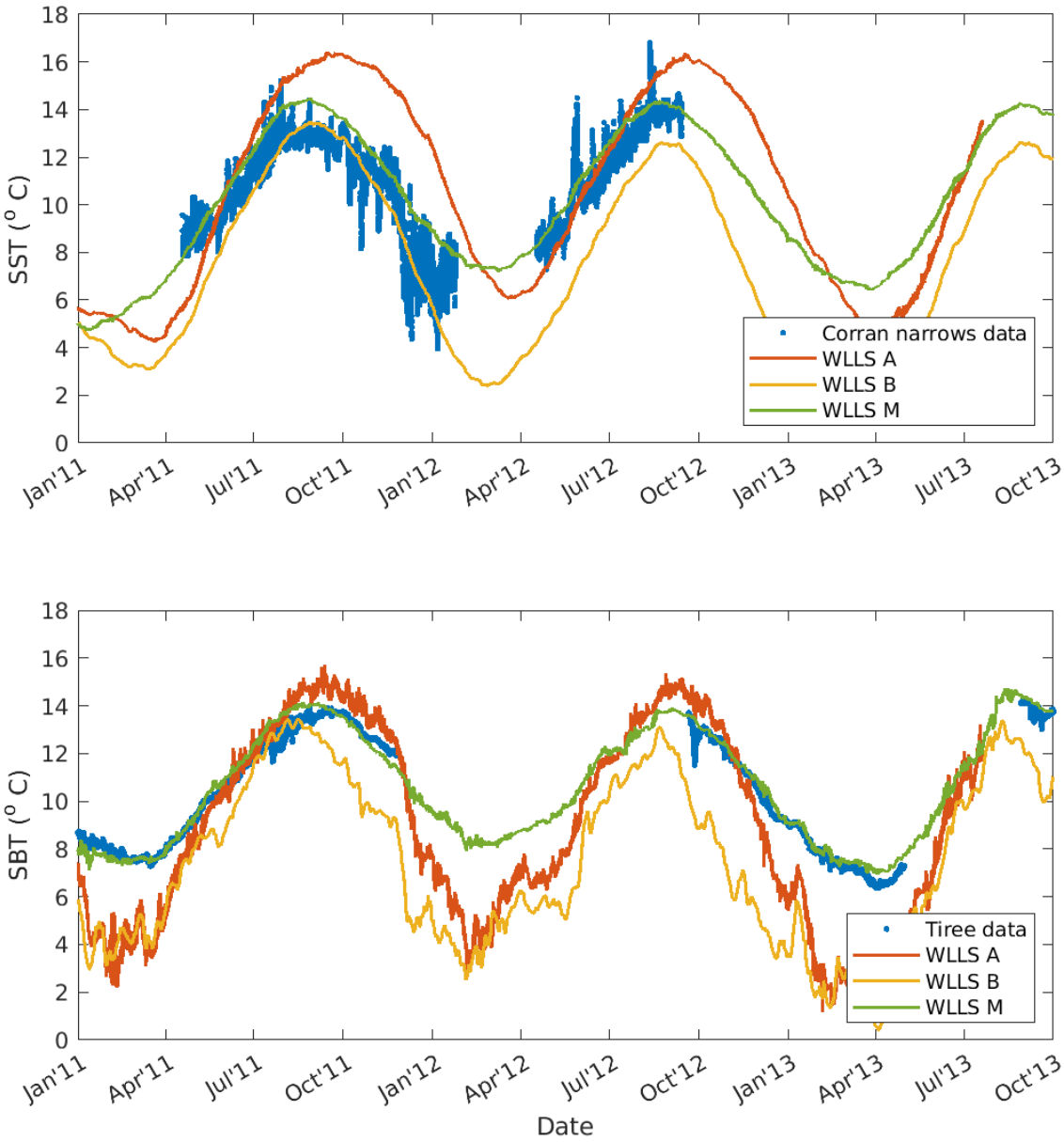


Figure 2.2: Comparison of sea surface temperature (SST) and sea bottom temperature (SBT) approximately 11 m above the seabed, with WLLS at Corran Narrows and Tiree moorings, respectively.

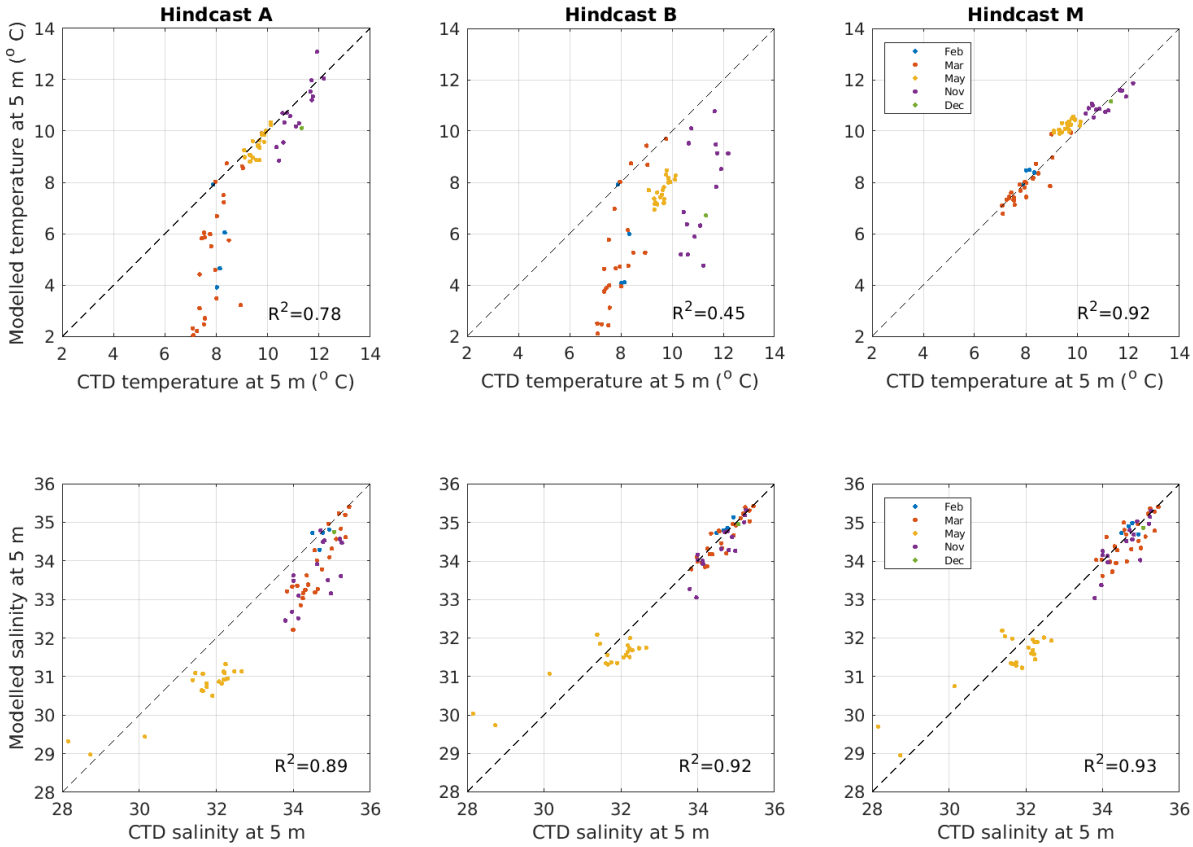


Figure 2.3: Scatter comparison plots comparing observed (CTD) and modelled near surface temperature and salinity, for the three hindcast simulations, with different months in the year indicated. The one-to-one line, indicating a perfect comparison, and the R squared values are shown.

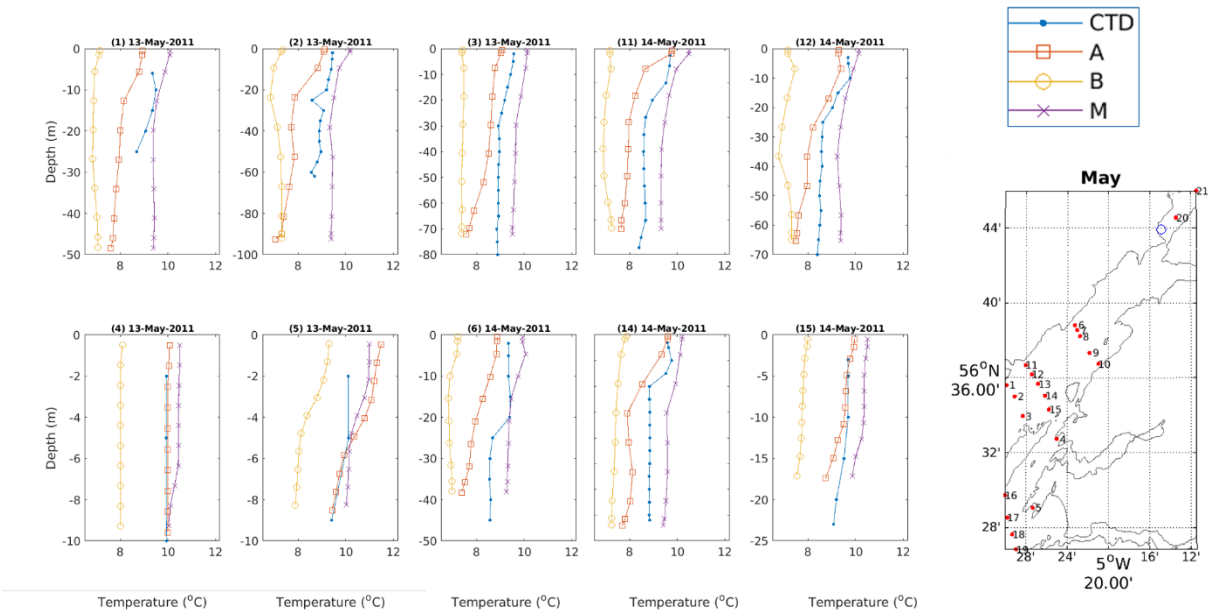


Figure 2.4: Comparison of measured temperature profiles with model hindcasts A, B and M from within Loch Linnhe. The location of each CTD station is indicated in the map on the RHS.

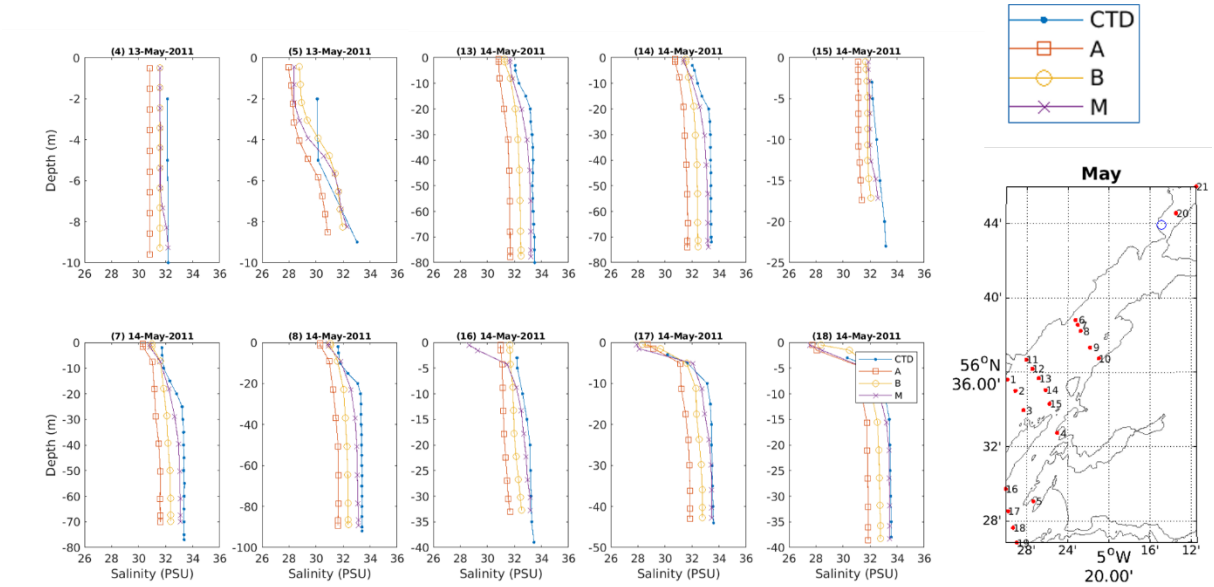


Figure 2.5: Comparison of measured salinity profiles with model hindcasts A, B and M from within Loch Linnhe. The location of each CTD station is indicated in the map on the RHS.

Water elevation data were gathered from two water level recorders (WLRs) and two tide gauges (TG) in the region. The WLRs were positioned in Loch Upper Loch Linnhe (Fort William) and in the Sound of Mull, and the TGs were positioned at Tobermory (Mull) and Portrush (Northern Ireland) (Figure 2.1). Figure 2.6 shows a times series comparison of mean sea level at each location over a 17.5-day period, spanning just over a spring-neap tidal cycle. There is very little difference between the water elevation results from the three model hindcasts, so only hindcast M is shown in Figure 2.6. The three locations well within the model domain compare very well with the data. Portrush compares less well. This is most likely due to its proximity to the model boundary and an Amphidromic point between Isla and the Mull of Kintyre. Table 2.2 shows statistics comparing the performance of the three hindcasts with the observations, and confirms that hindcast M is the best performing.

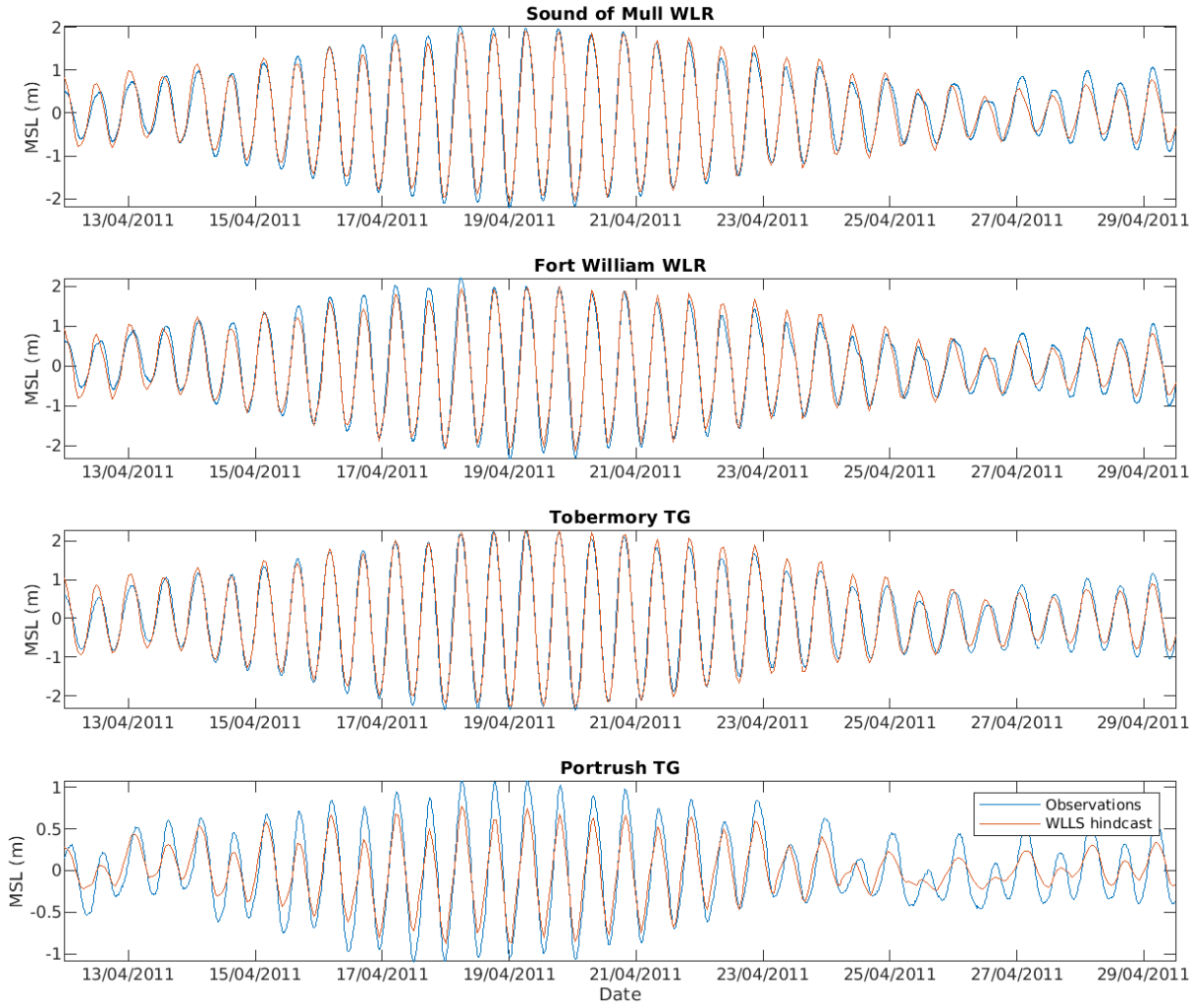


Figure 2.6: Mean Sea Level (MSL) from Water Level Recorders (WLR) and Tide Gauges (TG) compared with WLLS hindcast M.

Table 2.2: Summary of statistics comparing the performance of the three hindcasts with the observations, including difference in near surface (5 m) temperature (ΔT) and salinity (ΔS) and their standard deviations (std) using CTD data from March, May and November, and the Root Mean Square (RMS) difference between observed and modelled water elevations at four locations.

	Hindcast A	Hindcast B	Hindcast M
Mar ΔT ± std (°C)	-2.53 ± 1.95	-2.78 ± 1.79	-0.08 ± 0.35
May ΔT ± std (°C)	-0.18 ± 0.29	-1.93 ± 0.33	0.59 ± 0.21
Nov ΔT ± std (°C)	-0.39 ± 0.66	-3.50 ± 1.75	-0.04 ± 0.33
Mar ΔS ± std	-0.76 ± 0.48	-0.07 ± 0.22	-0.15 ± 0.35
May ΔS ± std	-0.88 ± 0.63	-0.14 ± 0.71	-0.14 ± 0.60
Nov ΔS ± std	-0.91 ± 0.58	-0.23 ± 0.31	-0.17 ± 0.35
Elevation RMS difference (m) Sound of Mull WLR	0.16	0.16	0.16
Elevation RMS difference (m) Fort William WLR	0.19	0.19	0.19
Elevation RMS difference (m) Tobermory TG	0.17	0.17	0.17
Elevation RMS difference (m) Portrush TG	0.19	0.20	0.20

Figure 2.7 shows scatter plots comparing observed near surface velocities from the MSS RDCP deployed near Kilmalieu in Loch Linnhe in April 2011 and modelled near surface velocities from WLLS hindcasts A, B and M. The RDCP data exhibits a broad north east – south west tendency (most probably a tidal ellipse), and also has a consistent westward (roughly onshore) component with speeds up to 0.3 m s^{-1}. This behaviour is reproduced by all three models although most of the modelled data appears to contribute to the north east – south west component. Hindcast M appears to best match the RDCP velocity, as hindcasts A and B produced more scatter outside the range of observed velocities. Figure 2.8 shows a two-

week time series comparison of the same near surface RDCP recorded velocities, and those from hindcast M. Clearly the model does not exactly reproduce the velocity time series, but the amplitude and phase of the velocity components is usually correct. Figures 2.7 and 2.8 both show that the model appears to do least well as reproducing the larger westward velocities (u<0). The time series comparisons of RDCP observations with hindcasts A and B are in Appendix 2.

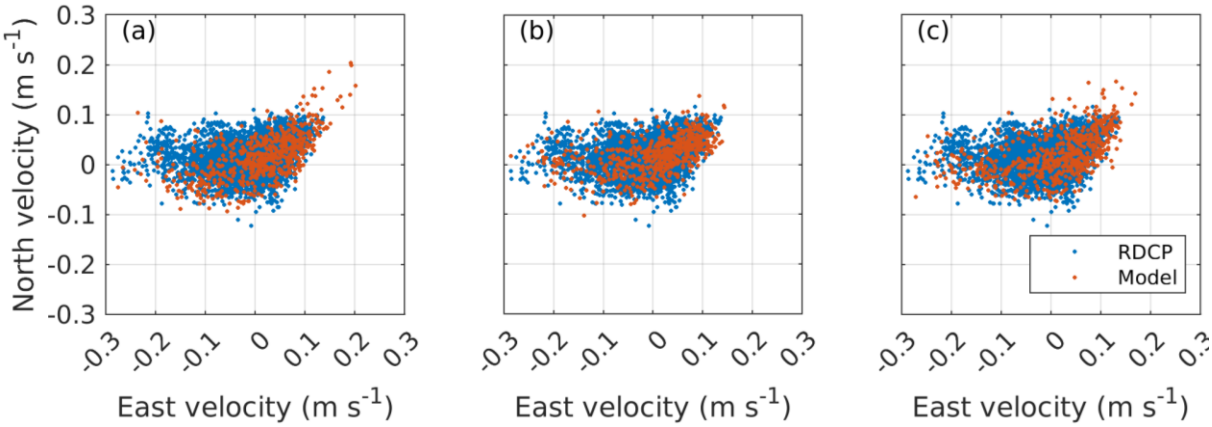


Figure 2.7: Scatter plots comparing observed near surface velocities from the RDCP deployed near Kilmalieu in Loch Linnhe in April 2011 and modelled near surface velocities from WLLS hindcasts A, B and M (a, b and c, respectively).

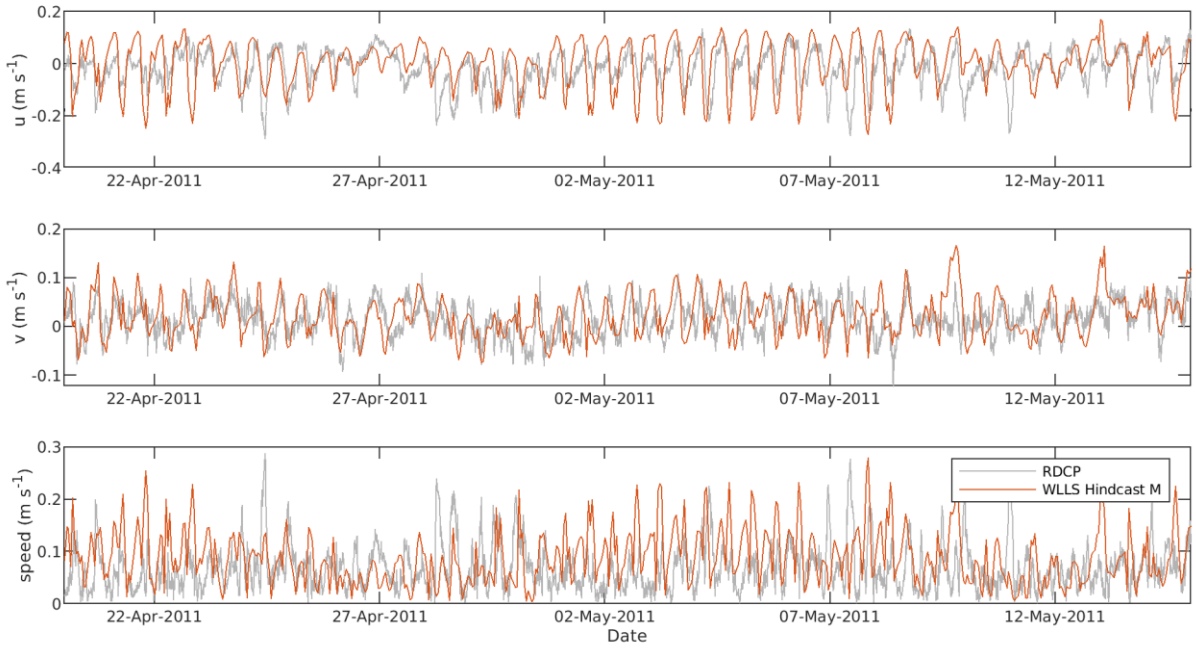


Figure 2.8: Comparison of measured near surface current speeds from a Recording Doppler Current Meter (RDCP) deployed near Kilmalieu in Loch Linnhe in April 2011 with the modelled near surface currents from hindcast M.

Figure 2.9 shows scatter plots comparing observed near surface velocities from the six RDCP time series examined and modelled near surface velocities from WLLS hindcast M. The similarity between observed and modelled values appears to vary within the model domain. The velocity magnitudes and general direction tends to agree in most cases, but in some cases there is often significant scatter in the observed data that is not represented in the model (e.g. Figure 2.9c). This scatter is most likely as short duration wind or fresh water driven event that was not captured by the model. Figure 2.8 shows a two-week time series comparison of RDCP observations near Kilmalieu with WLLS hindcast M. The equivalent time series comparisons at the other RDCP deployment locations and time periods are in Appendix 2.

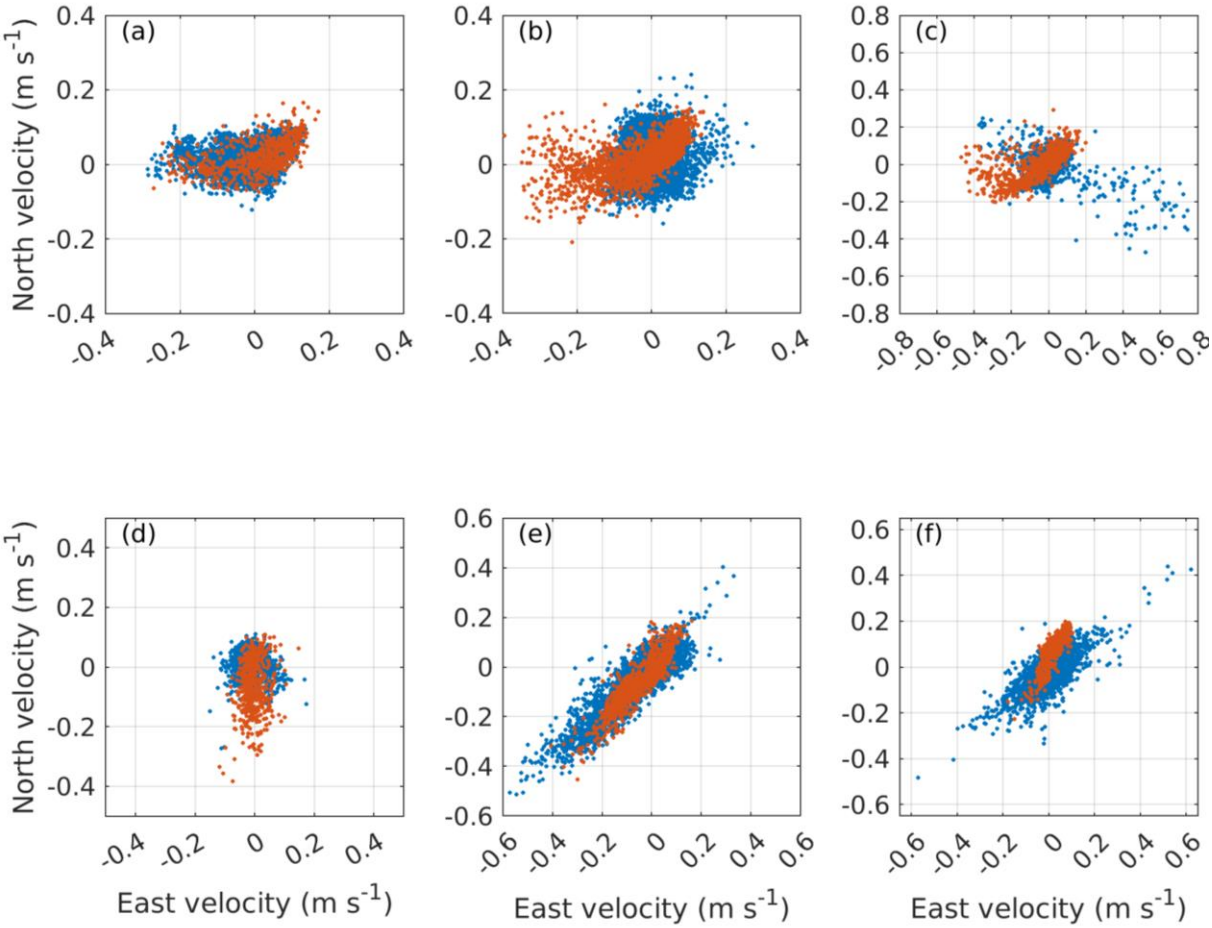


Figure 2.9: Scatter plots comparing observed near surface velocities from the six RDCP time series examined and modelled near surface velocities from WLLS hindcast M. (a) Kilmalieu Apr 2011, (b) Gearradh Apr 2012, (c) Gearradh Oct 2012, (d) Corran Jan 2013, (e) Gorsten Jan 2013, and (f) Gorsten North Mar 2013. Blue shows observed data from RDCP and orange shows data from WLLS model hindcast M.

2.5 Evaluation of hydrodynamic model hindcasts

Within this work package much effort was placed on hydrodynamic model (HDM) evaluation. Decisions on best hydrodynamic hindcast should be made *a priori* based on model validation – i.e. use the HDM which provides the best fit with field data. In this case, hindcast M provided the best agreement with the physical data. Hindcasting requires detailed historical data to drive and to validate the model. A full calibration and validation of the hydrodynamic model may include assessment of Eulerian

parameters (e.g. sea surface height, current velocity, water temperature and salinity) and, preferably, Lagrangian currents (e.g. from dye studies or drifter releases).

2.6 Conclusions and next steps from hydrodynamic modelling

- Flow fields from the hydrodynamic model provide the forcing to advect sea lice in the particle tracking simulation. It is therefore important that modelled currents reflect actual currents as accurately as possible.
- Use atmospheric forcing data (wind, heating) and methods (grid resolution, interpolation, heating calculation) that suit the model simulation and the scale of local processes.
- Appropriate turbulent mixing parameters must be applied for the area.
- To ensure this, use an iterative calibration process to ensure that the chosen atmospheric forcing and turbulent mixing parameters improve the comparison between modelled and observed velocity, temperature and salinity.
- It is challenging to reproduce nearshore currents in hydrodynamic models as the currents field can vary significantly in space and time. In hydrodynamically complex areas is it not always possible to achieve an exact fit to field data, but comparisons should be made to ensure the model is broadly reproducing the expected circulation patterns.

3. Biophysical Modelling in Loch Linnhe

Three particle tracking models were run using hindcast M from the hydrodynamic component of the project to infer the likely distribution of sea lice in the environment for the study period. This was compared to biological field data to quantify the performance of each model.

3.1 Description of Particle Tracking Models

Each model uses a Cartesian system of coordinates where the particles represent a cohort of sea lice in the nauplii stage, which mature into the infectious copepodid stage, this is carried out within UnPTRACK and Biotracker, and through a post processing step in FISCM. These particles move as a result of three processes. First, particles are advected using hourly vertical and horizontal velocity output from the WLLS hydrodynamic circulation model. Secondly, random perturbations due to eddy diffusion are taken into account using a random walk algorithm with specified diffusion coefficients. Finally, particles may be biologically propelled following vertical swimming behaviours. These parameters underwent a one-at-a-time sensitivity analysis detailed in section 3.2.

The simulation periods are consistent with the field study reference periods;

- Spring models run from 1st April - 31st May
- Autumn models from 1st September – 15th November

The time frames for the full simulations were used as they cover the period of field work, after a fifteen day “spin-up” time to allow analysis to occur on a fully established sea lice distribution. The source type and number of particles along with various modelling parameters are further tested using a sensitivity analysis (see section 3.2 for details).

3.1.1 BioTracker

SAMS’ BioTracker code (Adams et al. 2012, 2014, 2016) was written for use with the outputs of the ‘WeStCOMS’ hydrodynamic model (Aleynik et al. 2016, Aleynik 2020). Both this model and the WLLS are based on the Finite Volume Coastal Ocean Model (FVCOM) system (Chen et al. 2013), and as such the outputs are in a fairly similar form. However, some variables

differ between the two models, see Adams & Brain (2021) for details of methodology used to couple this particle tracking model with the SSM.

Particles are moved horizontally subject to the water currents predicted by the hydrodynamic model, in addition to random turbulent diffusion. Larval particles inhabit the upper layer of the water column, and were not allowed to move vertically between layers (Cantrell et al. 2019). Stage durations are dependent on water temperature (which typically vary between 8 – 14 °C in this locality), with particles moving from the non-infective nauplii stage to the infective copepodid stage after accumulating 40 degree-days (1 day at 10 degrees C equates to 10 degree-days). Particles are removed from the simulation after 150-degree days (Johnsen et al. 2016, Samsing et al. 2016). Particles are viewed as "super-particles". This means lice particles are able to infect multiple sites; i.e. they do not end their movement when an infection/arrival event occurs. Particles also have a density (reduced over time by mortality) which governs the predicted overall spatial extent. The weighting (number of lice represented by) of each larval particle was assumed to decay over time at a rate of 0.01 hr^{-1} (Amundrud & Murray 2009, Adams et al. 2016). The particle tracking code is available in an online repository (Adams, 2019).

3.1.2 FISCM

The tracking simulations undertaken by Marine Scotland Science (MSS) were performed by FISCM (FVCOM i-state configuration model), an offline Lagrangian / individual based model for FVCOM (Ji et al. 2011). Particles positions due to advection are updated by FISCM every 30 seconds using a 4th-order Runge-Kutta scheme to solve the Lagrangian equation of motion. FISCM also include a Gaussian random walk approach to introduce turbulent diffusion in the vertical and horizontal dimensions using respective diffusion coefficient. Boundary conditions ensure that particles stay within the model domain: when reaching the land at a new

time step the particle position is reset to the previous time step and reflecting conditions are used for the seabed and the free surface. Finally, MSS adapted FISCM to allow particles to reproduce a diel vertical migration using the following coefficient set by the user:

- maximum swimming depth

- upward velocity (effective between 6am and 6pm)

- downward velocity (effective between 6pm and 6am)

FISCM does not currently have a function for changing behaviour from non-infective to the infective stage, so this is done in a post processing stage, based on age in hours or degree days based on equations in Stien et al. (2005).

3.1.3 UnPTRACK

UnPTRACK (Unstructured mesh Particle Tracking model) is a multi-purpose lagrangian particle-tracking model designed to simulate the transport pathways of pelagic biota, chemical contaminants or particulate wastes using flow fields generated by unstructured mesh HDMs (Gillibrand, 2022; https://github.com/gillibrandpa/unptrack.git). UnPTRACK was developed from an earlier particle-tracking model that used hydrodynamic flow fields from regular grid models; the earlier version has been used to simulate the transport and dispersion of solute veterinary medicines (Willis et al., 2005) and dispersal of pelagic organisms, including sea lice larvae (Gillibrand and Willis, 2007) and harmful algal blooms (Gillibrand et al., 2016). In 2017-2018, the original model was developed to utilise flow fields from the unstructured mesh hydrodynamic models that are increasingly being used in coastal environments, to become "UnPTRACK".

The model runs offline, being forced by flow fields from previously-run HDMs. Advection can be treated using either a fourth-order Runge-Kutta algorithm or a simple Euler approach. A random walk model is used to

simulate horizontal and vertical eddy diffusion. Various aspects of biological development (e.g. temperature-dependent stage development, mortality) and behaviour (e.g. vertical migration, low salinity avoidance) can be simulated and relevant parameters specified in the input file. The basic advection, diffusion and biological algorithms in the model have been described by Gillibrand and Willis (2007) and Gillibrand et al. (2016).

3.2 Particle Tracking Model Development (Parameter Sensitivity Analysis)

To gain an understanding of spatial-temporal distributions of salmon lice populations we rely on modelling which describes both the physical environment and the biological processes that drive salmon lice behaviours. These models require parameterisation, and these parameters are calculated with various degrees of confidence, highlighted in Murray and Moriarty (2021). Prior to using these models for management, it is important to understand how sensitive the model outputs are to changing input parameter values. In addition to understanding how well they predict variability in lice abundances in the environment in space and time.

Here sensitivity analysis (SA) is used to investigate the main parameters and process affecting predictions of biophysical models, by reference to variation from a baseline simulation, and comparison to empirical data within a Scottish sea loch. The aim is to understand whether the current state of knowledge on model input and process uncertainty is sufficient to enable a decision to be taken (Maier et al., 2016) regarding the best fit model parameters or if further observational data is required to better inform our models.

A one-at-a-time sensitivity analysis (OATSA) of key parameters was carried out to get an understanding of how adjusting each parameter may change the message. This does not allow an assessment of how each

parameter interacts with the other model parameters. The model validation and calibration framework is outlined in Figure 3.2.1.

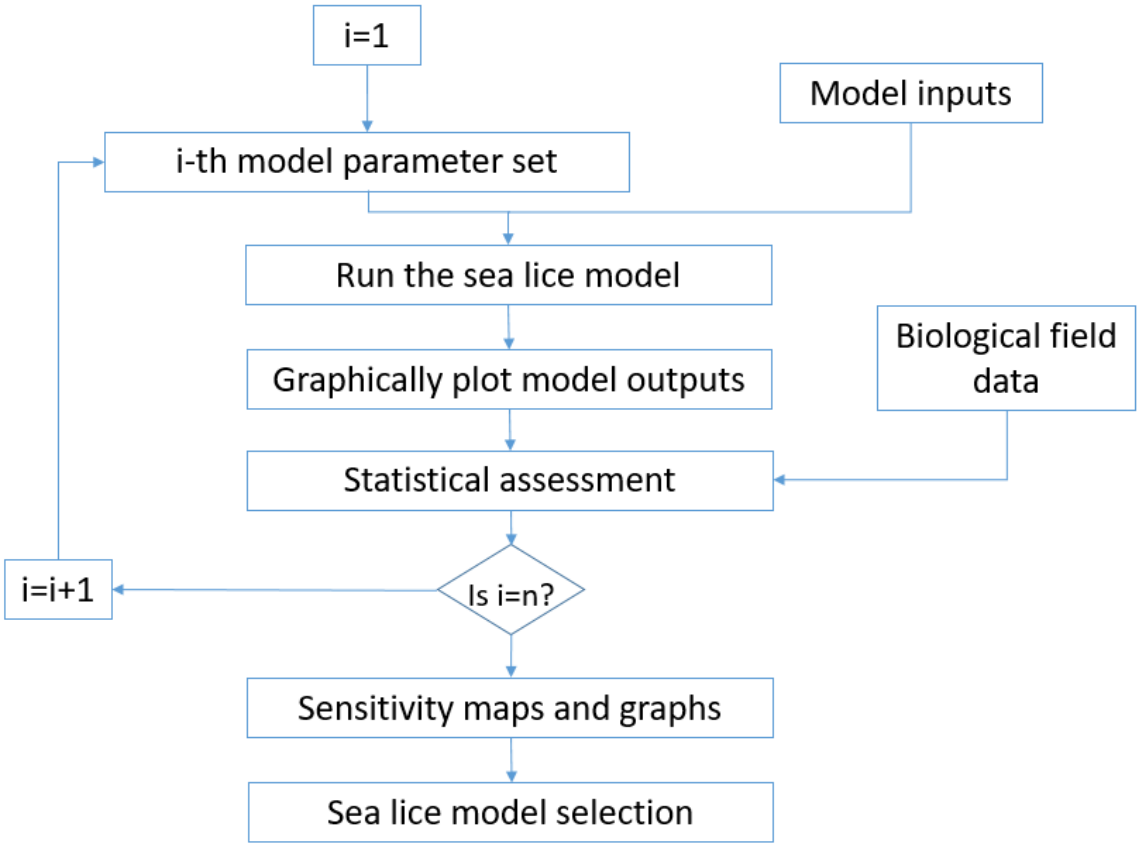


Figure 3.2.1: Framework outlining steps undertaken for bio physical model sensitivity analysis and model calibration. n is the total number of iterations preformed (n differs for each particle tracking model).

In this limited sensitivity test one simulation time period - 01st September – 15th November 2011 was assessed, this time frame was chosen because the data showed the most lice within the system during the time frame.

In all particle tracking models, each adult female louse is assumed to hatch 30 nauplii per day. Standard parameter values were set as outlined in section 3.1, with the parameter being assessed by varying them using a one-at-a-time approach.

Horizontal Diffusivity, K_H: Characterises processes on unresolved horizontal scales. To represent turbulence on a sub-grid scale and may account for fluctuating swimming motion by the salmon louse a range of values between 0 m^2s^{-1} – 10 m^2s^{-1} for the diffusion coefficients were tested in each particle tracking model. Additionally, within the UnPTRACK PTM, a Smagorinsky algorithm was used, which calculates the local horizontal diffusivity as a function of horizontal velocity shear and local grid cell size.

Vertical Diffusivity, K_V: In the vertical, particles can either be passively drifting with the local currents or given swimming behaviours to mimic observed behaviour of sea lice. To represent vertical diffusivity, a random walk in the vertical using a range of values between 0 m^2s^{-1} – 0.1 m^2s^{-1} were tested in Fiscm and UnPTRACK.

Number of Particles, N_P: Within each PTM there is a trade-off which must be made between model performance (speed and output file size) and model accuracy. Thus, we tested a particle source rate of between 2 – 100 Np $source^{-1}$ $hour^{-1}$.

Number of Particles, N_P, Consistency of Results: to assess the consistency between model runs 6 model runs were carried out for each case of 2,10, and 50 Np source $^{-1}$ $hour^{-1}$ in UnPTRACK.

Swimming Behaviour: Salmon lice have limited swimming capacity which can influence their horizontal transport through their vertical positioning. Here we assess the model fit of a range of swimming behaviours within each PTM.

UnPTRACK assessed the following scenarios:

1. “Fixed Depth”: Particles remain at a fixed depth of 1 m throughout (no swimming or vertical diffusion).

2. “Passive”: No swimming or sinking behaviour, but particles are subject to vertical diffusion.

3. “Slow”: Upward swimming speed of 0.5 mm s^{-1} for both nauplii and copepodids (Johnsen et al., 2014) from 6am – 6pm, with a maximum allowed depth of 50 m. Passive at night.

4. “Fast”: Upward swimming speed of 12.5 and 21.4 mm s^{-1} respectively for nauplii and copepodids (Brooker et al, 2018) from 6am – 6pm, with a maximum allowed depth of 50m. Passive at night.

5. “Extended”: “Fast” upward swimming period extended to 5 am – 7 pm, with a maximum allowed depth of 50m. Passive at night.

6. “+Sinking”: “Fast” upward swimming (6 am – 6 pm) with sinking at night, with a maximum allowed depth of 50m. Sinking speeds of 0.9 and 1.0 mm s^{-1} for nauplii and copepodids respectively (Brooker et al., 2018).

Fiscm assessed the following scenarios:

1. “Fast swimmer”: 50 mm s^{-1} upward velocity, 50 mm s^{-1} sinking velocity, with a 15 m maximum depth.

2. “Slow swimmer”: 0.5 mm s-1 upward velocity, 0.5 mm s^{-1} sinking velocity, with a 15 m max depth.

3. “Fast swimmer with sinking velocity”: 50 mm s^{-1} upward velocity, 5 mm s^{-1} sinking velocity , with a 15 m max depth.

4. “Swimmer with passive sinking”: 5 mm s^{-1} upward velocity, 0 mm s^{-1} sinking velocity and 15 m max depth.

5. “Shallow DVM”: 5 mm s^{-1} upward velocity, 5 mm s^{-1} sinking velocity and 5 m max depth.

6. “Deep DVM”: 5 mm s^{-1} upward velocity, 5 mm s^{-1} sinking velocity and 30 m max depth.

7. “No depth limit” DVM”: 5 mm s^{-1} upward velocity and 5 mm s^{-1} sinking velocity.

8. “Fixed Depth”: Particles remain at a fixed depth of 0 m or 15 m respectively (no swimming or vertical diffusion).

9. “Passive”: No swimming or sinking behaviour, but particles are subject to vertical diffusion.

Initial Spread: The exact location of a farm and its cages can vary from the reported mid-point of the farm. Particles were seeded at various radii (0-100m in UnPTRACK and 0-75m in Fiscm), around the farm location to assess the effect of spread.

Source Type: Lice on farm source information is now reported weekly in Scotland, previously it was reported monthly. To assess the impact of different source rates the UnPTRACK model considered the following:

1. “Variable”: Source rates of larvae lice vary in time (from weekly lice counts) and by site (based on lice counts at Mowi sites and weighted by biomass at SSF sites).

2. “Constant”: The total source of lice from each site is averaged over time and set as a constant rate. So total number of lice released from each site is the same as the “Variable” source, but is constant over time.

3. “Fixed”: The total number of lice released by all sites is averaged over time and across sites. So, all sites release the same constant source of lice, regardless of biomass. The total number of lice is the same as the “Variable” and Constant” runs.

4. “COGP”: Estimated number of lice based on reported biomass at each site in October 2011, an average fish weight of 3.33 kg, and a lice count at all sites of 0.5 adult females per fish.

Table 3.2.1 below summarises the parameters assessed and the values used within this limited sensitivity analysis, highlighting which PTM was used to test the parameter.

Table 3.2.1: Parameters tested in sensitivity analysis, PTM = Particle Tracking Model where U = UnPTRACK, B= Biotracker and F= Fiscm, Swimming behaviours are detailed in text only.

Parameter	Unit	PTM	Value Tested						
			1	2	3	4	5	6	7
Horizontal Diffusivity, K_H m^2s^{-1}		F	0	0.001	0.002	0.01	0.1	1	10*
		U	0	0.002	0.02	0.1	1	10	Smagorinsky 0.1/0.2
		B	0	0.001	0.01	0.1	1		
Vertical Diffusivity, K_V m^2s^{-1}		U	0	0.0001	0.001	0.01	0.1		
		F	Variable + DVM	Variable + Fixed Depth	Fixed + Fixed Depth				
Number of Particles, N_P #		B	2	5	10	25	50	100	
		F	5	10	25	50			
		U	2	5	10	25	50	100	
Initial Spread	metres	U	0	10m	50m	75m	100m		
Source Type	NA	U	Variable	Constant	Fixed	COGP			

3.3 Aquaculture study area

During the study period (2011-2013) the fish farms within the system were consented to produce approximately 10% of Scotland's salmon production. This production was undertaken by 10 farms in Loch Linnhe with additional farms feeding in to the system from side-lochs (Figure 3.3.1). In the 2011-2013 period these farms were operated by two companies and were managed in two separate farm management areas; however, these are considered as a single disease management area for regulatory disease control measures.

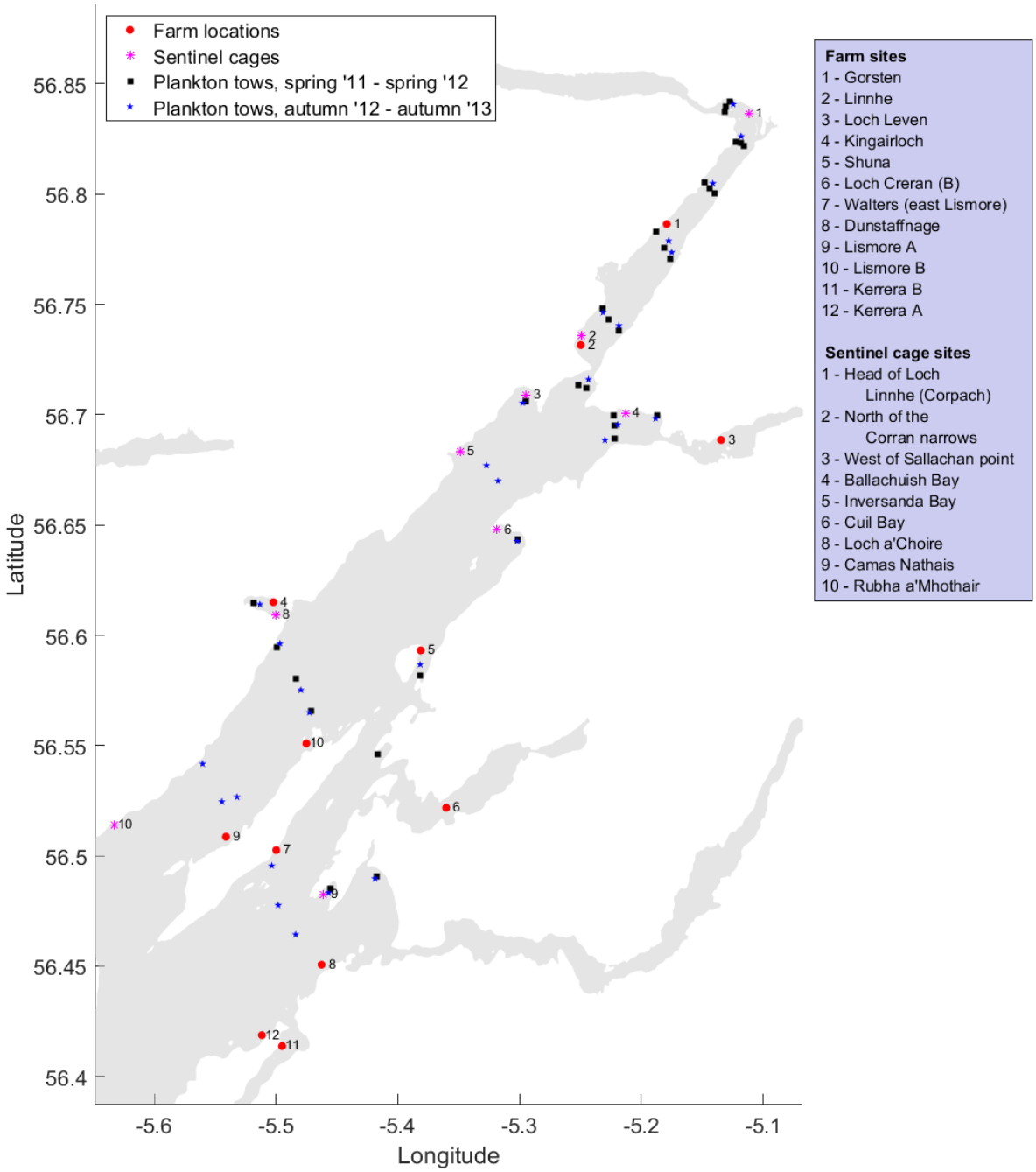


Figure 3.3.1: Map of Loch Linnhe showing locations of particle release points (aquaculture sites), sentinel cage and plankton tow sampling locations.

3.3.1 Farm Site Locations

Details of salmon farm sites in Loch Linnhe used for the model comparison are given in the Table 3.3.1 below. This is reflective of the sites from 2011-2013. These site locations are used to simulate the dispersal

patterns given the main sources of sea lice within the main body of water within the Loch Linnhe System.

Table 3.3.1: The salmon farm/location sites within the Loch Linnhe system used in the comparison study.

Farm Operator	*Farm Location*	*Longitude*	*Latitude*	*Consented Biomass (t)*
Mowi	Gorsten	-5.176	56.786	1990
Mowi	Linnhe	-5.24976	56.73151	2500
Mowi	Loch Leven	-5.12387	56.69248	1450
Mowi	Kingairloch	-5.50223	56.61498	1000
Scottish Sea Farms	Shuna	-5.38125	56.59316	800
Scottish Sea Farms	Loch Creran (B)	-5.36024	56.52181	1500
Scottish Sea Farms	Lismore N	-5.47515	56.55101	1130
Scottish Sea Farms	Lismore S	-5.54119	56.50869	1180
Scottish Sea Farms	Walters (east Lismore)	-5.49998	56.50267	999
Scottish Sea Farms	Dunstaffnage	-5.46273	56.45067	1300
Scottish Sea Farms	Kerrera A	-5.51173	56.41872	500
Scottish Sea Farms	Kerrera B	-5.49505	56.4138	100

3.3.2 Estimating Lice Loads on Farms for Model Validation

The source of lice larvae from the farms was estimated from lice counts on the four Mowi (then Marine Harvest Ltd) sites in Loch Linnhe. Weekly counts of adult females on each site were available (with just the occasional week missing; in total, across all four Mowi sites, five counts were missing from the 160 counts used for the four simulations). Lice count data were not available from the seven Scottish Sea Farms sites.

In order to provide an estimate of lice numbers from all sites, the Mowi lice count data were summed over the four sites and the total biomass of fish was calculated. A weekly average value of adult female lice per tonne of fish was calculated for the Mowi sites. This value was then combined with the reported biomass at the SSF sites over the modelled periods to give a weekly estimate of the number of adult female sea lice at each SSF site. These data are plotted in Figure S1.

Clearly, there is considerable uncertainty in these estimates for lice numbers at the SSF sites. This method assumes similar lice management approaches by the two operators and that the sites were at similar stages of production. A simulation using a fixed release of lice larvae from all sites (i.e., no biomass or time dependence) was performed as a null hypothesis.

Each adult female sea louse was assumed to hatch 30 eggs per day (Heuch et al., 2000, Rittenhouse et al., 2016) as nauplii. The daily number of larvae released can therefore be estimated from the total number of adult females on each site.

3.4 Biological field data

MSS have carried out extensive field sampling and experimental campaigns in Inner and Outer Loch Linnhe over several years, including sentinel cage deployments and plankton tows for the biological field sampling. Historical lice counts are also available from some of the farms. The biological field data for the period 2011-2013 were used here to compare the model results and separately validate three PTMs during that period. This allowed an opportunity to test three different dispersal models against an observational dataset to establish the robustness of dispersal modelling techniques and predictions against independent data.

During the spring and autumn 2011, 2012 and 2013 field seasons (see the six field study reference periods below) sea lice were sampled

using plankton trawls at either ~30 trawling stations sampling for planktonic phase lice (following Penston et al. 2008, described in Salama et al. 2018; see table S2 for locations) and 9 sentinel cage fish (as used by Pert et al. 2014) sites to sample settled stage lice (table S3 for locations) (Figure 3.3.1) (Pert et al. 2021). Details on the plankton station and sampling information and the sentinel cage station and sampling information can be also be found in Pert et al. (2021) supplemented with additional deployment records (Table 3.4.1).

The biological field data were collected during 2011 -2013 time period, and published in 2021 (see Pert et al. 2021). Precise hourly timings of deployments and recovery of sentinel cages are estimated to derive information for model-data comparisons (Table 3.4.1).

Table 3.4.1 Details of timings used for comparison between sentinel cage data and model runs. Temporal definition of deployment was resolved to 12h (for each model run deployment and recovery was set as 12:00h).

Field Study reference periods	Total Number of days	Deployment 1	Deployment 2	Deployment 3
May 9 – May 28, 2011	20	12:00 12/05/11 - 12:00 19/05/11	12:00 19/05/11 - 12:00 27/05/11	12:00 12/05/11 – 12:00 27/05/11
Oct 21 – Nov 09, 2011	20	12:00 26/10/11- 12:00 01/11/11	12:00 01/11/11 – 12:00 09/11/11	NA
Apr 29 – May 18, 2013	22	12:00 30/04/13 - 12:00 07/05/13	12:00 08/05/13 - 12:00 16/05/13	NA
Sep 30 – Oct 19, 2013	20	12:00 02/10/13 - 12:00 09/10/13	12:00 09/10/13 - 12:00 15/10/13	NA

3.5 Evaluation of Sensitivity Analysis Outputs

The following heat maps represent the variance explained (r value) in each model for each parameter, this is an output of the framework for biophysical model sensitivity analysis and model calibration (Figure 3.2.1). Answering the key question: “How much does the model performance change when we vary the parameters within in the model?” This gives us an understanding of the importance of each model parameter within each particle tracking model.

The metric chosen to evaluate the model performance here is the Pearson correlation coefficient (r), which is a measure of the direction and strength of the relationship between two continuous variables. For a correlation between variables x and y, the formula for calculating the sample Pearson's correlation coefficient is given by

$$r = \frac{\sum_{i=1}^{n}(x_i - x)(y_i - y)}{\sqrt{[\sum_{i=1}^{n}(x_i - \bar{x})^2][\sum_{i=1}^{n}(y_i - \bar{y})^2]}}$$

where x_i and y_i are the values of x and y for the ith individual.

Table 3.5.1 Rule of Thumb for Interpreting the Size of a Correlation Coefficient (Hinkle et al 2003, Mukaka (2012).

Size of Correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high positive (negative) correlation
.70 to .90 (−.70 to −.90)	High positive (negative) correlation
.50 to .70 (−.50 to −.70)	Moderate positive (negative) correlation
.30 to .50 (−.30 to −.50)	Low positive (negative) correlation
.00 to .30 (.00 to −.30)	negligible correlation

Pearson was chosen as it is a relatively simple statistical metric for interpretation (Table 3.5.1), which works for data that has an assumption of a linear relationship. The mean lice abundance on sentinel fish data compared to modelled infective pressure (lice days m^{-2}) should fit this assumption, but it is not possible to use r to fairly assess the planktonic field data against modelled density (lice m^{-2}), as these data are zero inflated. This means it would be beneficial to conduct further analysis on all model outputs using additional model performance metrics.

The biological field data (mean lice abundance on sentinel fish) for each time period is compared with the modelled infective pressure (lice days m^{-2}). This metric can be sensitive to outliers so the results were assessed for sites (top) and sites excluding No. 3 deployment 2 which was a consistent outlier in all sea lice dispersal models.

Horizontal Diffusivity, K_H

The value chosen to represent horizontal diffusivity within the model configuration influences the model's skill (Figures 3.5.1 – 3.5.3) . The inference provided by the UnPTRACK PTM suggests that lower constant K_H values (in the range 0 – 0.1 m^2s^{-1}), or using a Smagorinsky model with a coefficient of cs = 0.1 – 0.2, provided a good model fit within the Loch Linnhe system (Figure 3.5.1). The fit was markedly reduced for the (single) simulation with K_H = 10.0 m^2 s^{-1}, demonstrating that this value was too high to adequately represent sub-grid-scale mixing processes on the relatively high-resolution mesh of Loch Linnhe. Similarly, the horizontal diffusivity values taken directly from the FVCOM hydrodynamic model output and input into the particle tracking algorithm were not appropriate for the random walk model used to represent diffusion in the latter (Figure 3.5.1).

The simulation with zero horizontal diffusion (K_H = 0.0 m^2 s^{-1}) extended the radius at which high correlation coefficients were achieved when using UnPTRACK , particularly when Station 3 was excluded (Figure 3.5.1). But when all stations were included, the correlations at the smallest radial distance (50 m) were poorer with the lower values of K_H. Some sub-grid-scale diffusion is also needed to avoid the results becoming sensitive to the initial spread of the particles at the time of release (see 5. Initial Spread Below).

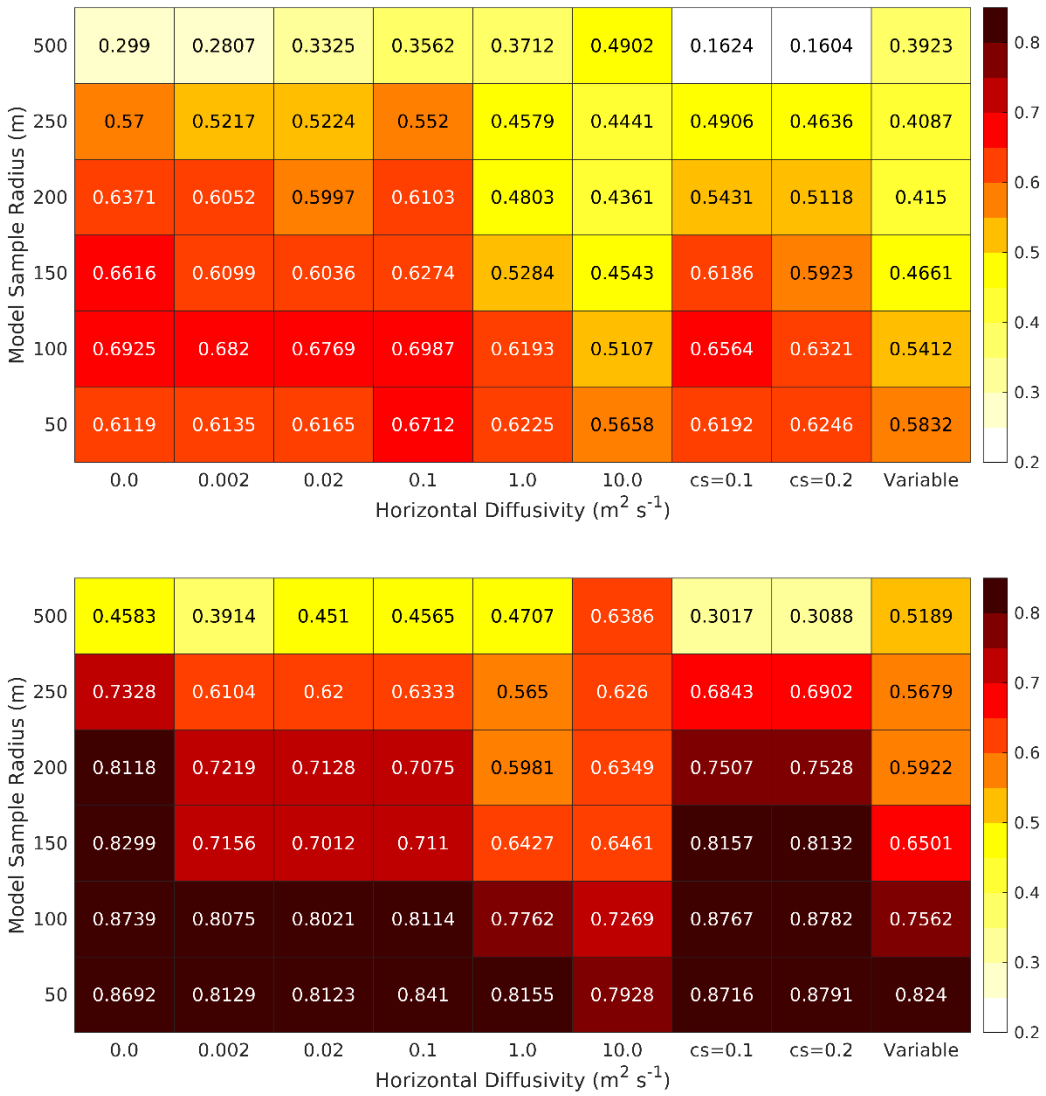


Figure 3.5.1 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for a range of horizontal diffusivity (K_H) values and for different radial distances about the sentinel cage location. The numeric values on the x-axis represent constant coefficients of horizontal diffusivity ($m^2\ s^{-1}$) used in the particle tracking model; the values of cs = 0.1 and cs = 0.2 are coefficient values in the Smagorinsky algorithm for horizontal diffusivity, which calculates diffusivity as a function of grid cell size and velocity shear; the final "Variable" label represents a simulation which used the horizontal diffusivity output by the WLLS hydrodynamic model. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

The inference from Biotracker suggests that K_H values in the range 0.1 -1 m^2s^{-1} produce a good model fit (Figure 3.5.2). Biotracker was run with particles at a fixed depth whereas UnPTRACK had swimming behaviours

included. A similar trend to UnPTRACK was seen with an improved fit when site 3 was excluded.

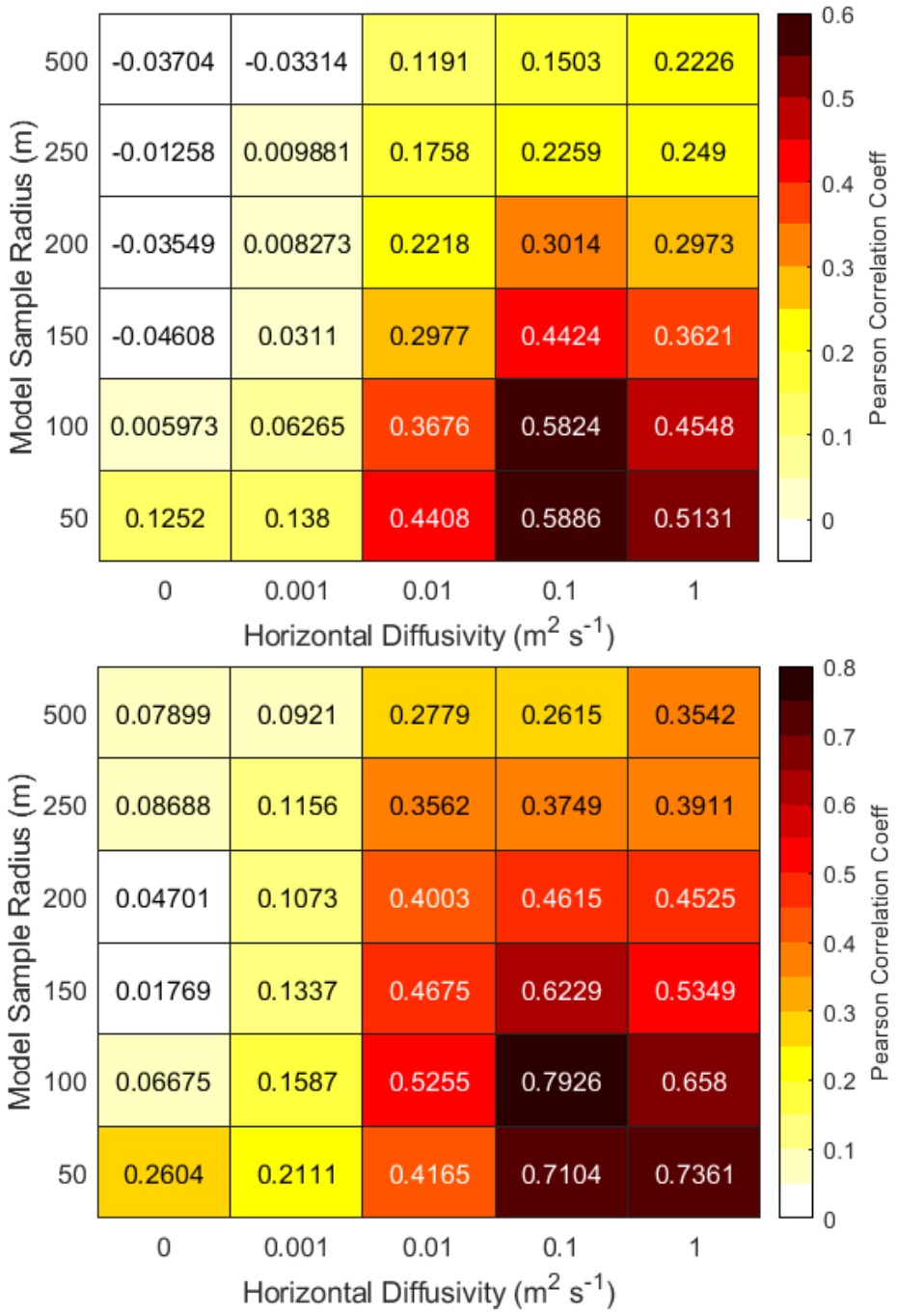


Figure 3.5.2 Biotracker; heatmap showing model skill (correlation coefficient, r) for a range of horizontal diffusivity (K_H) values and for different radial distances about the sentinel cage location. The numeric values on the x-axis represent constant coefficients of horizontal diffusivity ($m^2\ s^{-1}$) used in the particle tracking model. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

The correlations achieved in FISCM suggests that K_H values in the range of 0.002 m^2s^{-1} and 0.1 m^2s^{-1} produce a good model fit, with an optimal fit at 0.01 m^2s^{-1} (Figure 3.5.3). FISCM, similar to UpPTRACK was run with swimming behaviours included. A similar trend to UnPTRACK and biotracker was seen with an improved fit when site 3 was excluded.

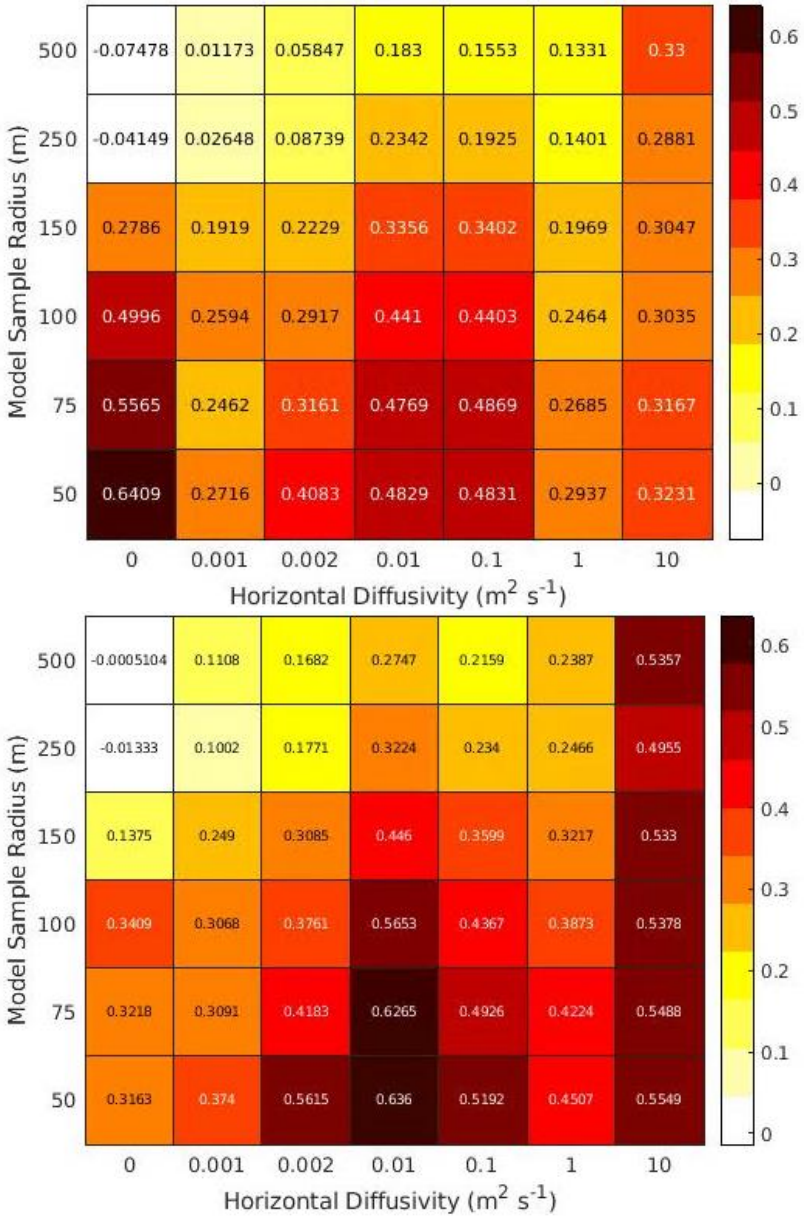


Figure 3.5.3 FISCM; heatmap showing model skill (correlation coefficient, r) for a range of horizontal diffusivity (K_H) values and for different radial distances about the sentinel cage location. The numeric values on the x-axis represent constant coefficients of horizontal diffusivity (m^2 s^{-1}) used in the particle tracking model. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

Vertical Diffusivity, K_V

UnPTRACK does not seem to be as sensitive to the vertical diffusivity parameter value, compared to other model parameter values tested here (Figure 3.5.4). This suggests that the baseline value K_V=0.001 m^2 s^{-1} is appropriate within this model setup. Larger values, greater than 0.01 m^2 s^{-1} did produce weaker correlation coefficients. As vertical diffusion increases, the modelled swimming behaviour of larval lice will be less effective at moving the lice to the sea surface during daylight hours, leading perhaps to more dispersion and, clearly, less agreement with the observed infection pressure at the sentinel cage locations. Using the vertical diffusivities output by the FVCOM hydrodynamic model also did not improve the comparison with data and this approach is not, therefore, recommended for the time being. Note that the vertical diffusivities calculated by FVCOM, or any baroclinic hydrodynamic model, will vary with depth; it is important, therefore, that the particle tracking model is able to correctly simulate particle diffusion in spatially-varying diffusivity (Visser, 1997; Ross and Sharples, 2004).

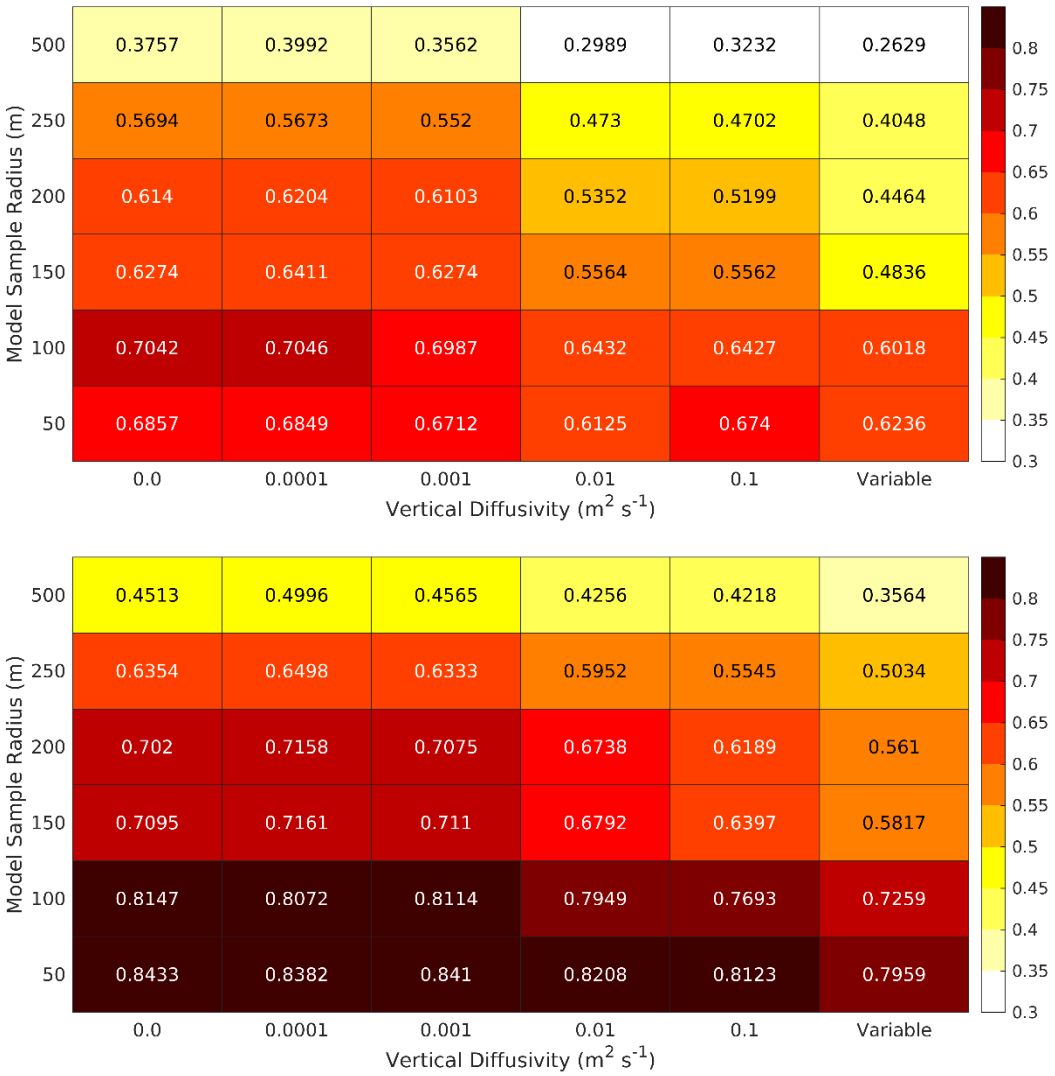


Figure 3.5.4 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for a range of vertical diffusivity (K_V) values and for different radial distances about the sentinel cage location. The numeric values on the x-axis represent constant coefficients of vertical diffusivity ($m^2\ s^{-1}$) used in the particle tracking model; the final "Variable" label represents a simulation which used the vertical diffusivity output by the WLLS hydrodynamic model. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

FISCM does not have the ability to vary vertical diffusivities at fixed rates, so with this particle tracking model the vertical diffusivities calculated by FVCOM are used for particle fixed at the surface and particle with DVM implemented (Figure 3.5.5). Simulating particle diffusion improved model fit when compared with FISCM.

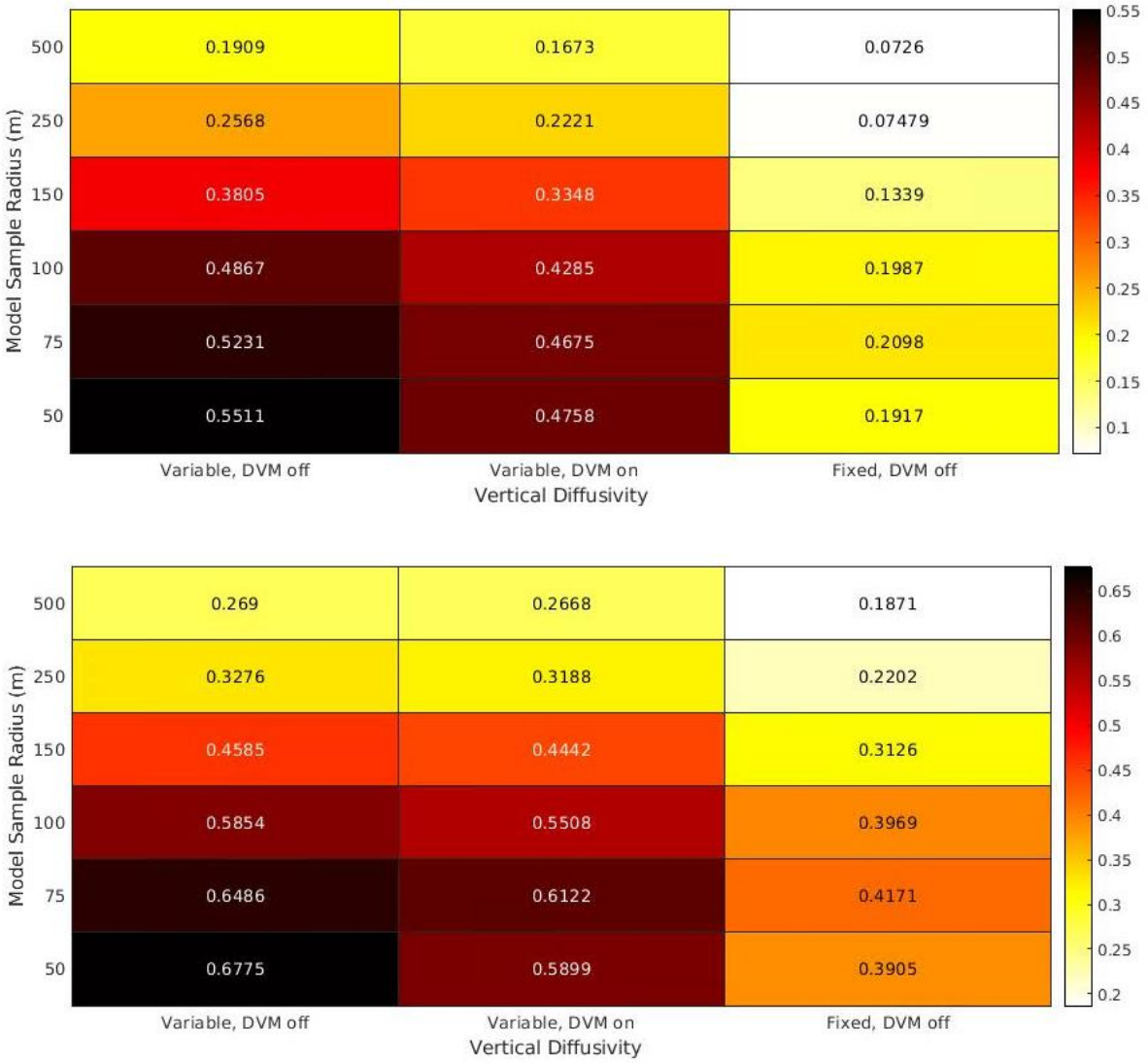


Figure 3.5.5 FISCM; heatmap showing model skill (Pearson correlation coefficient) for a range of vertical diffusivity (K_V) values and for different radial distances about the sentinel cage location. The numeric values on the x-axis represent parameter coefficients of vertical diffusivity ($m^2\ s^{-1}$) used in the particle tracking model; the term "variable" represents simulations which used the vertical diffusivity output by the FVCOM hydrodynamic model, while fixed represents no vertical diffusivity. "DVM on" swimming and sinking was set at 1.8 cm s^{-1}. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

Number of Particles, N_P

UnPTRACK nor Biotracker seem to be overly sensitive to the number of particles released, except for particularly low values (i.e. less than 5 particles per source per hour, Figure 3.5.6 – 3.5.7). Conversely results improve in FISCM at 50 particles per source per hour (Figure 3.5.8).

Using fewer particles leads to reduced consistency in results (tested in UnPTRACK). Six repeat simulations, using the identical model configuration for each particle source rate, led to greater variance around the calculated correlation coefficient (r) for smaller numbers of particles (Table 3.6.1). This shows that while N_P = 2 particles/source/hour may give good inference, the results were more variable than when N_P was increased to 10 or 50 particles/source/hour. The mean values for the correlation coefficient were similar for all three source rates, both radial distances of 50 m and 100 m and whether Station 3 was included or excluded. However, the standard deviation across each repeat set of simulations was markedly higher for lower particle numbers. For example, considering all sites and a radius of 50 m (results column 1 in Table 3.6.1), the standard deviation in r reduces from 0.033 for N_P = 2 to 0.005 for N_P = 50. Thus using more particles makes results more consistent (smaller standard deviation) and mean results from a single simulation are more reliable i.e. closer to the mean. Conversely results from a single simulation with small numbers of particles are less reliable. This is demonstrated in Figure 3.5.6 for sites excluding No. 3, where the correlation coefficient for Np = 5 is less than that for Np = 2, which is likely to be simply random variation in the results with these low particle numbers.

In practice, as many particles as can be handled within the computational limits of the computer should be used (Brickman et al., 2009) and the user should always be aware of increasing limitations and

uncertainties in model results from single simulations as particle numbers are reduced.

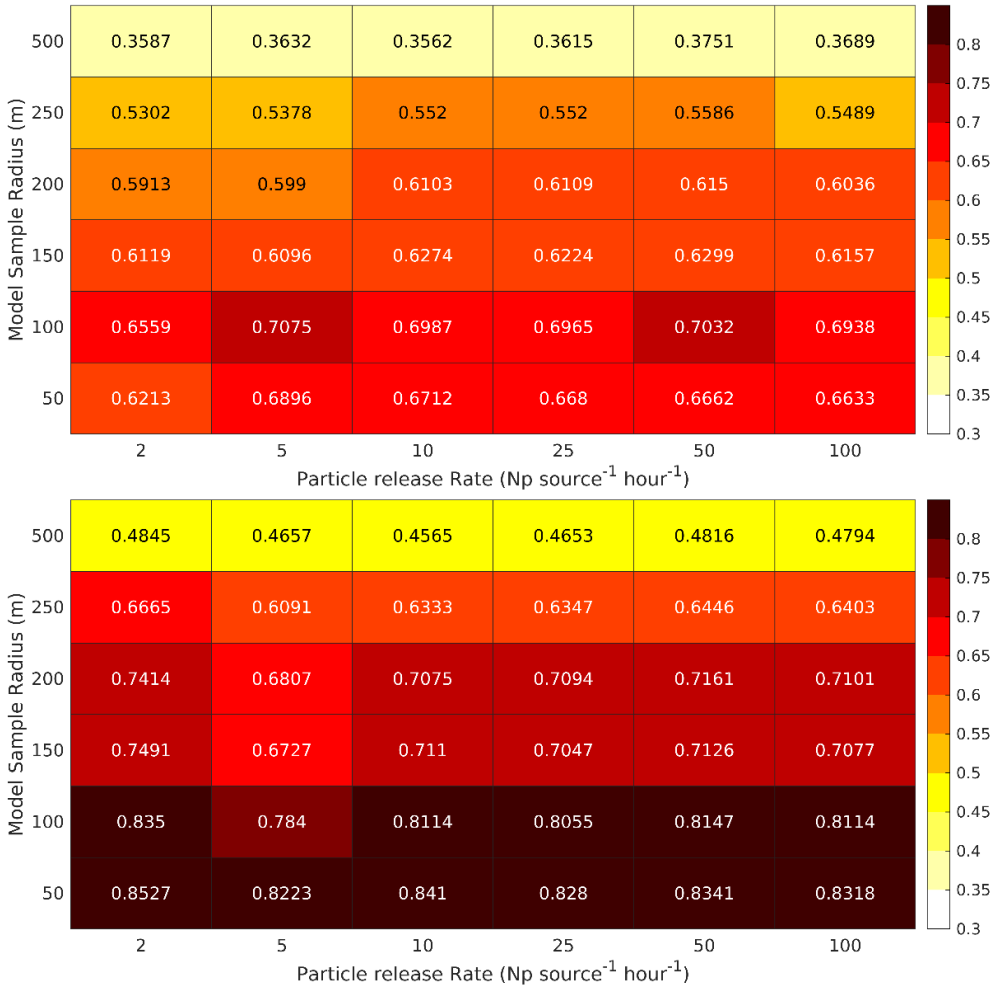


Figure 3.5.6 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for a range of number of particles, N_P (particles per source per hour), released and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

Table 3.6.1 Analysis for consistency of results carried out in UnPTRACK, 6 repeat simulations with identical model configuration for each particle source rate, mean and standard deviation in correlation coefficient.

Particle Source Rate	*Parameter*	*Correlation Coefficient r (all Sites)*		*Correlation Coefficient r (Exc Site 3)*	
(Np/source/hr)		r (50m)	r (100m)	r (50m)	r (100m)
2	Mean	0.646	0.686	0.822	0.809
	St. Dev.	0.033	0.025	0.019	0.014
10	Mean	0.658	0.690	0.839	0.815
	St. Dev.	0.017	0.016	0.008	0.005
50	Mean	0.660	0.695	0.829	0.809
	St. Dev.	0.005	0.004	0.004	0.004

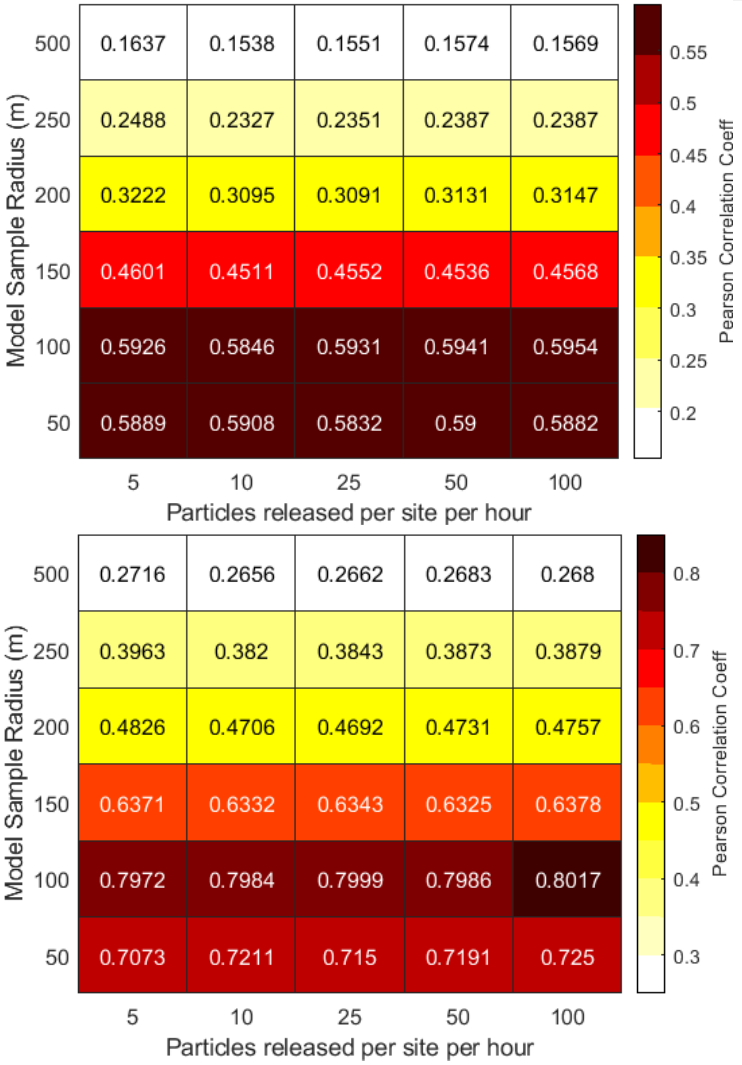


Figure 3.5.7 Biotracker; heatmap showing model skill (correlation coefficient, r) for a range of number of particles, N_P (particles per source per hour), released and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

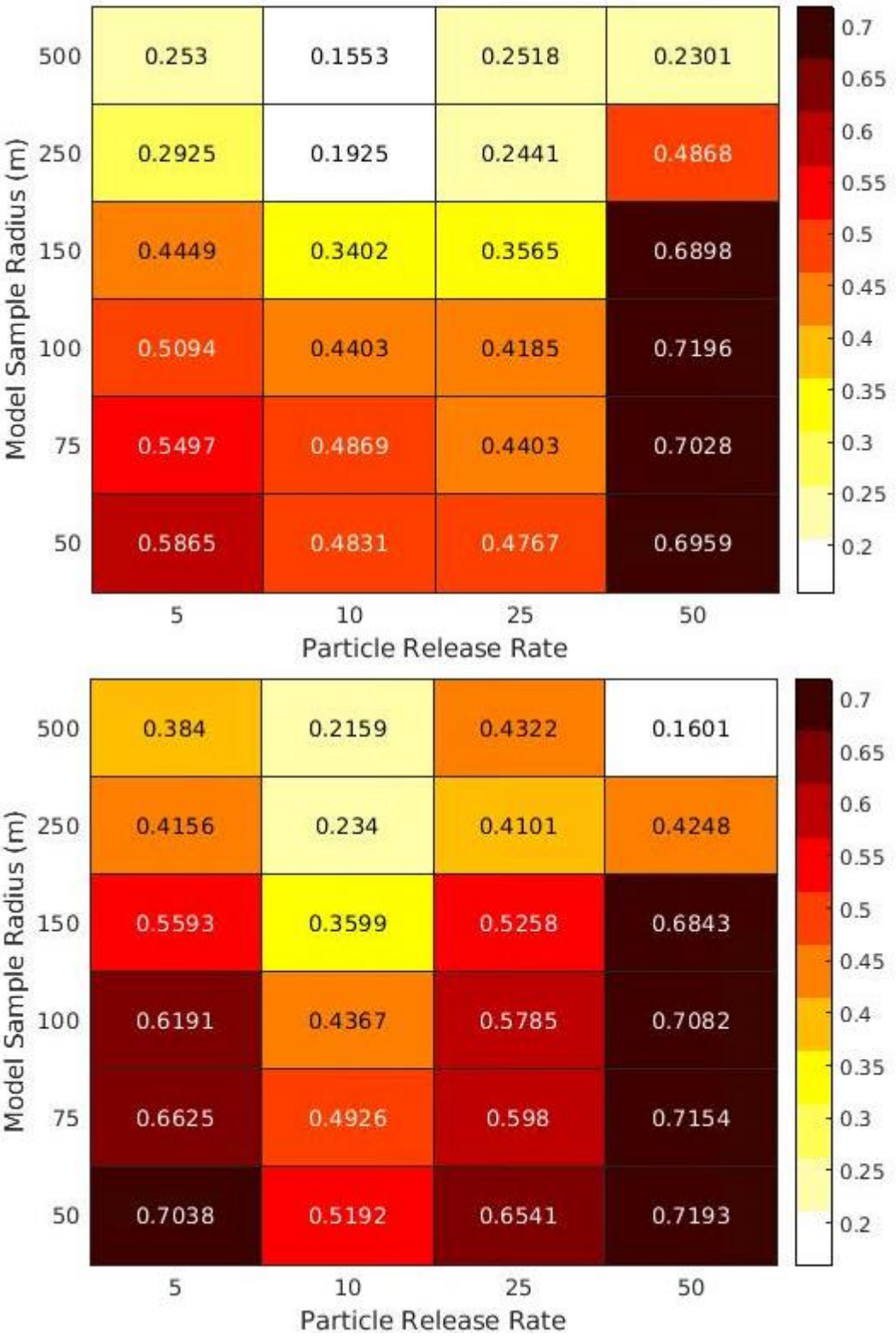


Figure 3.5.8 FISCM; heatmap showing model skill (correlation coefficient, r) for a range of number of particles, N_P (particles per source per hour), released and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011.

Swimming Behaviour

Swimming behaviour strongly influenced the model performance (Figure 3.5.9 and 3.6.10), thus is an important parameter to consider in model development. Within UnPTRACK, using "passive" particles with no biological behaviour, but subject only to vertical diffusion and vertical velocity, produced results with a poor fit to observational data (Figure 3.5.9). This was also the case in FISCM (Figure 3.5.10). Using a "slow" swimming speed of 0.05 cm s^{-1} (Johnsen et al, 2014; Sandvik et al., 2020) for both nauplii and copepodids did not improve model performance in UnPTRACK or FISCM. Releasing particles at a "fixed depth" of 1 m provided better performance than the "Passive" or "Slow" conditions for UnPTRACK, particularly when excluding station No. 3 from the analysis (Figure 3.5.9). When excluding station No. 3 from the analysis in FISCM the fixed depth runs also saw an improvement in fit (2.6.10). In FISCM the "Slow" swimming velocities in the "Shallow DVM", "Deep DVM", "No depth limit" and "Swimmer with passive sinking" did not provide good model fits (Figure 3.5.10).

The best model performance in UnPTRACK was seen for "Fast" swimmers (Figure 3.5.9), with upward swimming speeds of 1.25 and 2.14 cm s^{-1} from 6am – 6pm for nauplii and copepodids respectively (Brooker et al., 2018). Applying the fast upward swimming speeds over an "extended" period from 5am – 7pm, or adding "sinking" (a slow downward velocity of 0.09 and 0.10 cm s^{-1} for nauplii and copepodids respectively (Brooker et al., 2018) at night, produced slightly decreased model performance (Figure 3.5.9). FISCM, which does not have separate stages for nauplii and copepodids, performed best when using the "Fast swimmer with a sinking velocity" (Figure 3.5.10). Combining both "Extended" and "Sinking" in UnPTRACK in an "Extended, Sinking" simulation again improved model performance somewhat (Figure 3.5.9). This suggests that small scale DVM is

an important behaviour in simulating the distribution of sea lice in the environment. Further experimental work, improving understanding of larval behaviour, for both naupliar and copepodid stages, in the Scottish environment would be beneficial.

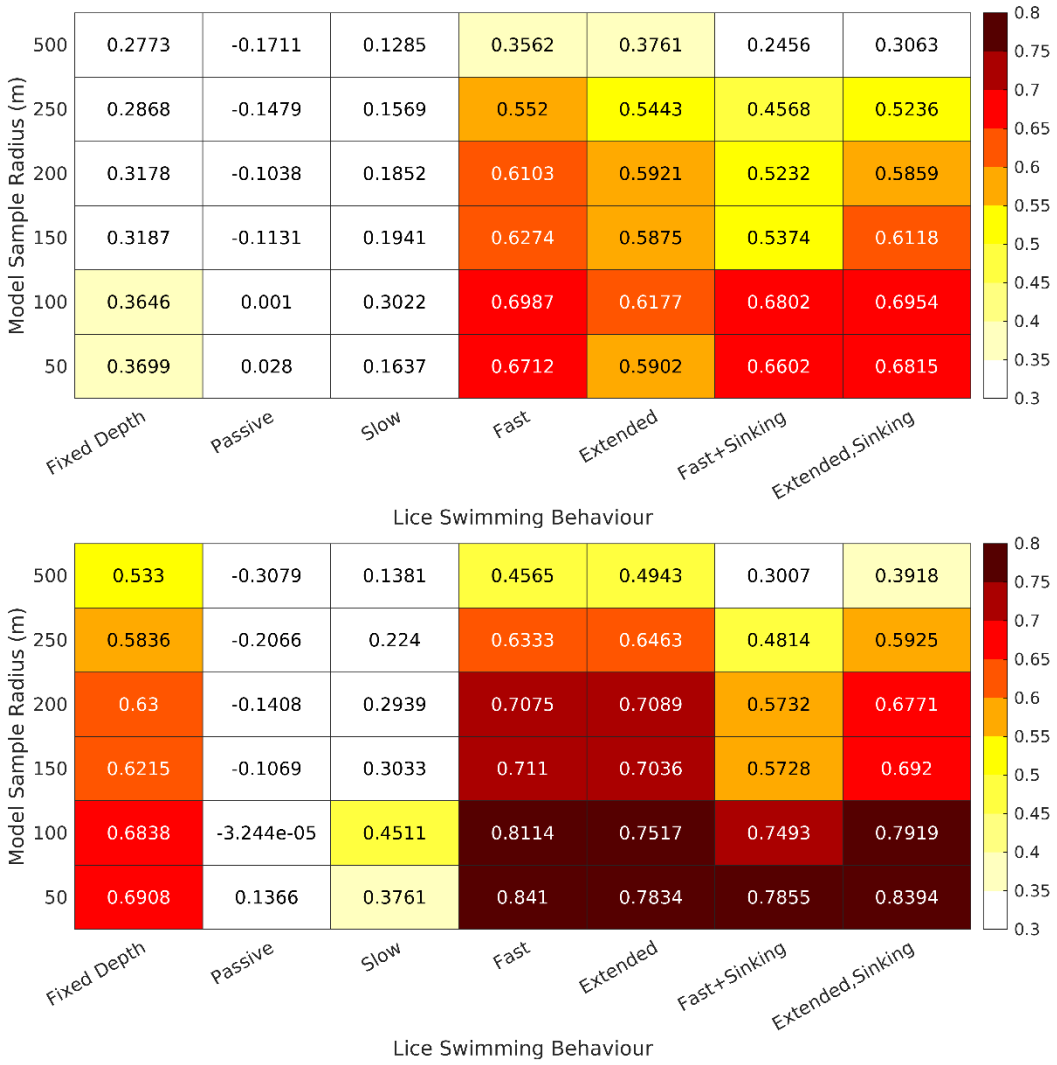


Figure 3.5.9 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for different swimming behaviours and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011. The different types of swimming behaviours are described in the text.

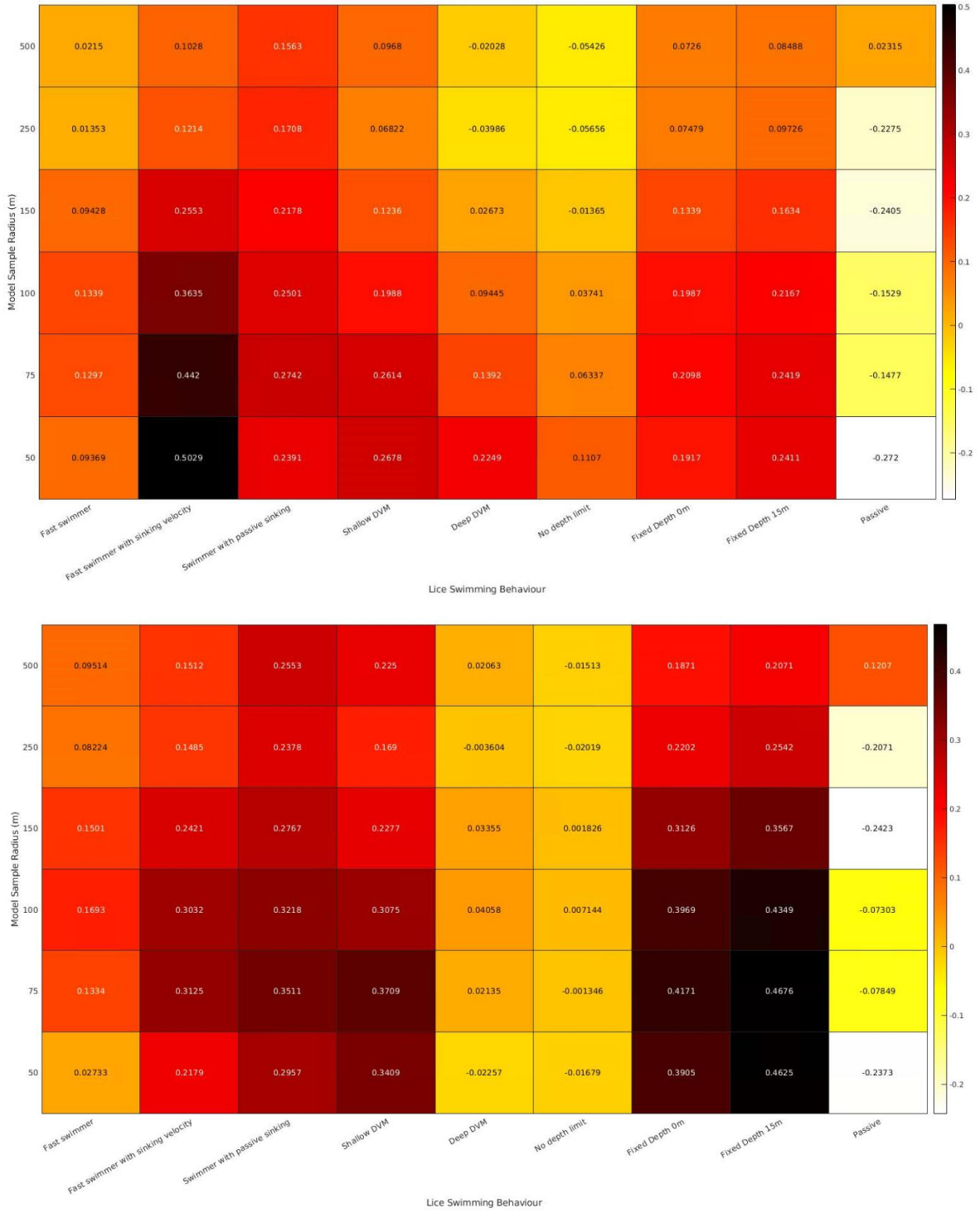


Figure 3.5.10 FISCM; heatmap showing model skill (correlation coefficient, r) for different swimming behaviours and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011. The different types of swimming behaviours are described in the text.

Initial Spread

Inference from the heat map showing the model skill for the range of initial particle positions tested here (Figure 3.5.11), suggests that this input parameter is not as influential as other input values, provided that some horizontal diffusion is included in the simulation. Without horizontal diffusion, the particle trajectories become more deterministic (less stochastic) and the initial release location becomes more significant. However, most sea lice dispersal models do include horizontal diffusion, so the initial spread of the particles at the farm location is not therefore a strong influence (given reasonable definitions of the initial spread).

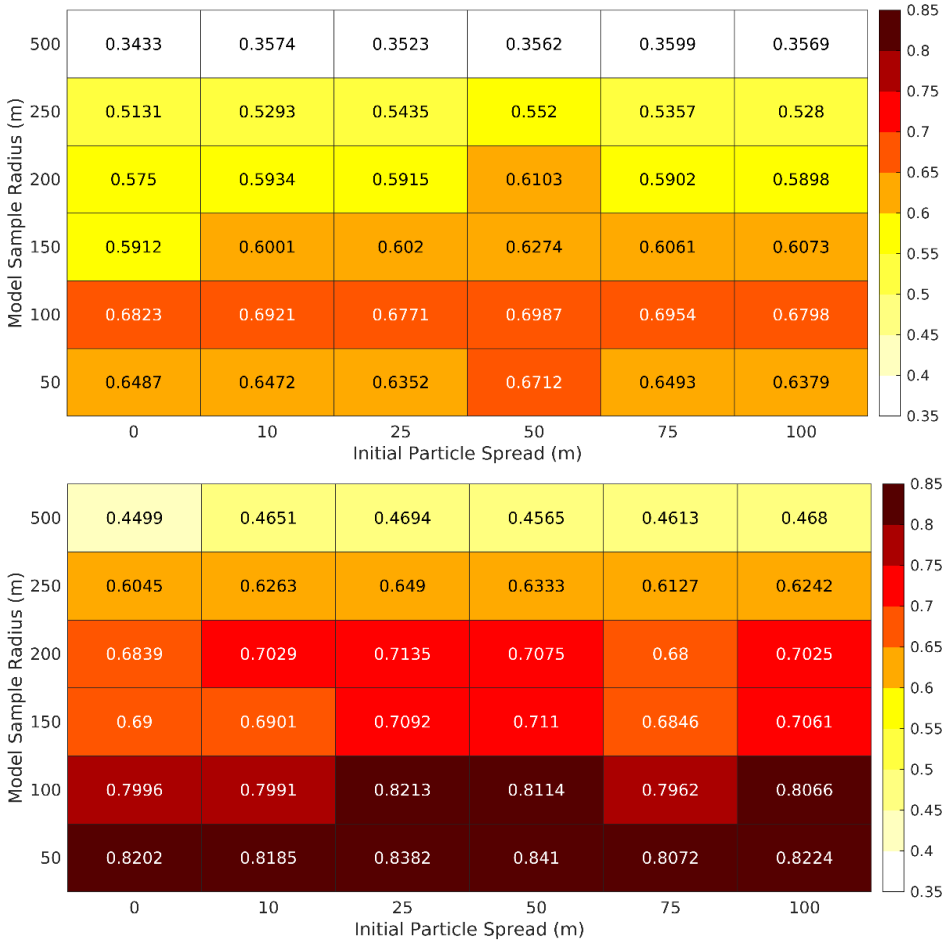


Figure 3.5.11 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for different initial spread of particles and for different radial distances about the sentinel cage location. The correlation coefficients for all sites (top) and sites excluding No. 3 (bottom) are shown for deployment 2, Autumn 2011. The initial spread, Δx, is the distance from the centre of each source location, X_0, within which area particles are initially located at random (i.e. $Xp = X_0 \pm \Delta x$).

Source Type

Simulations were performed with different source types: Variable, Constant, Fixed and COGP (Figure 3.5.12). The “Variable” source utilised the weekly lice counts from the Mowi farms during Autumn 2011 and the recorded fish numbers, extrapolated to the Scottish Sea Farm Sites by a biomass weighting, to provide a time-varying daily source of lice for the simulation. The “Constant” source took the total number of lice per site released over the simulation from the variable source and calculated a daily average rate which was applied. Thus the constant source varied by site, but was the same every day. The “Fixed” source rate took the total number of lice release over the variable simulation from all sites, and distributed them evenly between sites and averaged over time. Thus, each site released the same, constant, source of lice. In these three simulations, “Variable”, “Constant”, “Fixed”, the same total number of larval lice were released into the model domain over the simulation period. The fourth simulation, “COGP” used the reported biomass at each site in October 2011, converted that to a number of fish using as assumed average fish weight of 3.33 kg, and applied the Salmon Scotland Code of Good Practice maximum lice burden of 0.5 adult female lice per fish to given a number of lice per site during the simulation. This approach was to assess, crudely, the relevance of using the COGP approach to predict sea lice abundances in coastal waters. The COGP method resulted in 10.7% more lice larvae being released during the simulation, distributed according to biomass, than the first three simulations.

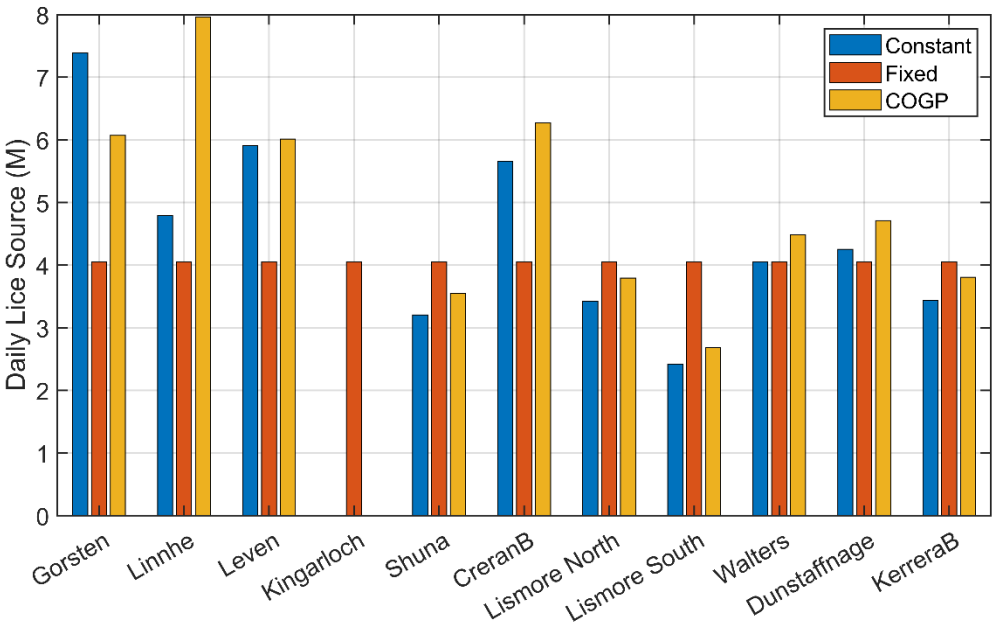


Figure 3.5.12. The “Constant”, “Fixed” and “COGP” sources of lice for the sensitivity analysis of source type.

The results from the source type simulations were mixed. When all sentinel cage locations were considered, the results from the “Variable” and “Constant” simulations produced slightly higher correlation coefficients (at the shorter radii of 50m and 100m, Figure 3.5.13). But when station No. 3 was excluded from the analysis, the “Fixed” source simulation performed better than all others. These results may possibly indicate that most of the sentinel cage locations were not particularly sensitive to individual sources of lice, but that station No. 3 was.

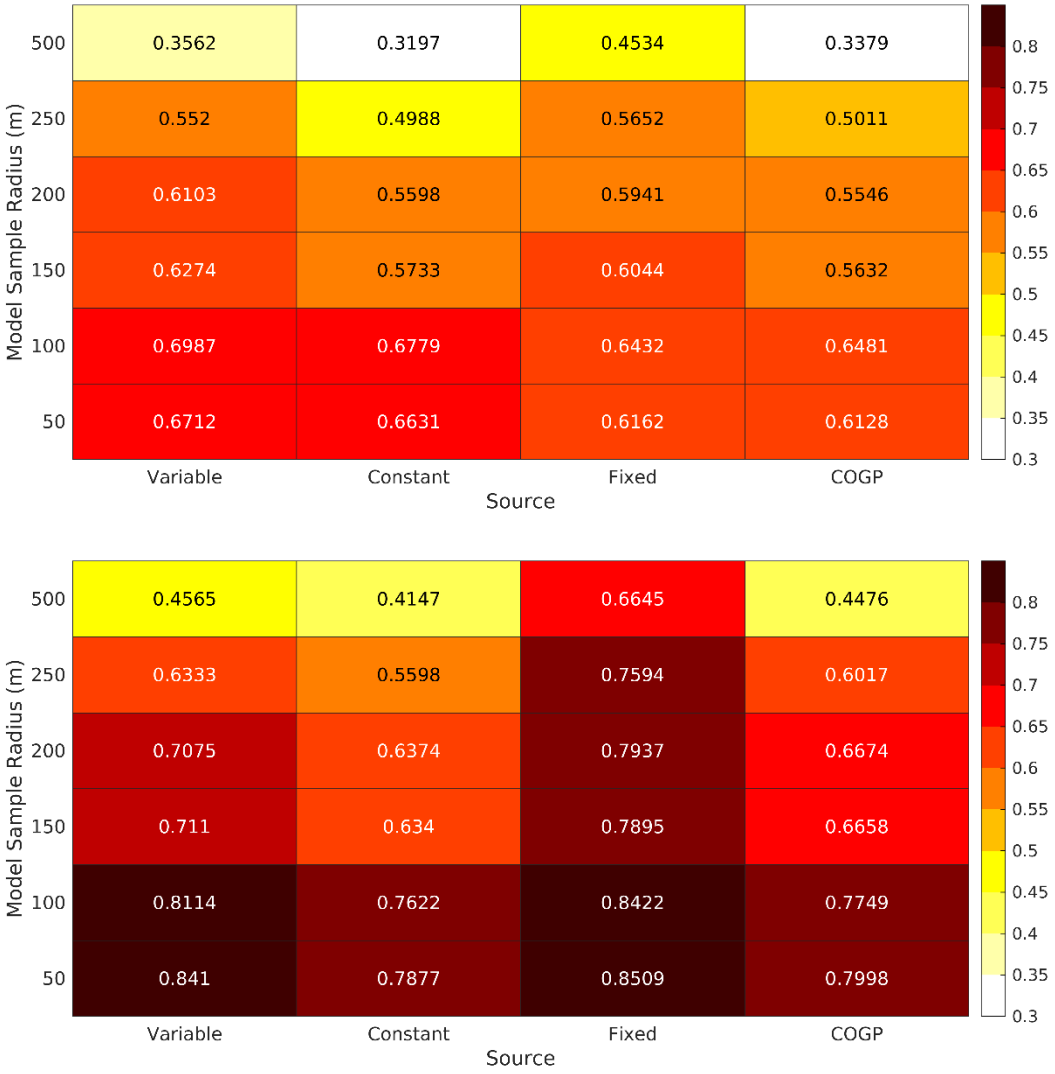


Figure 3.5.13 UnPTRACK; heatmap showing model skill (Pearson correlation coefficient) for a range of different data source types. Sites excluding No. 3 for deployment 2.

3.6 Lessons learned from sensitivity analysis

A comprehensive understanding of uncertainties in the projected distributions of salmon lice and associated impacts on wild fish is a key element in carrying out a robust assessment to allow the aquaculture industry to implement the most appropriate mitigation measures. Uncertainties arise from various sources in salmon lice model projections, such as structural differences in models, differences in initial conditions, different hydrodynamic scenarios, varying parameter values, output

resolution and bias correction. Further uncertainties arise in observational or field data, which can be difficult to quantify.

The one-at-a-time SA carried out here allows understanding how the model outputs which estimate sea lice density in space and time are related to and influenced by the input parameters tested above. The analysis carried out suggests the particle tracking models are not very sensitive to the number of particles released per source per hour. However, as the number is reduced below N_p = 10, more variance is introduced into individual simulations. Additional consideration is needed if inference on smaller spatial grids is required, here N_p = 10 was deemed appropriate for a calculation of densities on a spatial grid of 250 m^2.

The values tested for vertical diffusivity did not strongly impact model performance in this system, and a value of K_v=0.001 m s^{-1} is appropriate within this model setup. Here the interaction between vertical diffusivity and the various swimming behaviour implemented was not tested. Implementing swimming behaviour following Brooker et al, (2018), provided the best correlation with the field data, suggesting that these values are adequate for sea lice swimming behaviour implementation in Loch Linnhe. Including a user input parameter that recognises daylight hours for a given day of the year or incorporating surface irradiance information to drive a subsurface light intensity model that stimulates swimming/sinking behaviour (Johnsen et al., 2014) could potentially improve the model results. The source type influences the model performance in terms of how well the model correlates to the field data. Using the “Variable” option is likely helpful in understanding the actual sea lice distributions in space and time, however this information is only available after the lice counts have been recorded on the farm. The “COGP” source rates may be useful to managers/regulators in understanding the likely hotspots for lice in the water column should lice loads on all farms

reach these levels. This is helpful for planning, however inference on the actual lice loads is more useful in a research context.

3.7 Comparison of field data with particle tracking outputs

The three different models were individually validated through comparing modelled lice densities against plankton tow data and sentinel cage data, for various distances from the samples' location and a qualitative/quantitative model comparison was carried out for four time periods (Spring 2011, Autumn 2011, Spring 2013 and Autumn 2013). The following outputs were produced from each model; Map of mean lice density (lice km^{-2}) and plots showing comparison of modelled density to plankton tows at radii of 50m, 100m, 150m, 200m 250m, and 500m; Map of mean lice density (lice km^{-2}) and plots showing comparison of modelled density to sentinel cages at radii of 50m - 500m; Plots showing comparison of modelled infective pressure to sentinel cages at radii of 50m - 500m.

Key results for a subset of radii tested for each individual model are presented in the following sections 3.7.1 - 3.7.3. The sheer amount of data generated within this work package means it is not practical to include all figures within this report, Appendix 3 contains some of the additional plots. For each sample location the mean model lice densities were calculated using the lice in the top 2 m depth of the water column for the number of days the sentinel cages were deployed (see table 3.4.1 for details of dates). The modelled infective lice pressure shows the number of lice integrated over the number of days the sentinel cage was deployed.

3.7.1 BioTracker

The model comparison with mean lice abundance on sentinel cage for each time period from the coupled BioTracker- WLLS model are shown in figures 3.7.1.1 – 3.7.1.4.

In Spring 2011 the average number of lice per sentinel fish was between 0.02 and 0.92, the model infective lice density (See Appendix 3) and modelled infective lice pressure (Fig. 3.7.1.1) shown below do not show a strong linear correlation. This is not unexpected due to the signal-to-noise ratio, which limits the ability to detect weak signals (Silver 2012) as random variation grows relative to these signals. A similar pattern is seen in Spring 2013, with low average number of lice per sentinel fish and a poor fit with the linear regression with modelled data (Fig. 3.7.1.3). However the numbers of lice on fish are low, when compared to the autumn time periods (Fig. 3.7.1.2). Autumn 2013 shows a weaker linear relationship between field and modelled data (Fig. 3.7.1.4) than Autumn 2011. This suggests there may benefit from additional sensitivity analysis on each time period to better describe the relationships, and exploration of additional performance metrics as Pearson corelation is only one metric, with limitations.

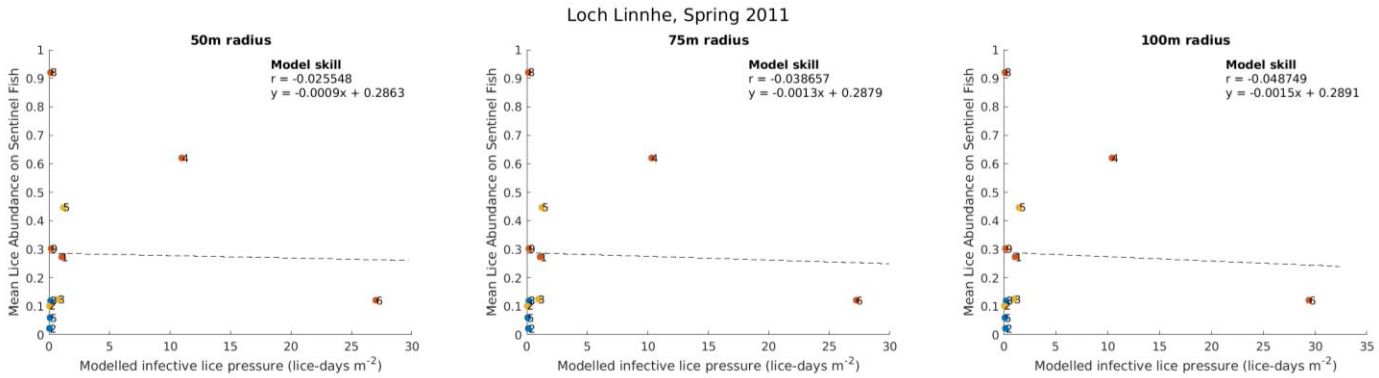


Figure 3.7.1.1: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the Biotracker PTM for Spring 2011, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployments 1 (●), 2 (●) and 3 (●).

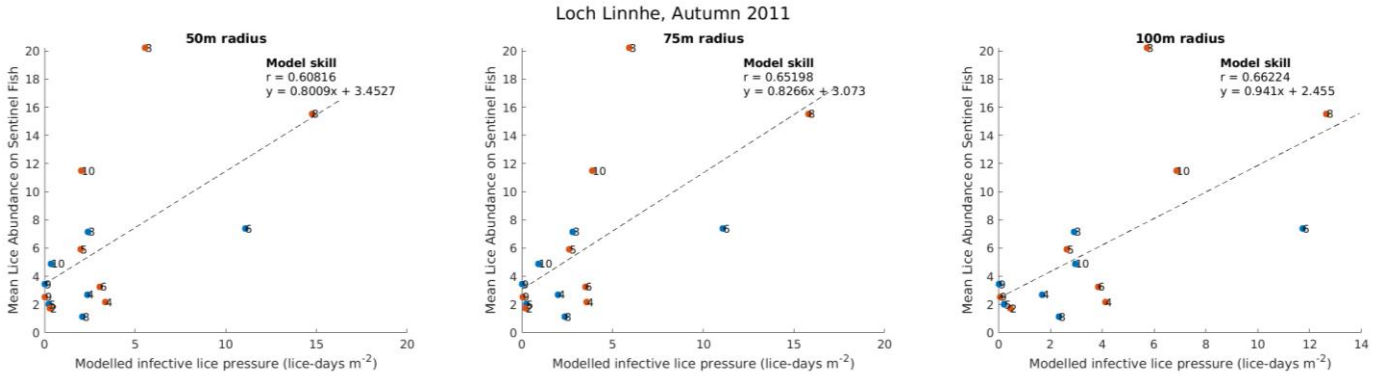


Figure 3.7.1.2: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the Biotracker PTM for Autumn 2011, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

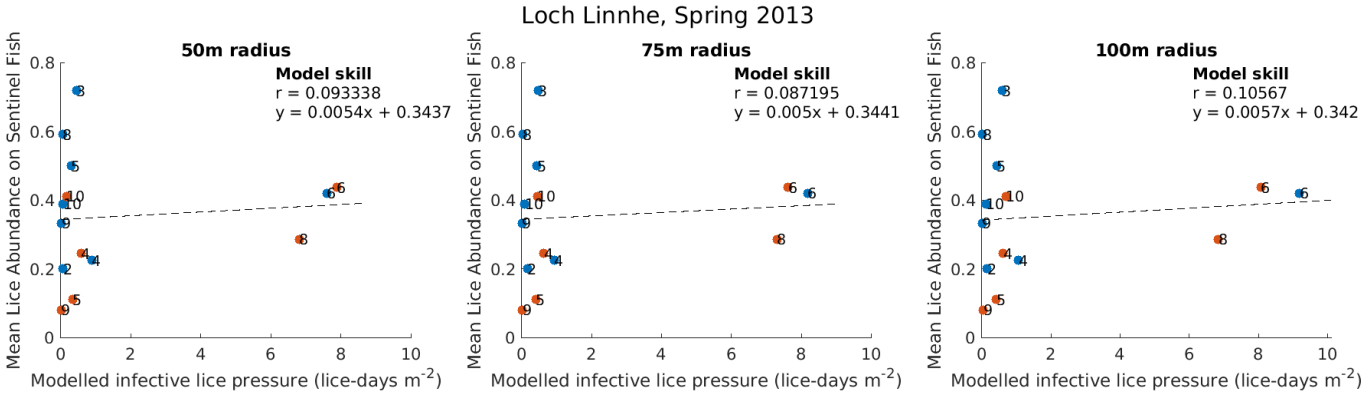


Figure 3.7.1.3: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the Biotracker PTM for Spring 2013, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

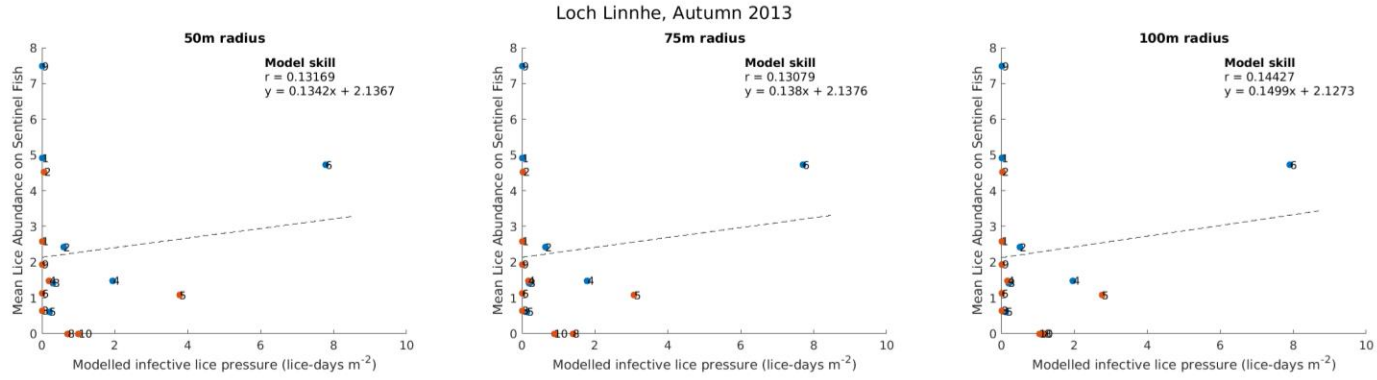


Figure 3.7.1.4: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the Biotracker PTM for Autumn 2013, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

3.7.2 FISCM

The model comparison with mean lice abundance on sentinel cage for each time period from the coupled FISCM- WLLS model are shown in Figures 3.7.2.1 – 3.7.2.4. The results from FISCM-WLLS model show a similar pattern to the Biotracker -WLLS, with Spring 2011 and Spring 2013 model infective lice density (Appendix 3) and modelled infective lice pressure (Fig. 3.7.2.1/3.7.2.3) shown below do not show a strong linear correlation. Autumn 2013 shows a weaker linear relationship between field and modelled data (Fig. 3.7.2.4) than Autumn 2011 (Figure 3.7.2.2).

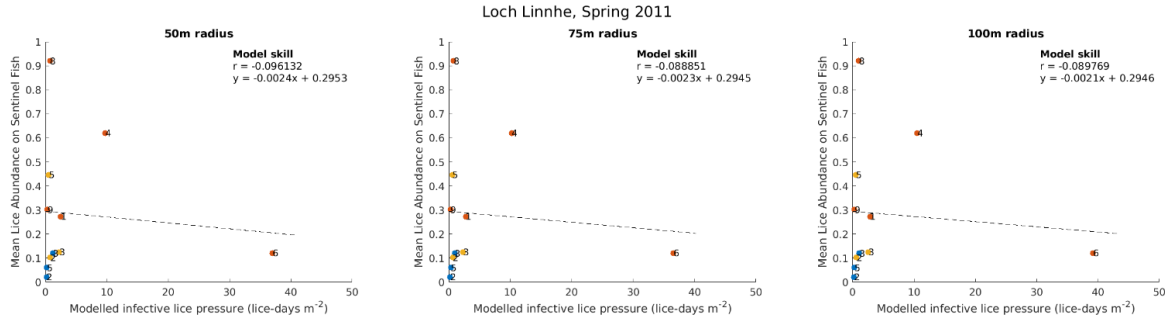


Figure 3.7.2.1 Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the FISCM PTM for Spring 2011, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployments 1 (●), 2 (●) and 3 (●).

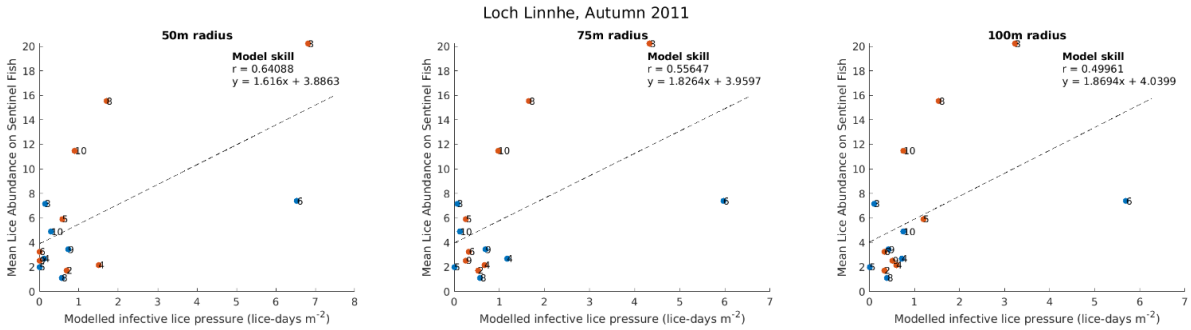


Figure 3.7.2.2: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the FISCM PTM for Autumn 2011, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

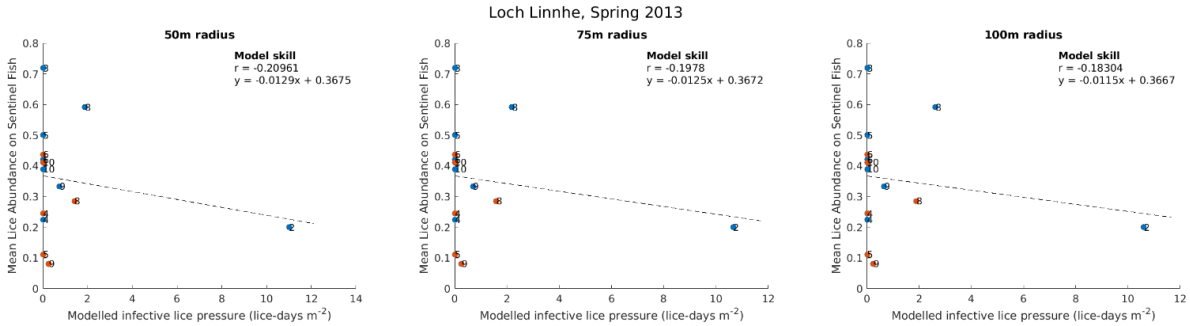


Figure 3.7.2.3 Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the FISCM PTM for Spring 2013, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

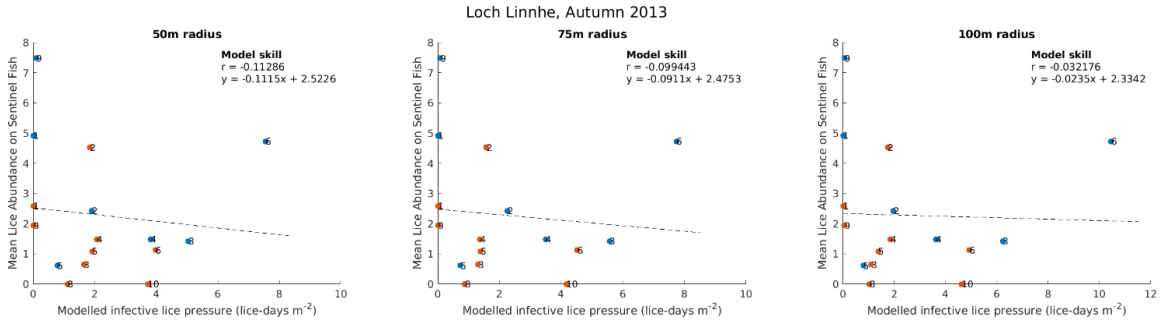


Figure 3.7.2.4: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the FISCM PTM for Autumn 2013, using 50m, 75m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

3.7.3 UnPTRACK

The model comparisons with mean lice abundance on sentinel cage for each time period from the coupled UnPTRACK - WLLS model are shown in Figures 3.7.3.1 – 3.7.3.4. The results from UnPTRACK-WLLS model show a similar pattern to the Biotracker -WLLS model and the FISCM- WLLS model, where the Spring 2011 and Spring 2013 model infective lice pressure (Fig 3.7.3.1/3.7.3.3) do not show a strong linear correlation. Autumn 2013 shows a weaker linear relationship between field and modelled data (Fig. 3.7.3.4) than Autumn 2011 (Figure 2.7.3.2). An additional plot 3.7.3.5 shows the combined inference of all four simulations, using a 50m radius circles to calculate modelled lice density.

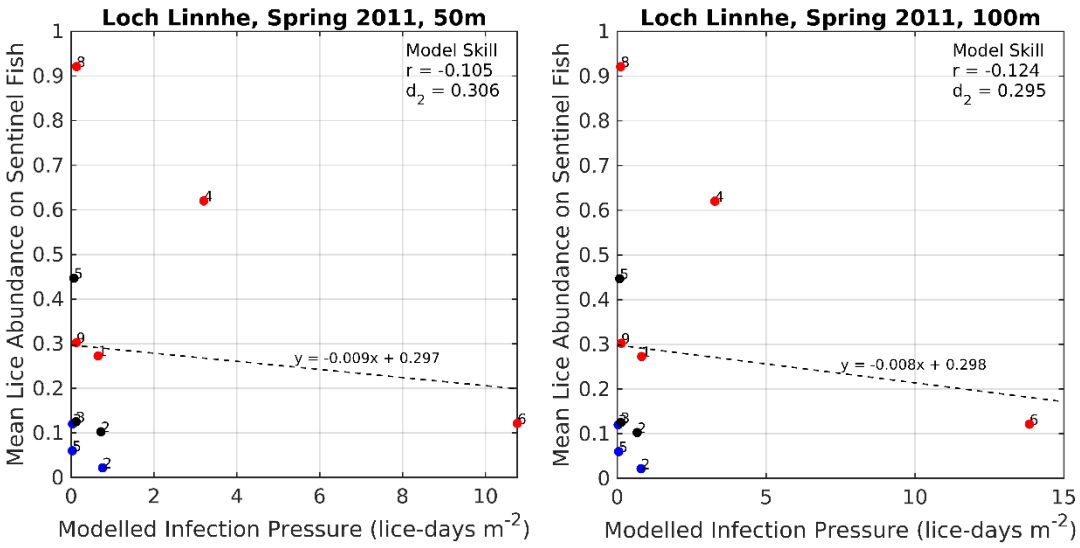


Figure 3.7.3.1. Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the UnPTRACK PTM for Spring 2011, using 50m and 100m radius circles to calculate modelled lice infective pressure for Deployments 1 (●), 2 (●) and 3 (●).

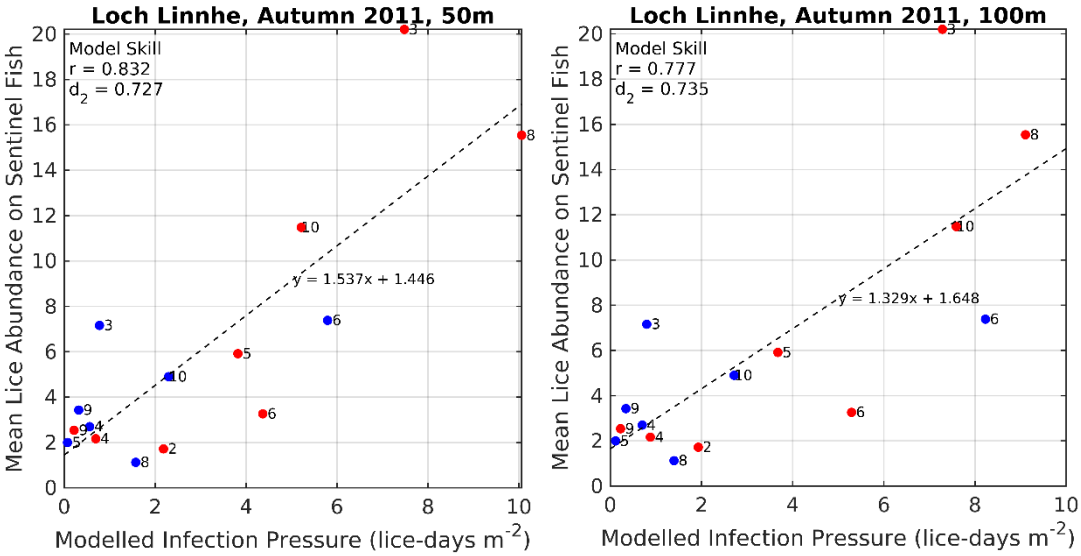


Figure 3.7.3.2: Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the UnPTRACK PTM for Autumn 2011, using 50m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

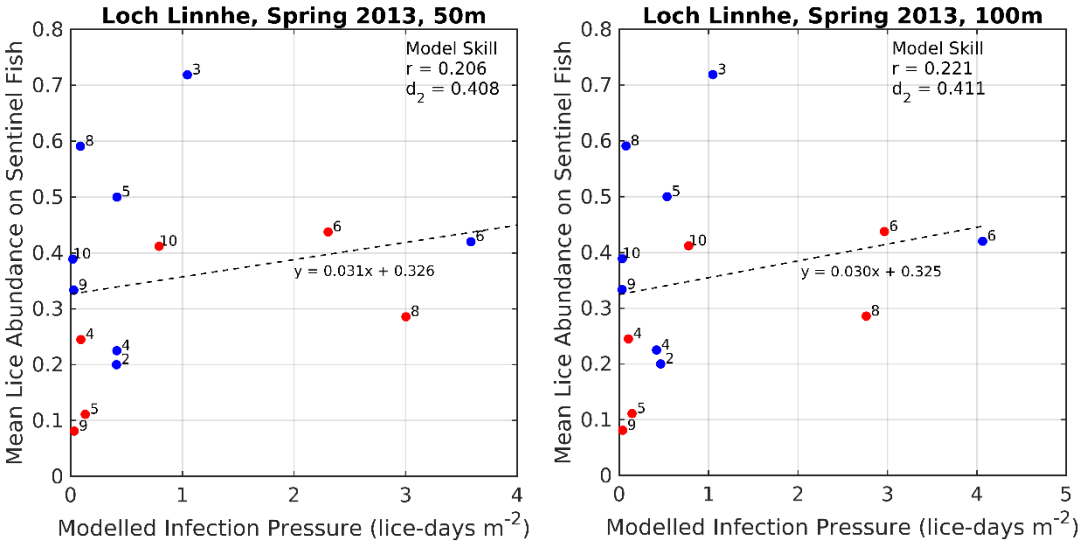


Figure 3.7.3.3. Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the UnPTRACK PTM for Spring 2013, using 50m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

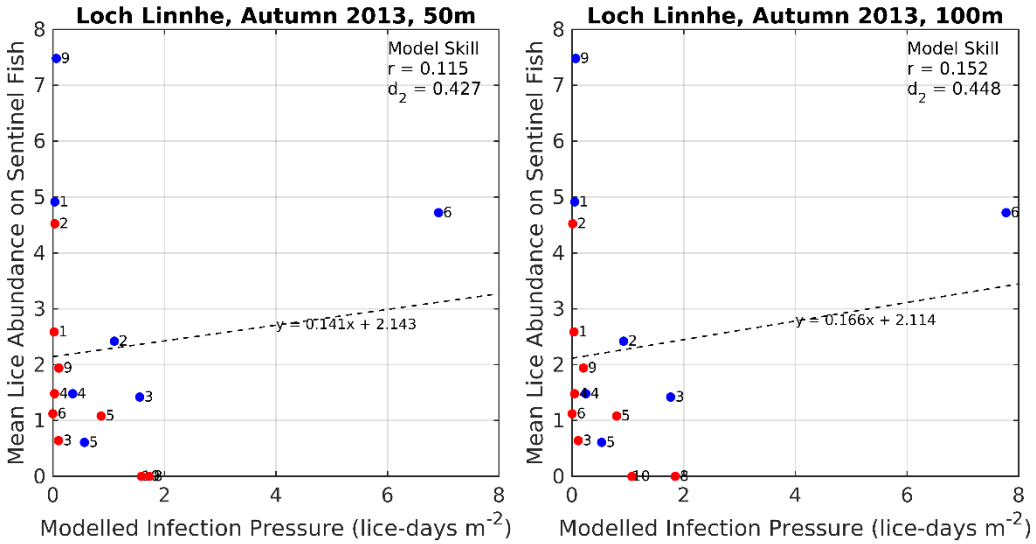


Figure 3.7.3.4. Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the UnPTRACK PTM for Autumn 2013, using 50m and 100m radius circles to calculate modelled lice infective pressure for Deployment 1 (●) and Deployment 2 (●).

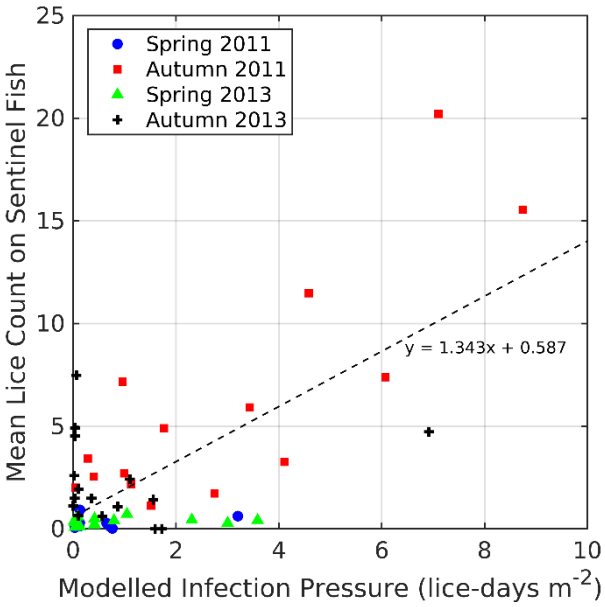


Figure 3.7.3.5. Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from the UnPTRACK PTM for all four simulations, using a 50m radius circles to calculate modelled pressure.

3.8 Summary of Individual Model Validation

1) Each model fitted data for lice infestation of salmon in most sentinel cages well in Autumn 2011, the time period for which sensitivity analysis was undertaken. This suggests that further sensitivity analysis should be undertaken to investigate if any sea lice dispersal model parameter fits the data better.

2) Fits depended on signal:noise ratios and therefore best fits were when lice on farms was highest. Targeted future sampling in areas that tend to have higher lice numbers on farms may be beneficial for further model development.

3) This validation enables the application of particle tracking models as tools for management of risk of lice from salmon farms to be considered with varying degrees of confidence depending on model fit.

4) Given multiple models exist, it is necessary either to select the best model or combine all or a sub-set of the models as an ensemble to assess model uncertainty for application in management.

The next section explores how the models can be combined in an ensemble and how this process can allow visualisation of how closely models match the predictions of each other across space in Loch Linnhe.

4. Ensemble Modelling

Robust assessments of sea lice in the environment are important for assessing the scale of lice management required on farms, for estimating the impact on wild fish populations, and for understanding transfer of lice among farms. Hydrodynamic Ocean Circulation Models coupled with Lagrangian particle tracking models are an important tool in our understanding of sea lice dispersal from fish farms. Sea lice dispersal models have been developed in Norway (Sandvik et al. 2016), Scotland (Salama et al. 2018), Canada (Nelson et al. 2018), Ireland (Costelloe et al. 2009) and the Faroe Islands (Kragesteen et al. 2018). Further to these nationally developed models, a standardised framework has been agreed for sea lice dispersal modelling through international collaboration (Murray et al. 2022). Some direct measures of sea lice are available such as lice counts from farms. Given challenges in collecting relevant empirical information on distributions of lice in the environment, work to date has relied on a numerical modelling approach. The availability of multiple models necessitates an evaluation of the implications of how they differ in predicting lice distributions across space and time.

In this part of the SPILLS study, the uncertainty in model outputs for sea lice dispersal modelling is investigated among a suite of 3 particle tracking models coupled with the Wider Loch Linnhe System hydrodynamic model to assess model variability. This is important as an understanding of uncertainties in the modelled sea lice distributions with respect to sentinel cage observations, is a key element in carrying out a robust assessment. Thus, in order to gauge the amount of trust we can have in our models we measure the types of uncertainty within our model predictions.

For this first usage of ensemble modelling in Scottish aquaculture, we consider simple ways to describe the types of uncertainty to present final outputs in a succinct and intuitive way. Average and weighted ensemble models using these individual models are calculated to assess ensemble model performance and uncertainty.

4.1 Ensemble Model Development

Ensemble modelling is a technique that has the potential to decrease bias and variance of predictions from individual models, through combining the predictions of different models (Fig. 4.1.1). The approach is used in fields such as climate (Tebaldi and Knutti 2007), fisheries management (Jardim et al. 2021), and epidemiological (Oidtman et al. 2021) modelling to improve understanding and management of uncertainty. The inference from a simple ensemble prediction, where all models are weighted equally and mean weighted ensemble prediction, where models are weighted based on their fit with the observational data sets are both investigated as a proof of concept.

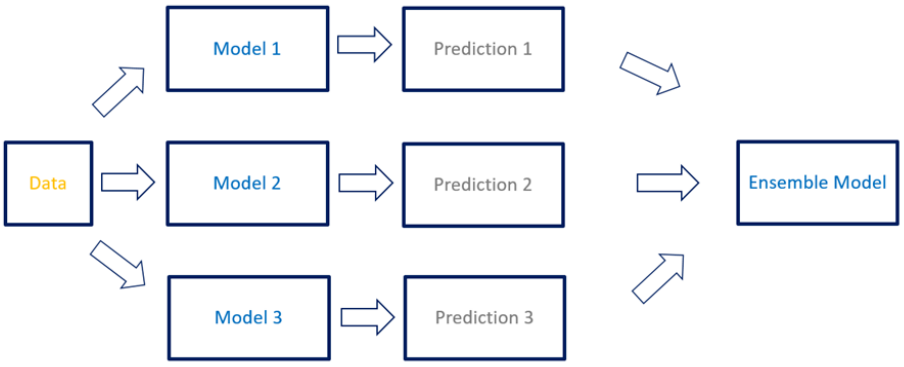


Figure 4.1.1: Visualisation showing how models come together to form an ensemble model.

To develop the ensemble methods the focus below is on one time period, Autumn 2011, for modelled infective pressure calculated at a 100 m radius around the sentinel cages (Figure 4.1.2) these figures are akin to the prediction 1, 2 and 3 shown in figure 4.1.1. This time period was chosen to

help to develop the approach, as it was the time period where the OATSA (One At a Time Sensitivity Analysis) was carried out, chosen *a priori* based on the higher mean number of lice per sentinel fish and therefore displaying a greater signal-to-noise ratio which minimises the influence of random variability. Confidence in the both the hydrodynamic forcing and PTM performance for the Autumn 2011 deployment is high, due to the time spent on the validation and sensitivity analysis for this time period. Here, the three models underwent varying levels of scrutiny through the OATSA, all three predict a similar range of values at the sentinel cage locations. Additionally, in this instance, the underlying hydrodynamics are the same in all three models, so differences between models are due to the particle tracking models only. Thus, in this case, we have the opportunity to assess if different particle tracking models cause spatial variability in the predicted sea lice distributions. This approach should be expanded to investigate other time periods, where practical. Table 4.1.1 shows the parameter values used in each model considered within the ensemble. We use the model with the highest Pearson correlation values at 100m radius as developed from each organisation at the time of reporting, using the release of 10 particles per source per hour, according to the sensitivity analysis undertaken in section 3.5. Further model developments and calibration may lead to changes in these parameter values in Table 4.1.1, so these should not be considered prescriptive, rather, used as a starting point for future work. It is important to understand that the model outputs in lice – days per m^2 while observational data are mean lice counts on fish, this leads to variation in the relationships to observations between outputs from different models, thus further analysis would be beneficial.

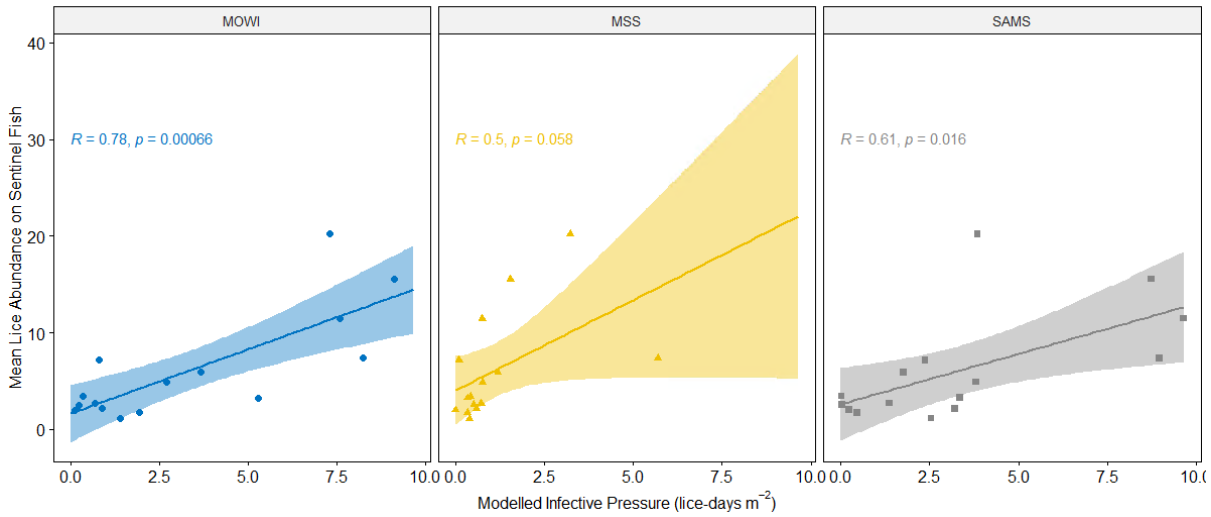


Figure 4.1.2 Comparison of modelled infective lice pressure with mean lice abundance on sentinel fish from each of the three PTM at a 100m radius around the cages for Autumn 2011. 95% confidence intervals indicate confidence in position of the linear regression.(See section 3.7.1 -3.7.3 for more details on different radius or time periods).

Table 4.1.1 : Summary of setup parameters in each dispersal (particle tracking) model used in development of ensemble approach.

Parameter	Biotracker Value	Fiscm Value	UnPTRACK Value
Horizontal Diffusivity, K_H (m^2 s^{-1})	0.1	0	0.1
Vertical Diffusivity, K_V (m^2 s^{-1})	-	-	Variable
Initial Spread (m)	0	0	50
Source Type	Variable (weekly lice counts)	Variable (weekly lice counts)	Variable (weekly lice counts)
Swimming Behaviour (cm s^{-1})	-	All particles: 1.8 cm s^{-1}	Nauplii: 1.25 cm s^{-1} Copepodid: 2.14 cm s^{-1} 6am – 6pm
Sinking Behaviour (cm s^{-1})	-	All particles: 1.8 cm s^{-1}	Nauplii: 0.09 cm s^{-1}

			Copepodid: 0.10 cm s^{-1} 6pm – 6am
Low-Salinity Avoidance Threshold (PSU)	-	-	20
Maturation rate	86 hrs	86 hrs	40 degree-days
Time step (s)	120 s	30 s	600 s (with sub-stepping when required)
Particle source rate (Np per farm per hour)	10	10	10
Mortality constant (h^{-1})	0.01	0.01	0.01
Depth Limit (m)	1	20	50
Settlement time	336 hrs	336 hrs	170 degree-days

Relationships between predicted lice in the water column (infective pressure, lice-days.m^{-2}) in 100m radius around sentinel cages and actual mean number on fish held in sentinel cages are compared for the three different particle tracking models in Fig 4.1.2 (note the different axes scales). The UnPTRACK and BioTracker models are broadly similar whereas the slope of the FISCM model may be steeper, such that if the mean alone is used, it tends to be more conservative- predicting lower infective pressures than the other two models. The FISCM model is non-significant and has a relatively low correlation coefficient. However, it is notable that a single cage appears to be an outlier that tends to reduce the slope, resulting in an uneven distribution of residuals at the lower end of the regression fit, this reduces the correlation coefficient. For FISCM particularly, further work is recommended to explore potential causes of such outlying points, additional modelling is required to fully incorporate the lessons learned from the sensitivity analysis. Additional calibration of the FISCM modelling algorithm is also needed, as this model was originally designed for a different particle tracking application in a different system. Model calibration and validation are key steps which should be undertaken prior to using any particle tracking model as a decision making tool.

Two versions of the ensemble model were calculated using the predictions shown in Figure 4.1.2. The first was a simple mean, in this version all models were weighted equally, the second was a weighted mean, where each model was weighted according to the variance explained by the modelled (r value). Figure 4.1.3 shows the normalized values for these ensembles, these values are then back calculated to reflect the mean range of values within the models (min= 0.047, max = 8.15). Here the two versions of the ensemble models are showing a very similar range and very similar results. Where this becomes important is when models present very different ranges within a system and decision makers need to understand the potential consequence of using one model over another. The mean weighted ensemble approach provides a marginally better R value and p-value than the simple mean model.

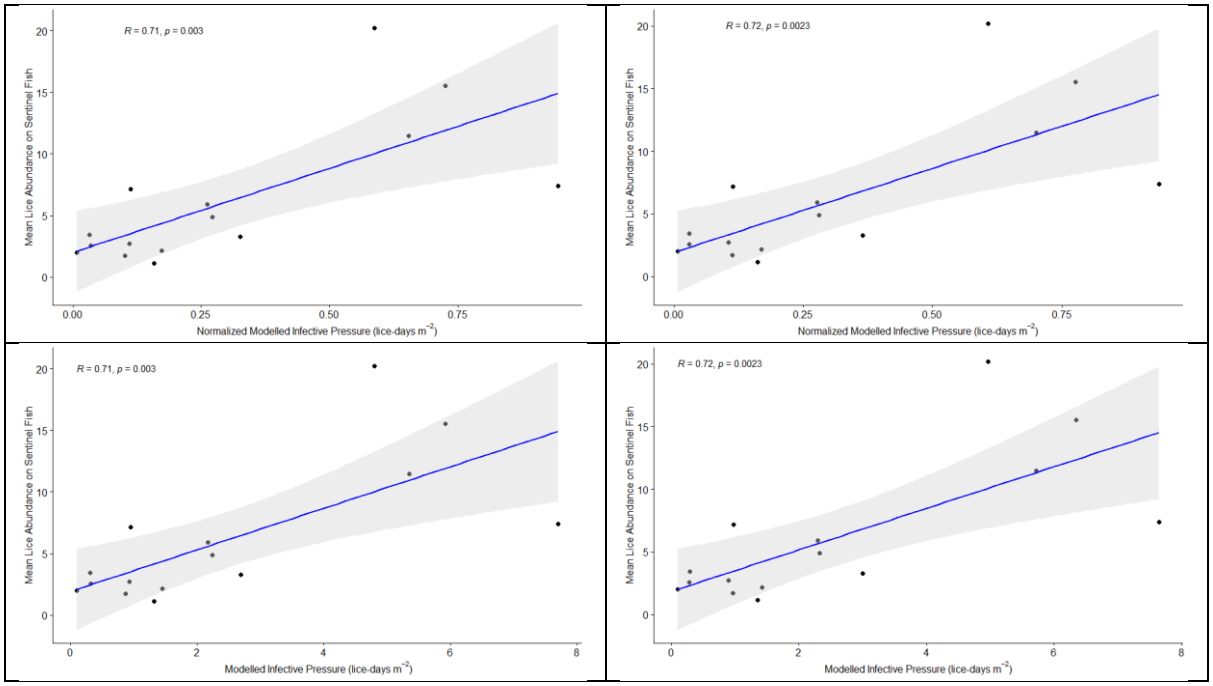


Figure 4.1.3 Calculating a simple mean ensemble and a weighted mean ensemble using predicted values shows in figure 4.1.2 by normalizing the data to calculate the mean values first and then translating that into a predicted value for Autumn 2011. 95% confidence intervals around the linear regression (blue line) shown in grey.

4.2 Evaluation of Ensemble Approach

Visualising the variance around the mean values in the ensemble is one way to think about the uncertainty for decision making. Maps showing the predicted density or prevalence of particles are the output of interest from the bio-physical modelling exercise for managers and scientists. An assessment of similarity in the spatial predications between the mean ensemble model prediction and the three individual model predictions allows us to assess the spatial uncertainty in the predictions.

The uncertainty from the three coupled hydrodynamic-dispersal models were assessed by (i) comparing predicted lice densities/infection-pressure generated by each model with relevant field data (see section 3.7 and Figure 4.2.1), (ii) individual model comparisons with ensemble model output (Figure 4.2.1) and (iii) comparing the individual maps of infective pressure to the ensemble map of infective pressure (Figure 4.2.3).

In the evaluation of model fit, the average lice abundance on sentinel fish per cage per deployment is used. In each cage the aim was to sample 50 fish, in reality between 20 and 50 fish could be sampled due to avian and seal predation. The variance around the mean in each cage for each deployment is not represented in the evaluation of model fit shown in Figure 4.1.3 To provide context the field data showing the number of lice counted on each individual sentinel fish in each cage for each deployment is plotted beside the models predicted values and the confidence intervals (Figure 4.2.1).

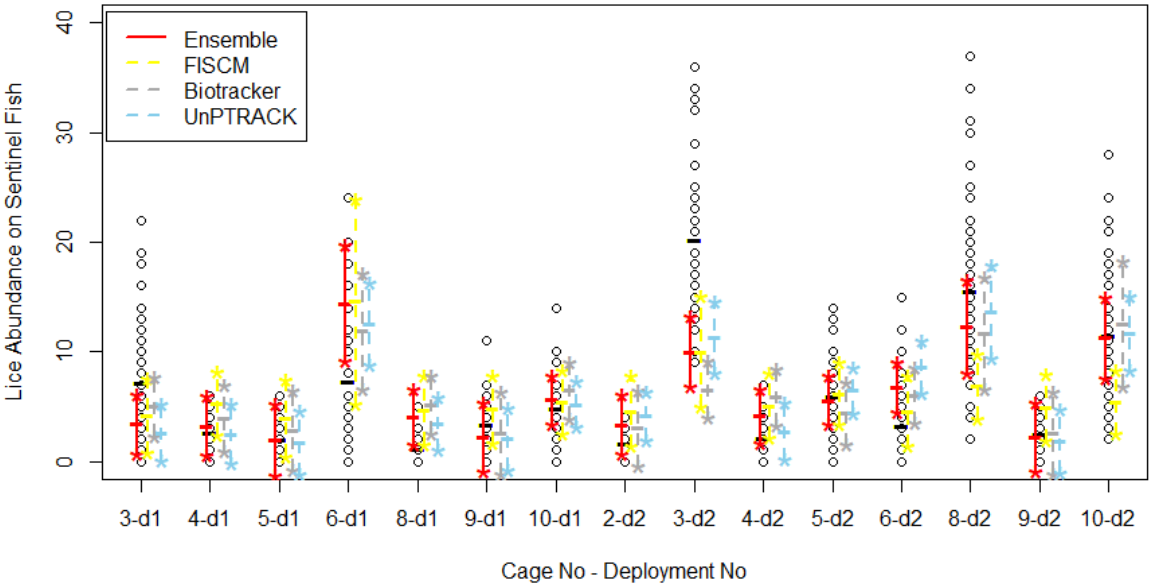


Figure 4.2.1 Qualitative comparison of PTM outputs at a 100m radius around the cage with field data for the deployment in Autumn 2011. Mean lice abundance on sentinel fish represented with black horizontal line. The number of lice counted on each individual sentinel fish represented by black open circles. The predicted mean lice abundance on sentinel fish is represented by small coloured horizontal line and upper and lower 95% confidence interval values show with *Mean ensemble results are shown in red, UnPTRACK model prediction results are in blue, FISCM model prediction in yellow, and BioTracker model prediction in grey. Each individual model is fitted using its unique relationship between lice in the water column and lice on sentinel fish illustrated in Figure 4.1.2. The mean and weighted-mean fits are from models specified in Figure 4.1.3.

The ensemble prediction models are doing a good job at predicting the lice abundance on sentinel fish at many of the sentinel cage locations, highlighting that generally we are able to represent lice in the environment. For cage 3, deployment 2 the models all underestimated the number of lice on sentinel fish, while at cage 9, deployment 2 the Mean Ensemble, UnPTRACK model prediction and BioTracker model prediction model overestimated the number of lice on sentinel fish. This highlights that sentinel cages in highly dynamic areas may not always be well represented by any model. For each of these model outputs the relationship between infective pressure lice – days per m^2 and lice counts on sentinel fish are different. And so the models all replicate the pattern of variation in sentinel

cage observations even though they vary in predictions of infective pressure.

Figure 4.2.1 highlights the variance around mean in the field data is well represented in the models. This suggests that the models are generally well calibrated. This, coupled with the sensitivity analysis undertaken, gives a good understanding of how the parameters individually impact the model outputs. The sea lice dispersal models individually are able to predict the distribution of sea lice in the environment. The ensemble as presented here helps to contextualise the level of variability in the different models, providing understanding of areas where the models are not in agreement. This approach becomes important when validation against field data is not possible, for example for future predictions or areas where data are not available. In this case, as the models do provide similar fits to sentinel data, this should lead to higher confidence in decision making. Conversely if all three models provided very different ranges and there were no data available to validate them, then decision makers should have lower confidence in the models.

In order to assess the geographical consistency between pairs of maps, a Pearson's correlation metric was applied. This is a spatially explicit metric comparing the local correlation coefficient between two maps using a focal neighbourhood search (n=5). The “corLocal” function from the “raster” package in R (Hijmans 2020) was used. The closer correlation coefficients get to -1.0 or 1.0, the stronger the agreement between the two individual data sets. This helps in understanding where in space differences occur in model outputs. Figure 4.2.2 and 4.2.3 show areas where all models have similar predictions to the ensemble mean values (green to blue areas), thus giving higher confidence in the models predictive ability. Conversely, there are areas where models differ significantly from the ensemble mean values (yellow to red), in these there is less confidence

in the individual models predictive ability. Understanding the variation in geographical consistency between models output may help to inform future sampling strategies.

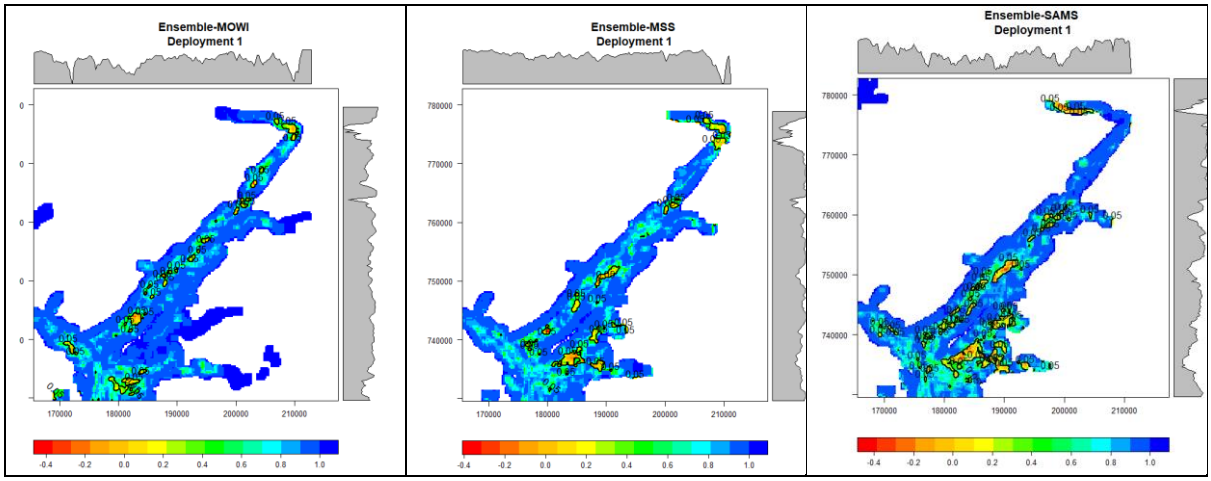


Figure 4.2.2 Deployment 1 showing each individual model geographical consistency with mean ensemble. The pearson's correlation provides an estimate of the strength of the linear relationship between two variables. Correlation coefficients range from -1.0 (a perfect negative correlation) to positive 1.0 (a perfect positive correlation). Values of p = > 0.05 are generally considered significant, these areas are outlined in black. Appendix 4 provides larger versions of these plots.

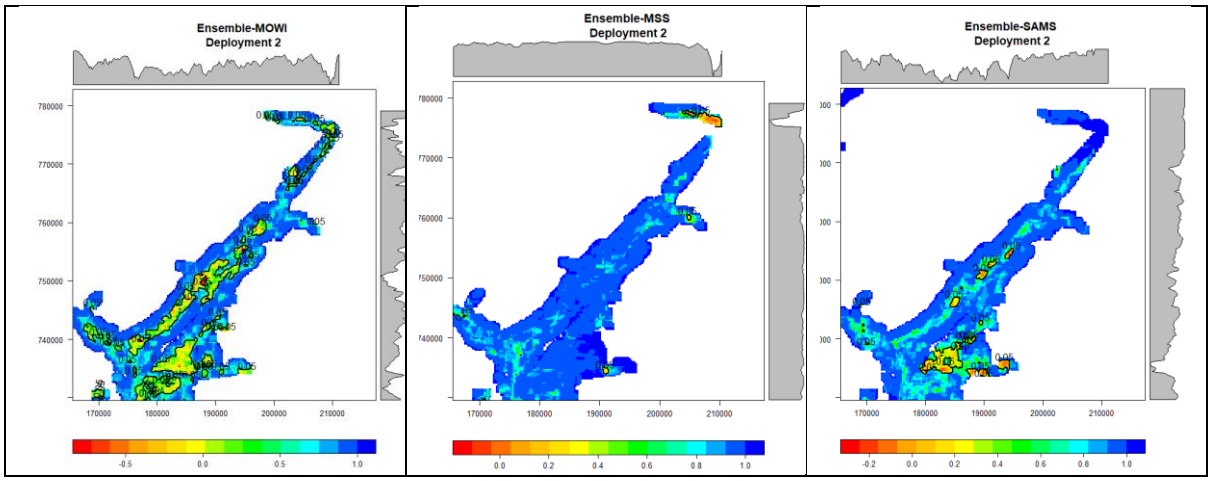


Figure 4.2.3 Deployment 2 each individual model geographical consistency with mean ensemble. The pearson's correlation provides an estimate of the strength of the linear relationship between two variables. Correlation coefficients range from -1.0 (a perfect negative correlation) to positive 1.0 (a perfect positive correlation). Values of p = > 0.05 are generally considered significant, these areas are outlined in black. Appendix 4 provides larger versions of these plots.

Each of the dispersal models has been developed by different organisations and each has its own assumptions and limitations. All models have used the same oceanographic forcing. However, the biological conditions to simulate their sea lice dispersal rates differ. This allows direct comparison of each PTM's results on the sea lice survival and movement within the study area. Using Figure 4.2.1 we can assess the range of values predicated by the ensemble and assess what that means in relation to each individual PTM. Figure 4.2.2 and 4.2.3 can be used to further infer areas of high and low agreement in space, for given time periods, here averaged over the sentinel cage deployment period.

4.3 Lessons learned from applying ensemble approach

Quantifying model uncertainty using an ensemble approach is possible. We have multiple models which have given similar fits to sentinel cage data for the same scenario. These models all have reasonable parameter estimates. Each model has predicted the observed data well (Figure 4.2.2), this is promising given the uncertainty in the source data.

Models always need refinement, and these sea lice dispersal models would benefit from further validation with contemporary data. This is a reason to collect more field data in an open and systematic way, to better support sustainable growth of the aquaculture industry.

There is measurement error in the sentinel cage data, and there is some inherent randomness in the system that we have not fully quantified. There is some natural variation in space and time. Combining results from the sensitivity analysis (i.e. using this information to infer if models are likely to be well calibrated) and the ensemble results (i.e. do models give similar results) with expert opinion (i.e. consider the quantitative rigour of our models, individually and as an ensemble) and insight about other potentially relevant factors outside of the models (historical data, unknowns about the source lice etc.) to make an assessment of model

selection is necessary, particularly in the short term, as a data set for validation is built up in Scottish waters.

5. Conclusions and Next Steps

5.1 Hydrodynamic modelling for sea lice particle tracking

- The quality and resolution of the underlying hydrodynamic model is the singularly most important component of a particle-tracking model. Different parameterisations of the underlying HDM resulted in different fits to observational oceanographic data. The version finally adopted was the one which presented the best bit of field measurements.

5.2 Sea lice particle tracking models

- The importance of calibrating the different PTMs with some simple test case prior to the scenarios simulation, to make sure the basic model performance is good, should not be underestimated.
- Careful consideration of the specificity of each PTM and how they are suitable for sea lice modelling should be made:
    - Land boundary conditions (restoring behaviour vs reflecting behaviour vs sub-timestep method when nearing land boundary), which one is the best for sea lice modelling needs to be explored.
    - Spherical vs cartesian coordinates – cartesian coordinates are more suitable for the SPILLS spatial scales whereas spherical coordinates are acceptable for larger scale simulations.

- Fiscm performance is the lowest of all models, a lot has been learnt on how it is coded and some things have already been improved during SPILLS. Some others identified problems are:

    - Land boundary handling needs to be revised

- Extra care needs to be taken to know what forcing variables are given as input of the particle tracking simulation. After the Fiscm simulation were performed and the data post processed, we realised that Fiscm was not reading the same time variable as Biotracker and UnPTRACK from the hydrodynamic model. The time variable read by Fiscm was not as precise as the one used by Biotracker and UnPTRACK, so Fiscm did not consider the velocity field forcing at the exact same time as Biotracker and UnPTRACK were doing, which would have an impact on the particle trajectories, but the exact magnitude of the has not yet been quantified.
- Some difference between models in the handling of the vertical dimension has been identified. These will likely make a difference when doing the interpolation of the hydrodynamic velocity forcing to compute the particle trajectory.
- Fiscm Runge Kutta solver uses sigma coordinates, if the particles' initial position is given with z-coordinates they will be converted to sigma coordinates. On the contrary, UnPTRACK can do the computation with either z-coordinates or sigma coordinates depending on the coordinates chosen for the particles' initial position.
- Fiscm uses z-coordinates with z=0 at mean sea level while UnPTRACK set z=0 at the moving free sea surface.

5.3 Using an ensemble model approach to quantify model uncertainty

Within SPILLS a sensitivity analysis was carried out. This has improved the understanding of important modelling parameters and highlighted areas where further research could further reduce uncertainty and thus aid decision making. The ensemble model was developed by

selecting the best fit research models from each PTM. These 3 models predicted 3 different “modelled infective pressure (lice days m^{-2})” ranges, with 3 different linear equations representing the relationship between the modelled infective pressure and the number of lice on sentinel cage fish. This provides a measure of uncertainty around the model performance for the Autumn 2011 time period.

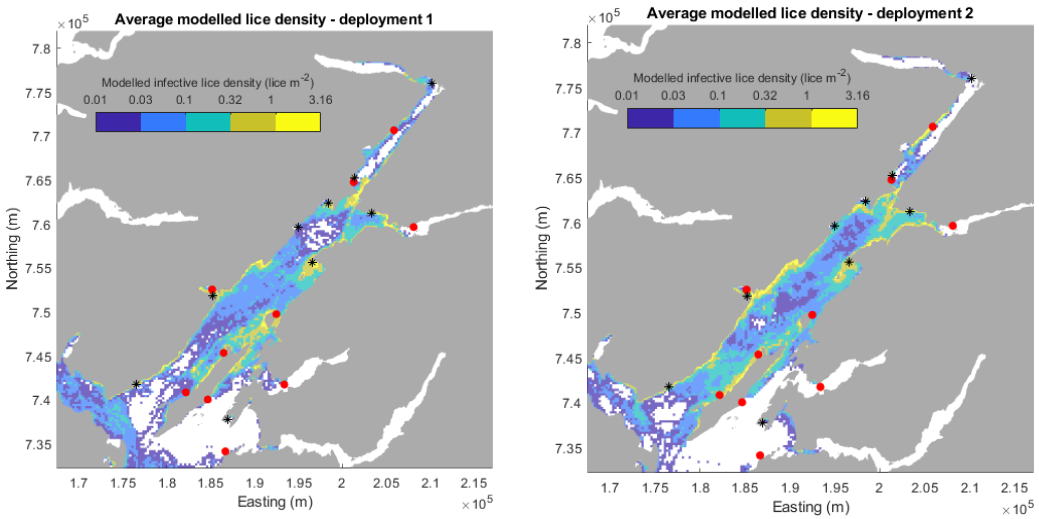


Figure 5.2.1: Example of modelled infective lice density from the Biotracker PTM for Autumn 2011 (see Table 3.4.1 for dates).

The purpose of sea lice dispersal models is to represent the dispersal of sea lice from sources of origin as accurately as necessary for a given criteria. Every modelling decision involves aspects of aggregation and exclusions. Generally, aggregations are required to provide a simplified representation of a complex real-world problem, while exclusions can be used to omit a portion not important to the modelling objectives. For example, within SPILLS the inclusion of linking a population model to our sea lice dispersal models was not explored, as the primary objectives were to calibrate and validate our PTMs. In the context of sea lice modelling, aggregations over various spatial and temporal scales are necessary to represent a manageable view of the highly dynamic (in time and space) sea lice infestation pressure. Here figures representing modelled lice infective pressure over timescales relevant to sentinel deployments (generally

averaged over 1 week, e.g., figure 5.2.1) have been produced. Additionally, maps showing predicted lice density over a 12-hour period to compare with our plankton tow data were drawn. These average plots do not allow the viewer to develop an understanding of the patchy distributions of sea lice in the environment, so videos showing hourly locations of sea lice have also been produced.

Using plots showing geographical consistency (Figs 4.2.3 & 4.2.4) along with density (e.g. Fig. 5.2.1) or prevalence maps we can begin to answer the following questions:

i. Where do we have high confidence that lice persist because prevalence of lice is high across all models? Where do we have high confidence that lice do not persist because prevalence of lice is low across all models?

ii. Where do we have high confidence that high densities of lice persist (i.e., density is consistently high across all models)?

iii. Where do we have high confidence that low densities of lice persist?

iv. What areas in the models show inconsistent results, thus low confidence in the models (where do we have low agreement between model runs)?

This work represents the first step in developing ensemble modelling for describing sea lice distributions. More work is needed to develop this ensemble approach for management purposes, but this report and the lessons learned through SPILLS will help to drive the next steps.

6. References

Adams T (2019) BioTracker - biological particle tracking in unstructured and structured hydrodynamic grids. Scottish Association for Marine Science, Oban, UK; BioTracker

Adams T, Aleynik D, Black K (2016) Temporal variability in sea lice population connectivity and implications for regional management protocols. Aquaculture Environment Interactions 8:585–596

Adams T, Black K, MacIntyre C, MacIntyre I, Dean R (2012) Connectivity modelling and network analysis of sea lice infection in Loch Fyne, west coast of Scotland. Aquaculture Environment Interactions 3:51–63

Adams T, Brain S (2021) Preliminary report on Loch Linnhe dispersal model simulations setup using Biotracker and SSM

Adams TP, Aleynik D, Burrows MT (2014) Larval dispersal of intertidal organisms and the influence of coastline geography. Ecography 37:698–710

Aleynik D (2020) SCOATS: Scottish Coastal Ocean and ATmospheric Modelling Service. Scottish Association for Marine Science, Oban, UK; SCOATS

Aleynik D, Dale AC, Porter M, Davidson K (2016) A high resolution hydrodynamic model system suitable for novel harmful algal bloom modelling in areas of complex coastline and topography. Harmful Algae 53:102–117

Amundrud, TL, & Murray, AG (2009) Modelling sea lice dispersion under varying environmental forcing in a Scottish sea loch. Journal of fish diseases, 32(1), 27-44.

Bell, VA, Kay, AL, Jones, RG, & Moore, RJ (2007) Development of a high resolution grid-based river flow model for use with regional climate model output. Hydrology and Earth System Sciences. https://doi.org/10.5194/hess-11-532-2007

Brickman, D, Ådlandsvik, B, Thygesen, U, Parada, C, Rose, K, Hermann, A, and Edwards, K, (2009) Particle tracking. Manual of Recommended Practices for Modelling Physical-biological Interactions in Fish Early-life History, 295, p.14e31.

Brooker, AJ, Skern-Mauritzen, R, & Bron, JE (2018) Production, mortality, and infectivity of planktonic larval sea lice, *Lepeophtheirus salmonis* (Krøyer, 1837): current knowledge and implications for epidemiological modelling. ICES Journal of Marine Science, 75(4), 1214-1234.

Cantrell DL, Filgueira R, Revie CW, Rees EE, Vanderstichel R, Guo M, Foreman MGG, Wan D, Grant J (2019) The relevance of larval biology on spatiotemporal patterns of pathogen connectivity among open-marine salmon farms. Can J Fish Aquat Sci. 77: 505-519 cjfas-2019-0040. https://doi.org/10.1139/cjfas-2019-0040

Chen C, Beardsley RC, Cowles G, Qi J, Lai Z, Gao G, Stuebe D, Xu Q, Xue P, Ge J, Hu S, Ji R, Tian R, Huang H, Wu L, Lin H, Sun Y, Zhao L (2013) An unstructured grid, finite-volume community Ocean Model FVCOM user manual https://wiki.fvcom.pml.ac.uk/lib/exe/fetch.php?media=source:fvcom_small.pdf

Chen, C, Liu, H, and Beardsley, RC (2003) An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries. J. Atmospheric Oceanic Technol. 20, 159–186.

Cole, SJ, & Moore, RJ. (2009) Distributed hydrological modelling using weather radar in gauged and ungauged basins. Advances in Water Resources. https://doi.org/10.1016/j.advwatres.2009.01.006

Costelloe M, Costelloe J, O'Donohoe G, Coghlan NJ, Oonk M, Van Der Heijden (2009) Planktonic Distribution of Sea Lice Larvae, *Lepeophtheirus Salmonis*, in Killary Harbour, West Coast of Ireland. Journal of the Marine Biological Association of the UK 78, 853-874

Dee, DP, Uppala, SM, Simmons, AJ, Berrisford, P, Poli, P, Kobayashi, S, et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. https://doi.org/10.1002/qj.828

Edwards, KP, Barciela, R, & Butenschön, M. (2012). Validation of the NEMO-ERSEM operational ecosystem model for the North West European Continental Shelf. Ocean Sci. https://doi.org/10.5194/os-8-983-2012

Gillibrand, PA. (2022); UnPTRACK: A multi-purpose particle tracking model for unstructured grids. Mowi Scotland Ltd, 33 pp. https://github.com/gillibrandpa/unptrack.git

Gillibrand, PA, and Willis, KJ, (2007) Dispersal of Sea Lice Larvae from Salmon Farms: A Model Study of the Influence of Environmental Conditions and Larval Behaviour. Aquatic Biology, 1, 73-75.

Gillibrand, PA, Siemering, B, Miller PI, and Davidson, K. (2016) Individual-Based Modelling of the Development and Transport of a Karenia mikimotoi Bloom on the North-West European Continental Shelf. Harmful Algae, DOI: 10.1016/j.hal.2015.11.011

Hersbach, H, Bell, B, Berrisford, P, Hirahara, S, Horányi, A, Muñoz-Sabater, J, Nicolas, J, et al. (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146: 1999–2049. https://doi.org/10.1002/qj.3803

Heuch PA, Nordhagen JR, Schram TA (2000) Egg production in the salmon louse [*Lepeophtheirus salmonis* (Krøyer)] in relation to origin and water temperature. Aquac Res 31:805 – 814.

Hijmans, RJ (2020). Raster: Geographic Data Analysis and Modeling. R package version 3.4-5. https://CRAN.R-project.org/package=raster

Hinkle DE, Wiersma W, Jurs SG. (2003) Applied Statistics for the Behavioral Sciences. 5th ed. Boston: Houghton Mifflin; 2003.

Jackson FL, Malcolm IA, Hannah DM (2016) A novel approach for designing large-scale river temperature monitoring networks. Hydrol. Res., 47, pp. 569-590, 10.2166/nh.2015.106

Jardim, E, Azevedo, M, Brodziak, J, Brooks, EN, Johnson, KF, Klibansky, N, Millar, CP, Minto, C, Mosqueira, I, Nash, RD and Vasilakopoulos, P, (2021). Operationalizing ensemble models for scientific advice to

fisheries management. ICES Journal of Marine Science, 78(4), pp.1209-1216.

Ji, R, Ashjian, CJ, Campbell, RG, Chen, C, Gao, G, Davis, CS, Cowles, GW, and Beardsley, RC (2011) Life history and biogeography of calanus copepods in the Arctic Ocean: An individual-based modeling study. Progress In Oceanography 96(1): 40–56.

Johnsen IA, Asplin LC, Sandvik AD, Serra-Llinares RM (2016) Salmon lice dispersion in a northern Norwegian fjord system and the impact of vertical movements. Aquacult Environ Interact 8: 99-116. https://doi.org/10.3354/aei00162

Johnsen IA, Fiksen Ø, Sandvik AD, 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. https://www.jstor.org/stable/24864124

Kragesteen TJ, Simonsen K, Visser AW, Andersen KH (2018) Identifying salmon lice transmission characteristics between Faroese salmon farms. Aquaculture Environment Interactions 10, 49-60

Liu, Cowles, Churchill, & Stokesbury, (2015) FVCOM i-state configuration model: offline Lagrangian / IBM model for FVCOM. https://github.com/GeoffCowles/fiscm

Maier, HR, Guillaume JHA, vanDelden, H, Riddell GA, Haasnoot, M, Kwakkel JH (2016) An uncertain future, deep uncertainty, scenarios, robustness and adaptation: how do they fit together? Environ. Model. Software, 81(2016),pp.154-164, 10.1016/j.envsoft.2016.03.014

Mukaka MM. (2012) Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J. 2012 Sep;24(3):69-71. PMID: 23638278; PMCID: PMC3576830.

Murray, AG, Shephard, S, Asplin, L, Adams, TP, Ådlandsvik, B, Gallego, AG, Hartnett, M, Johnsen, IA, Jones, SRM, Moriarty, M, Nash, S, Pert, CC, Rabe, B, Gargan, PG, (2022). A standardised generic framework of sea lice

model components for application in coupled hydrodynamic-particle models. In: Treasurer, J, Bricknell, I and Bron, J (Eds.), Sea Lice Biology and Control, 5M Books Ltd, pp.167 – 187.

Murray, AG, & Moriarty, M. (2021). A simple modelling tool for assessing interaction with host and local infestation of sea lice from salmonid farms on wild salmonids based on processes operating at multiple scales in space and time. Ecological Modelling, 443, 109459.

Nelson EJ, Robinson SMC, Feindel N, Sterling A, Byrne A, Pee Ang K (2018) Horizontal and vertical distribution of sea lice larvae (*Lepeophtheirus salmonis*) in and around salmon farms in the Bay of Fundy, Canada. Journal of Fish Diseases 41, 885-899

O'Dea, EJ, Arnold, AK, Edwards, KP, Furner, R, Hyder, P, Martin, MJ, et al. (2012). An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf. Journal of Operational Oceanography. https://doi.org/10.1080/1755876X.2012.11020128

Oidtman, RJ, Omodei, E, Kraemer, MU, Castañeda-Orjuela, CA, Cruz-Rivera, E, Misnaza-Castrillón, S, Cifuentes, MP, Rincon, LE, Cañon, V, Alarcon, PD. and España, G, (2021). Trade-offs between individual and ensemble forecasts of an emerging infectious disease. Nature communications, 12(1), pp.1-11

Penston MJ, Millar CP, Zuur A, Davies IM. (2008) Spatial and temporal distribution of *Lepeophtheirus salmonis* (Kroyer) larvae in a sea loch containing Atlantic salmon, *Salmo salar* L., farms on the north-west coast of Scotland. Journal of Fish Diseases 31, 877-877.

Pert C, Collins C, Salama N, Dunn J, Wallace S, MacDonald P, Murray A & Rabe B. (2021). Loch Linnhe Biological Sampling Data Products. DOI: 10.7489/12361-2

Pert CC, Fryer RJ, Cook P, Kilburn R, McBeath S, McBeath A, Matejusova I, Urquhart K, Weir SJ, McCarthy U, Collins C, Amundrud T,

Bricknell IR. (2014) Using sentinel cages to estimate infestation pressure on salmonids from sea lice in Loch Shieldaig, Scotland. Aquaculture Environment Interactions 5, 49-59.

Price, D, Stuiver, C, Johnson, H, Gallego, A, & Murray, ROH. (2016). The Scottish Shelf Model. Part 5: Wider Loch Linnhe System Sub-Domain. Scottish Marine and Freshwater Science, 7(7). https://doi.org/10.7489/1696-1

Rabe, B. and Hindson J. (2017). Forcing mechanisms and hydrodynamics in Loch Linnhe, a dynamically wide Scottish estuary. Estuarine, Coastal and Shelf Science. 196, 159-172.

Rittenhouse MA, Revie CW, Hurford A (2016) A model for sea lice (Lepeophtheirus salmonis) dynamics in a seasonally changing environment. Epidemics 16:8–16.

Ross, ON, Sharples, J, (2004). Recipe for 1-D Lagrangian particle tracking models in space-varying diffusivity. Limnology and Oceanography: Methods, 2, 289-302.

Salama, N, Collins, C, Fraser, J, Dunn, J, Pert, C, Murray, A, Rabe, B, (2013). Development and assessment of a biophysical dispersal model for sea lice. J. Fish Dis. 36, 323e337.

Salama, N, Dale, A, Ivanov, V, Cook, P, Pert, C, Collins, C, Rabe, B (2018). Using biological-physical modelling for informing sea lice dispersal in Loch Linnhe, Scotland. Journal of Fish Diseases, 901-919. https://doi.org/10.1111/jfd.12693

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. Aquat. Sci. 73: 1841–1851 dx.doi.org/10.1139/cjfas-2016-0050

Sandvik AD, Bjørn PA, Ådlandsvik B, Asplin L, Skarđhamar J, Johnsen, IA, Myksvoll M. Skogen M.D. (2016) Towards a model-based prediction

system for salmon lice infestation pressure. Aquaculture environment interactions, 8, 527-542 doi: 10.3354/aei00193

Sandvik AD, Johnsen IA, Myksvoll MS, Sævik PN, Skogen MD (2020) Prediction of the salmon lice infestation pressure in a Norwegian fjord. ICES J Sea Res 77: 746-756. DOI: 10.1093/icesjms/fsz256

Scottish Government (2016) Scottish Shelf Model. Part 5: Wider Loch Linnhe System Sub-Domain. Scottish Government, St. Andrew's House, Regent Road, Edinburgh EH1 3DG Tel:0131 556 8400 ceu@scotland.gsi.gov.uk.

Stien A, Bjørn, PA, Heuch, PA, Elston, DA. (2005) Population dynamics of Salmon lice *Lepeophtheirus salmonis* on Atlantic salmon and sea trout. Marine Ecological Progress Series 290, 263 – 275.

Tebaldi, C. and Knutti, R., (2007). The use of the multi-model ensemble in probabilistic climate projections. Philosophical transactions of the royal society A: mathematical, physical and engineering sciences, 365(1857), pp.2053-2075.

Vieno M, Heal MR, Hallsworth S, Famulari D, Doherty RM, Dore AJ, Tang YS, Braban CF, Leaver D, Sutton MA et al. (2014). The role of long-range transport and domestic emissions in determining atmospheric secondary inorganic particle concentrations across the UK. Atmospheric Chemistry and Physics. 14

Vieno M, Heal MR, Williams ML, Carnell EJ, Nemitz E, Stedman JR, Reis S. (2016). The sensitivities of emissions reductions for the mitigation of UK PM2.5. Atmospheric Chemistry and Physics. 16

Visser, AW, (1997). Using random walk models to simulate the vertical distribution of particles in a turbulent water column. Mar. Ecol. Prog. Ser., 158, 275-281.

Willis, KJ, Gillibrand, PA, Cromey CJ, and Black, KD, (2005). Sea lice treatments on salmon farms have no adverse effect on zooplankton communities: A case study. Marine Pollution Bulletin, 50, 806 – 816.

Wolf, J, Yates, N, Brereton, A, Buckland, H, De Dominicis, M, Gallego, A., et al., (2016). The Scottish shelf model. part 1: shelf-wide domain. Scottish Mar. Fresh. Sci. 7:151.

7. Appendix 1 – Details of Field Data

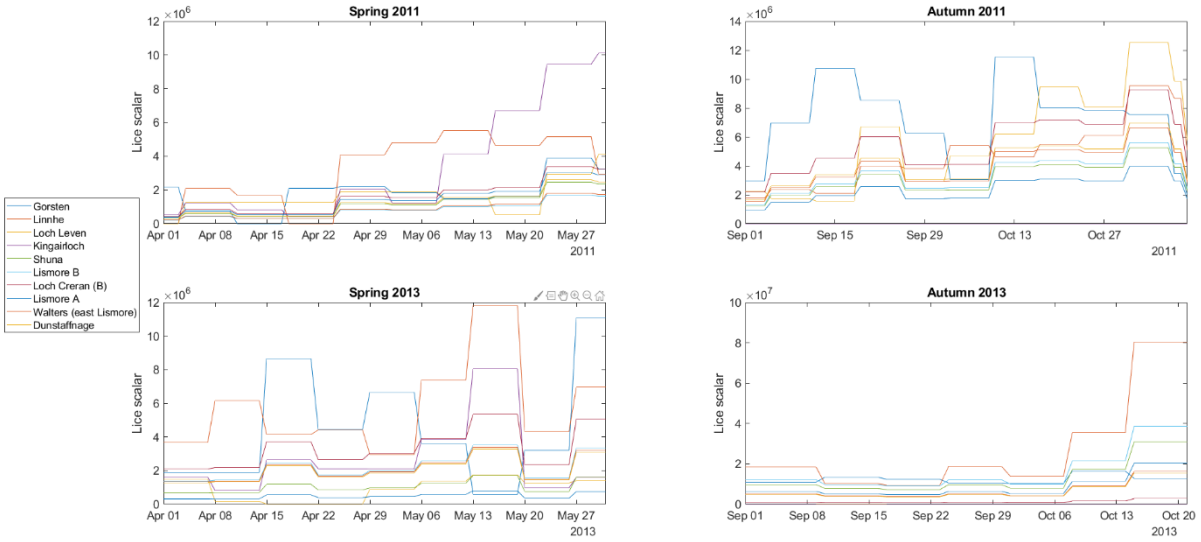


Figure S1.1 Time series of lice scaler for four simulation periods.

Table S1.2 The plankton sampling locations used in the 2011-2013 field study. Data sets can be found and downloaded from Pert et al. (2021)

Plankton Station Number	*Latitude (DDM)*	*Longitude (DDM)*
1	56° 50.33.954 N	5° 7.40.39 W
2	56° 50.23.924 N	5° 7.49.508 W
3	56° 50.15.718 N	5° 7.54.067 W
4	56° 49.26.483 N	5° 7.21.243 W
5	56° 49.23.748 N	5° 7.6.655 W
6	56° 49.20.101 N	5° 6.54.802 W
7	56° 48.21.747 N	5° 8.54.243 W
8	56° 48.10.806 N	5° 8.39.655 W
9	56° 48.1.689 N	5° 8.24.155 W
10	56° 46.58.777 N	5° 11.16.479 W
11	56° 46.32.335 N	5° 10.54.597 W
12	56° 46.15.012 N	5° 10.34.538 W
13	56° 44.52.952 N	5° 13.55.127 W
14	56° 44.35.629 N	5° 13.36.892 W
15	56° 44.18.305 N	5° 13.8.627 W
16	56° 42.48.952 N	5° 15.8.98 W
17	56° 42.42.57 N	5° 14.42.539 W
18	56° 42.1.54 N	5° 13.22.303 W

19	56° 41.42.393 N	5° 13.19.568 W
20	56° 41.20.51 N	5° 13.17.744 W
21	56° 41.58.805 N	5° 11.12.832 W
22	56° 42.22.511 N	5° 17.40.334 W
23	56° 38.37.304 N	5° 18.4.951 W
24	56° 34.53.921 N	5° 22.55.806 W
25	56° 36.54.274 N	5° 31.3.602 W
26	56° 35.39.509 N	5° 29.56.131 W
27	56° 34.50.091 N	5° 29.2.336 W
28	56° 33.57.209 N	5° 28.15.836 W
29	56° 32.47.003 N	5° 24.59.806 W
30	56° 29.27.325 N	5° 25.3.453 W
31	56° 29.5.443 N	5° 27.19.307 W

Table S1.3. The sentinel cage locations used in the 2011-2013 field study. Site 3a was moved after the first samples were collected in May 2011. Site 7 did not produce any samples as it was too difficult due to the current speeds. Data sets can be found and downloaded from Pert et al. (2021).

Site Number	Site_Name	Latitude (DDM)	Longitude (DDM)	Longitude	Latitude	Depth (m)
1	Head of Loch Linnhe (Corpach)	56 50.184N	5 06.721W	-5.112	56.836	8
2	North of the Corran narrows	56 44.143N	5 14.949W	-5.249	56.736	12
3a	South of the Corran narrows	56 42. 612N	5 15. 509W			8
3b	West of Sallachan point	56° 42.292N	5° 17.684W	-5.295	56.709	12
4	Ballachuish Bay (off from Onich hotel)	56 42.032N	5 12.783W	-5.213	56.701	20
5	Inversanda Bay	56 40.992N	5 20.919W	-5.349	56.683	7
6	Cuil Bay	56 38.876N	5 19.122W	-5.319	56.648	16
7	Shuna Island	56 35. 254N	5 23. 049W			12
8	Loch a'Choire (east of Kingarloch)	56 36.546N	5 30.021W	-5.500	56.609	17
9	Camas Nathais	56 28.959N	5 27.667W	-5.461	56.483	13
10	Rubha a'Mhothair	56 30.841N	5 37.955W	-5.633	56.514	9 - 22

8. Appendix 2 – Additional hydrodynamic model validation figures

'Gorsten Jan 2013', 'Gorsten North Jan 2013'

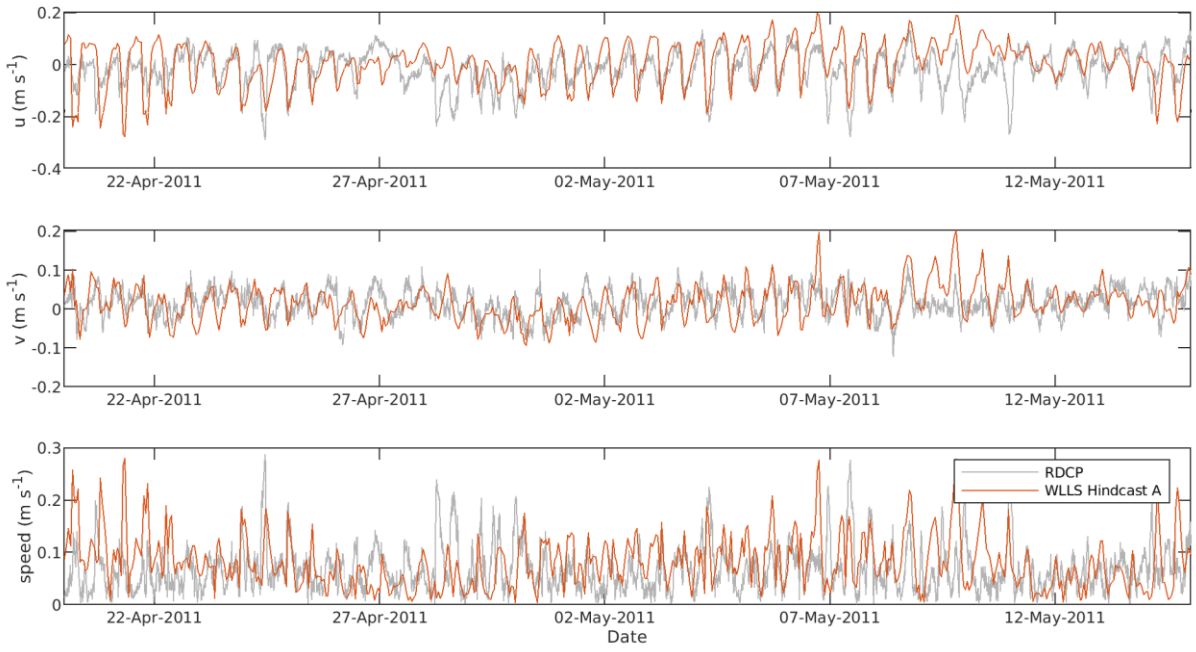


Figure S2.1: Comparison of measured near surface current speeds from RDCP deployed near Kilmalieu in Apr 2011 with the modelled near surface currents from WLLS hindcast A.

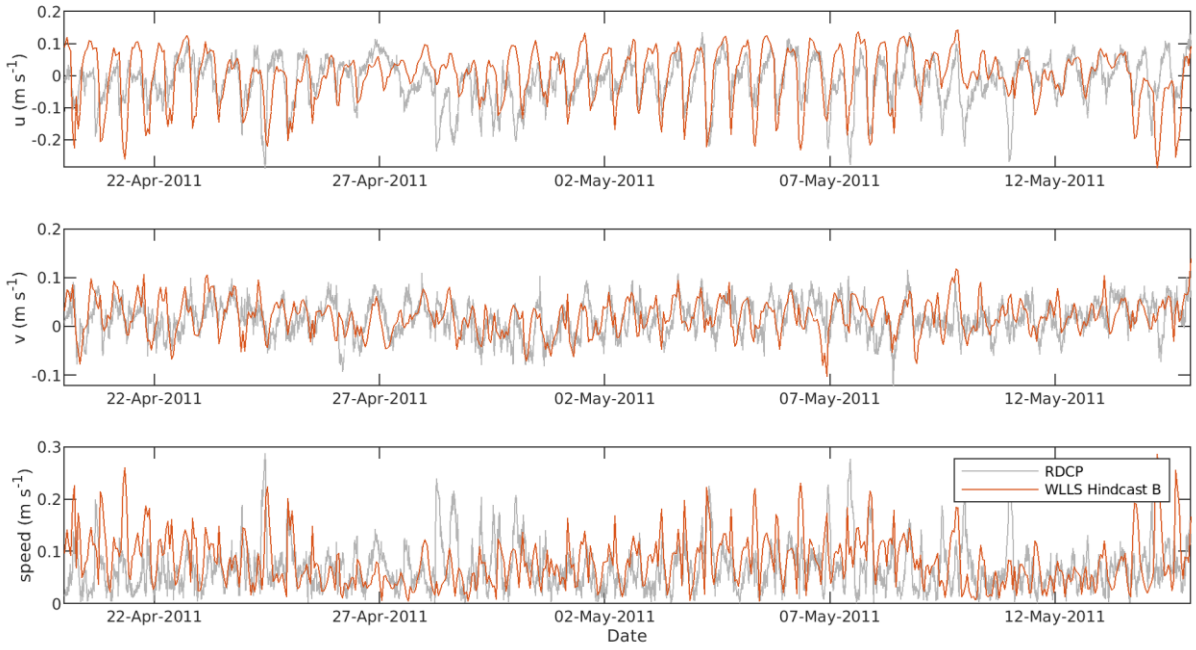


Figure S2.2: Comparison of measured near surface current speeds from RDCP deployed near Kilmalieu in Apr 2011 with the modelled near surface currents from WLLS hindcast B.

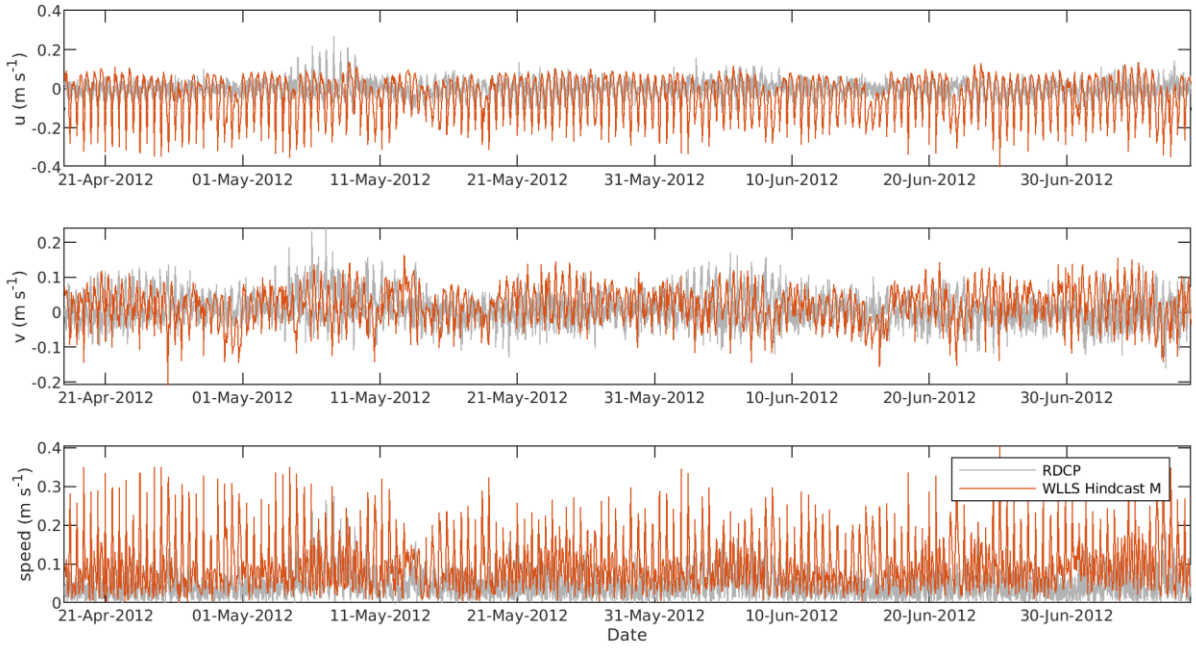


Figure S2.3: Comparison of measured near surface current speeds from RDCP deployed near Gearradh in Apr 2012 with the modelled near surface currents from WLLS hindcast M.

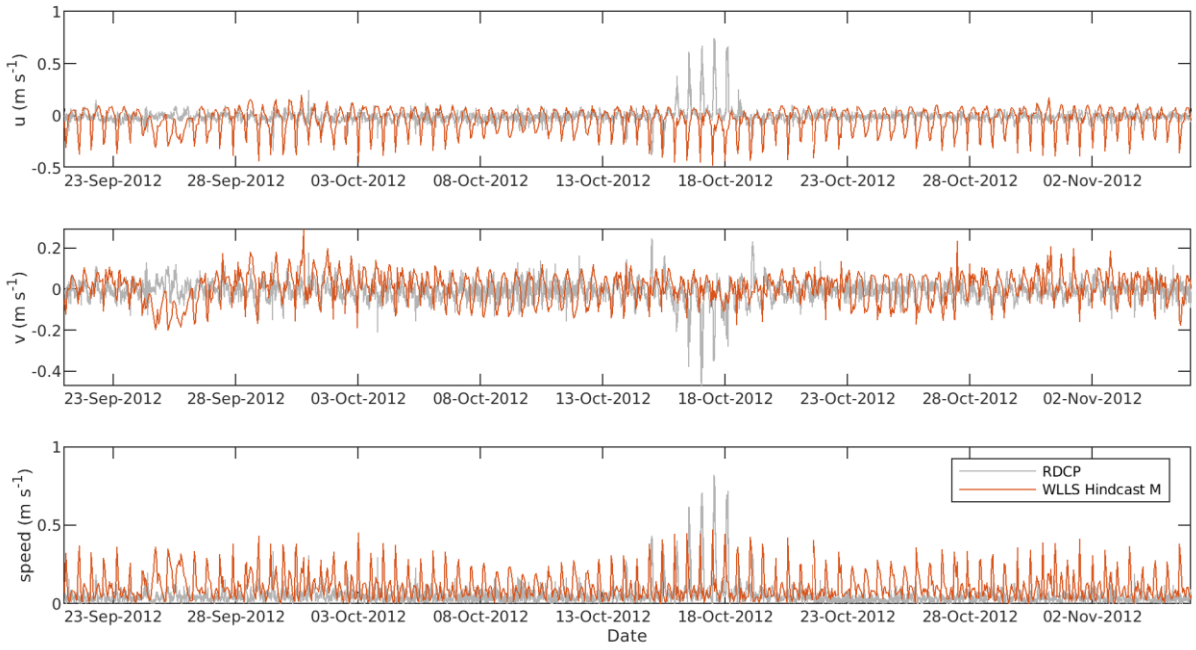


Figure S2.4: Comparison of measured near surface current speeds from RDCP deployed near Gearradh in Oct 2012 with the modelled near surface currents from WLLS hindcast M.

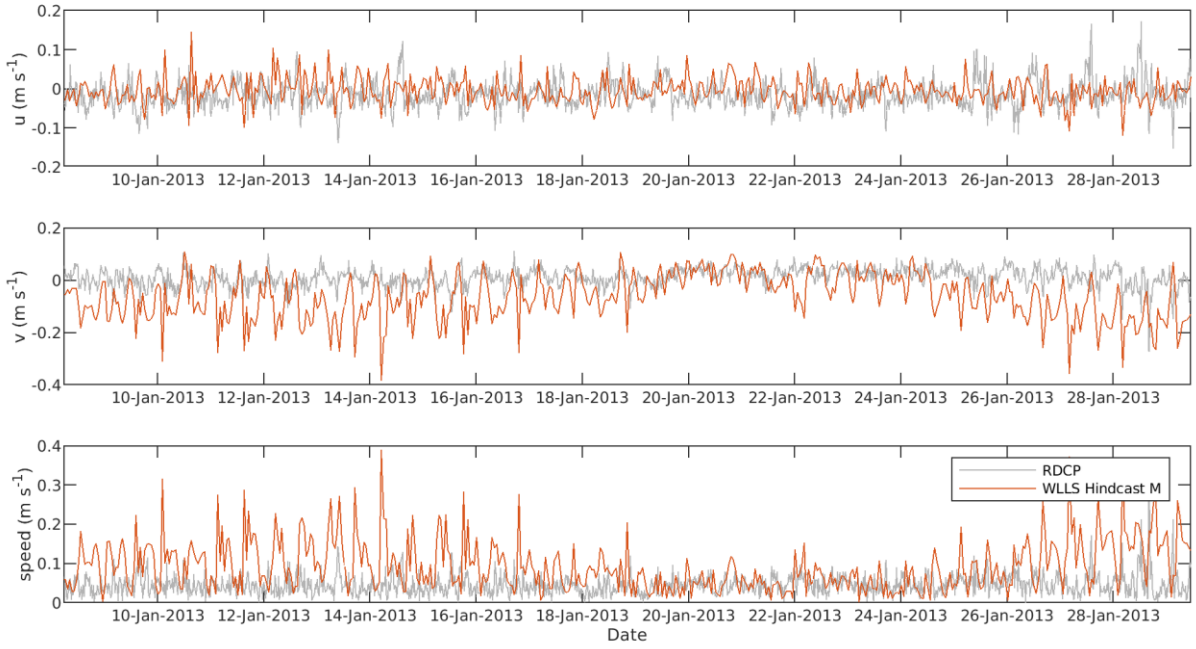


Figure S2.5: Comparison of measured near surface current speeds from RDCP deployed near Corran in Jan 2013 with the modelled near surface currents from WLLS hindcast M.

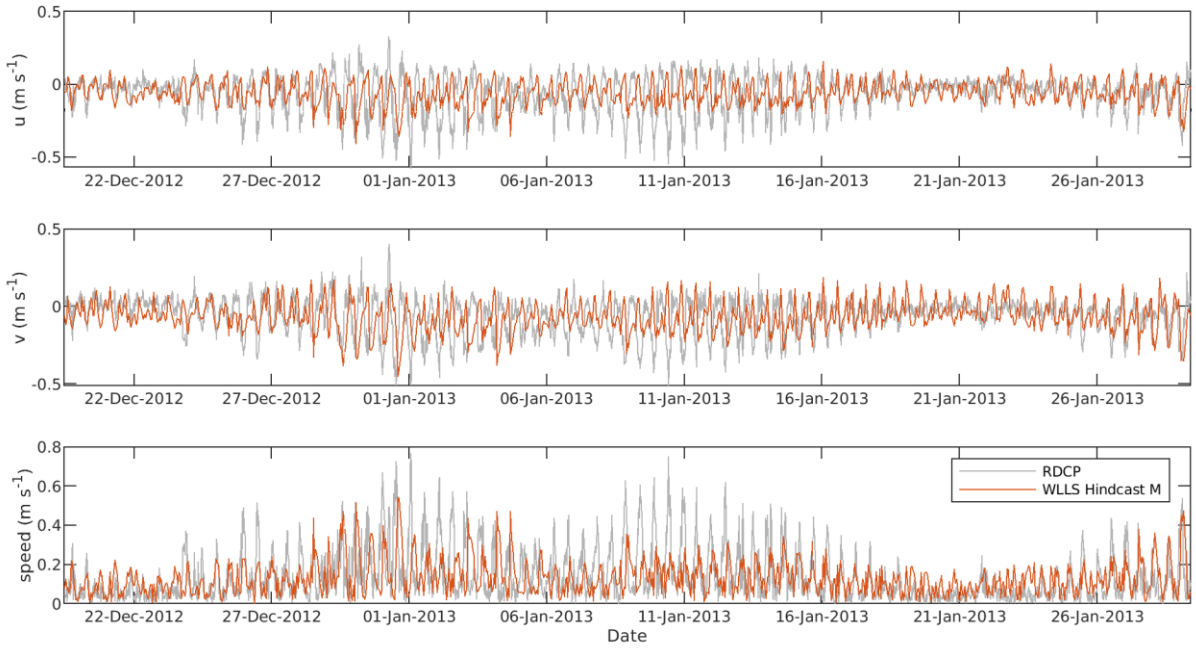


Figure S2.6: Comparison of measured near surface current speeds from RDCP deployed near Gorsten in Jan 2013 with the modelled near surface currents from WLLS hindcast M.

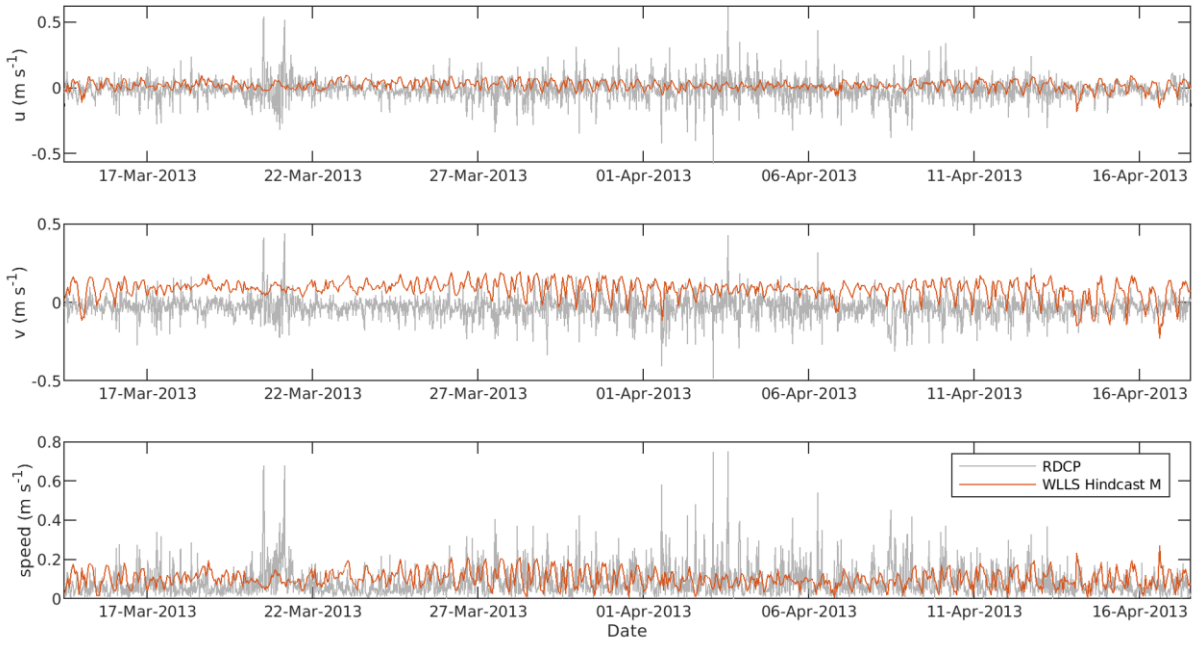


Figure S2.7: Comparison of measured near surface current speeds from RDCP deployed near Gorsten (Gorsten North) in Mar 2013 with the modelled near surface currents from WLLS hindcast M.

9. Appendix 3 – Additional sea lice dispersal model figures

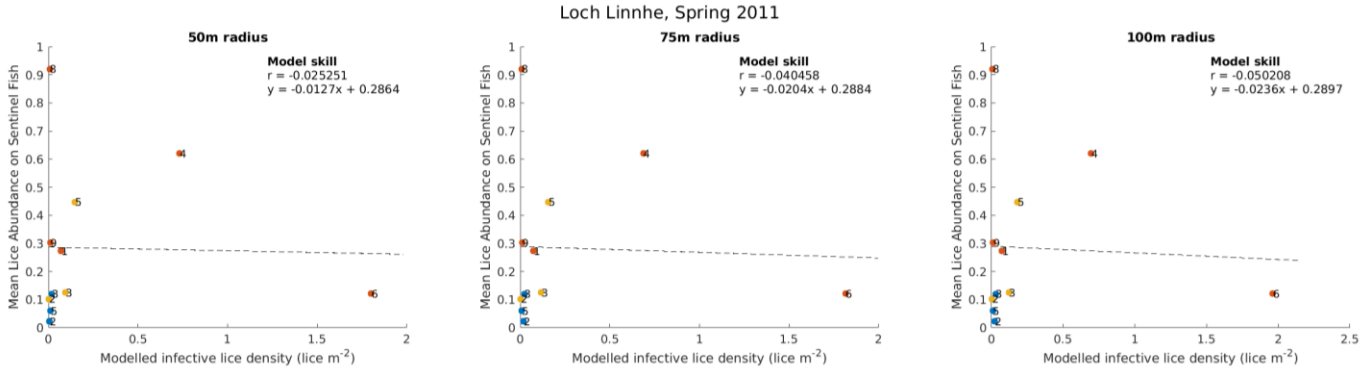


Figure S3.1: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the Biotracker PTM for Spring 2011, using 50 m, 75 m and 100 m radius circles to calculate modelled lice density for Deployments 1 (●), 2 (●) and 3 (●).

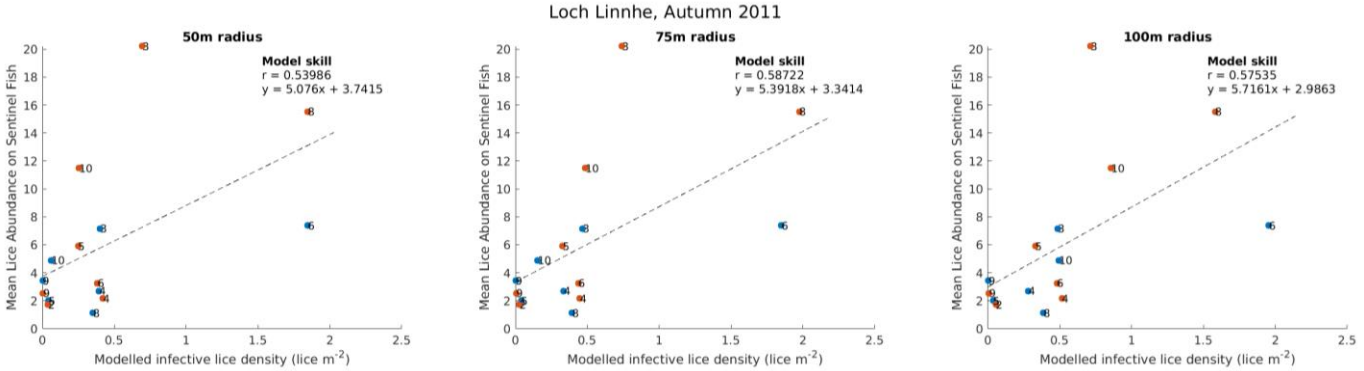


Figure S3.2: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the Biotracker PTM for Autumn 2011, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

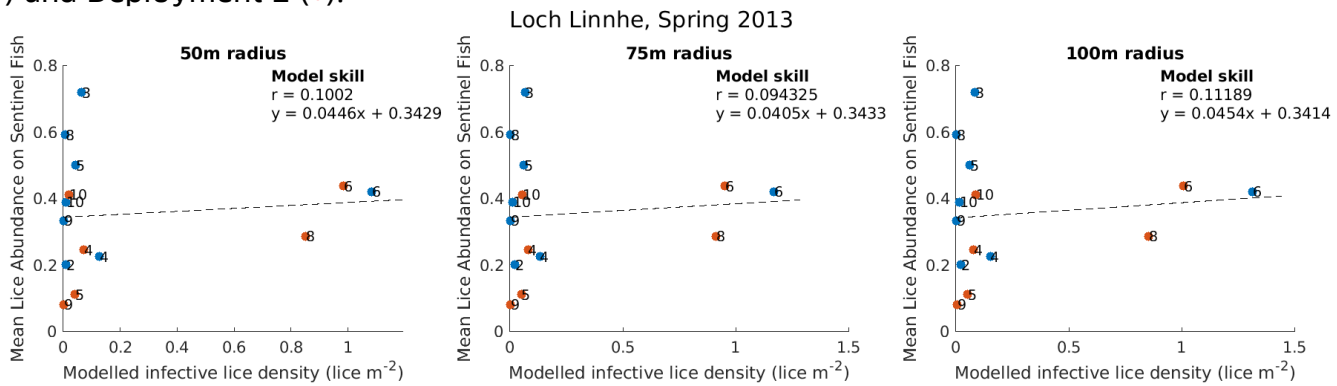


Figure S3.3: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the Biotracker PTM for Spring 2013, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

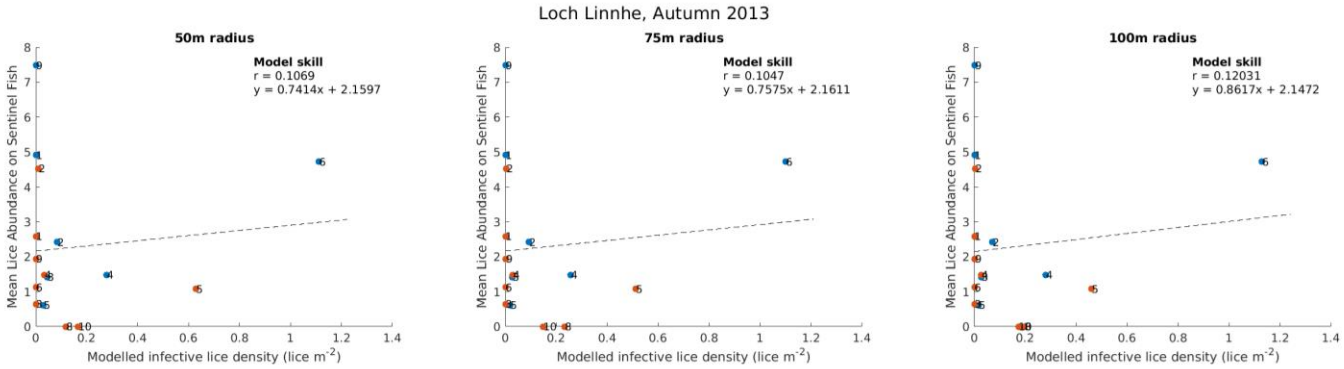


Figure S3.4: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the Biotracker PTM for Autumn 2013, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

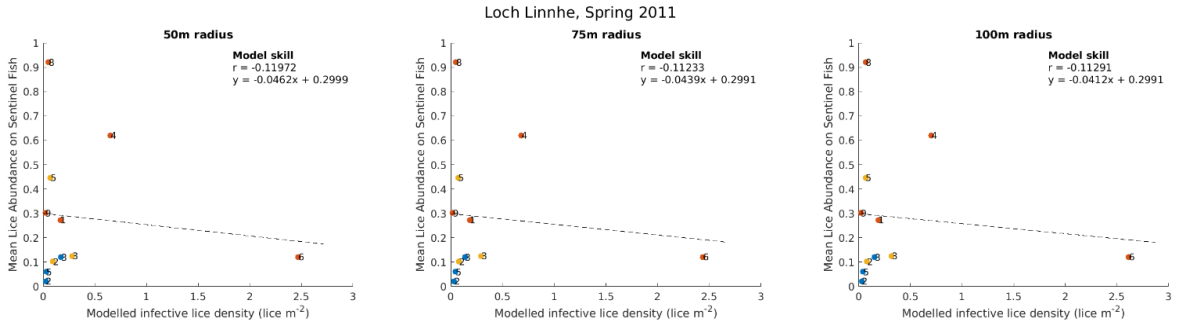


Figure S3.5: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the FISCM PTM for Spring 2011, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployments 1 (●), 2 (●) and 3 (●).

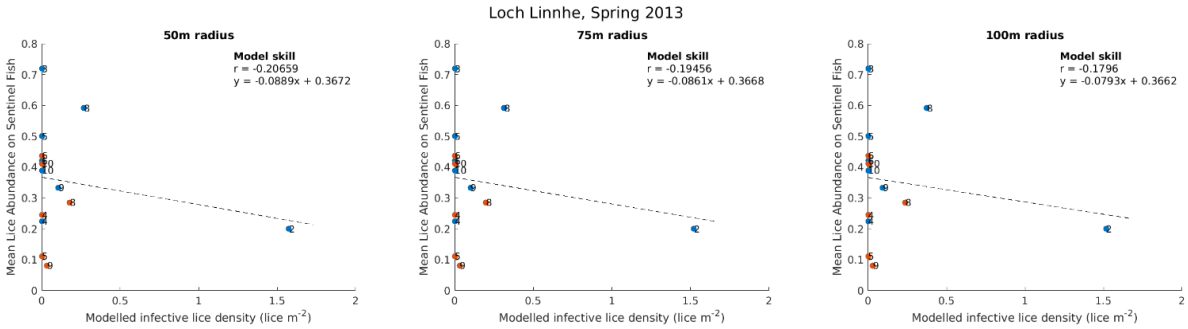


Figure S3.6: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the FISCM PTM for Spring 2013, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

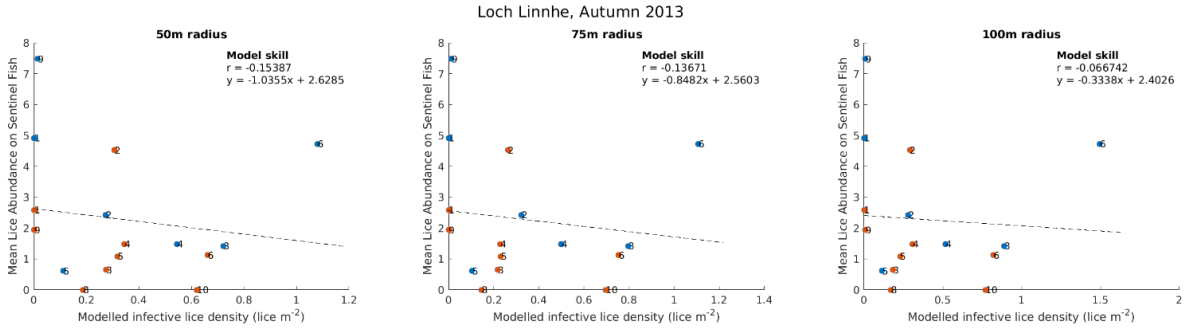


Figure S3.7: Comparison of modelled infective lice density with mean lice abundance on sentinel fish from the FISCM PTM for Autumn 2013, using 50m, 75m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

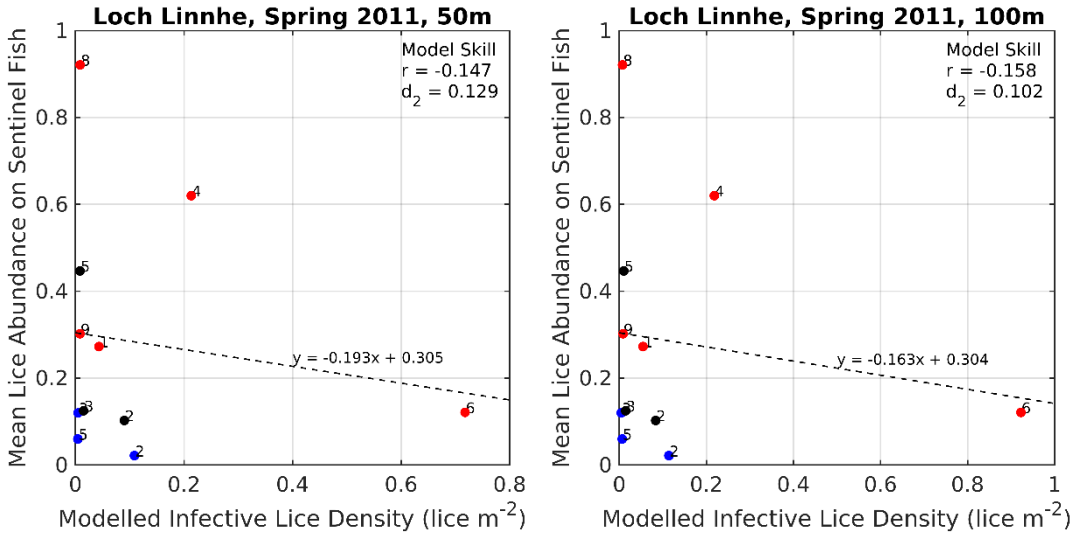


Figure S3.8: Comparison of modelled mean infective lice density with mean lice abundance on sentinel fish from the UnPTRACK PTM for Spring 2011, using 50m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

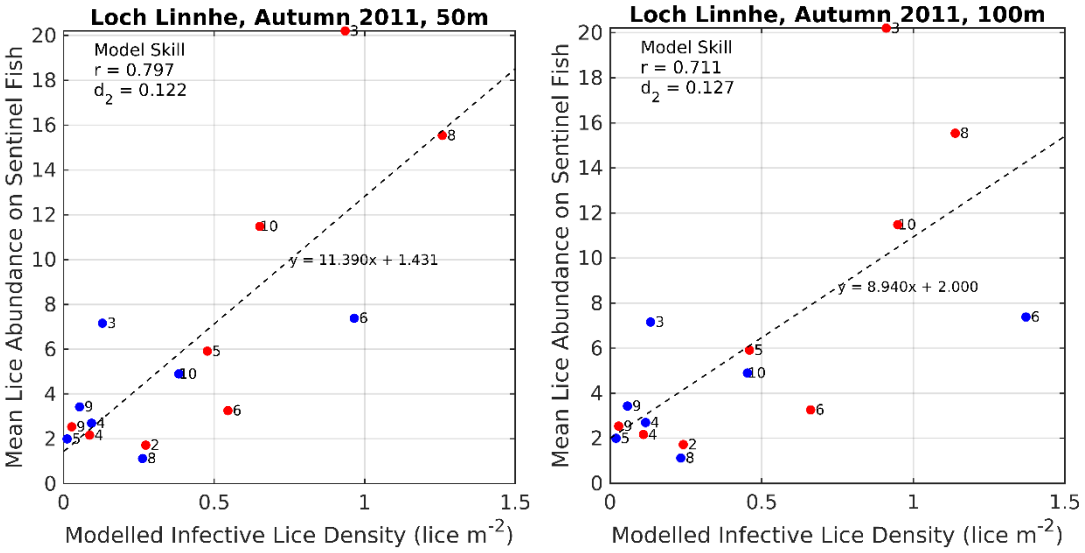


Figure S3.9: Comparison of modelled mean infective lice density with mean lice abundance on sentinel fish from the UnPTRACK PTM for Autumn 2011, using 50m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

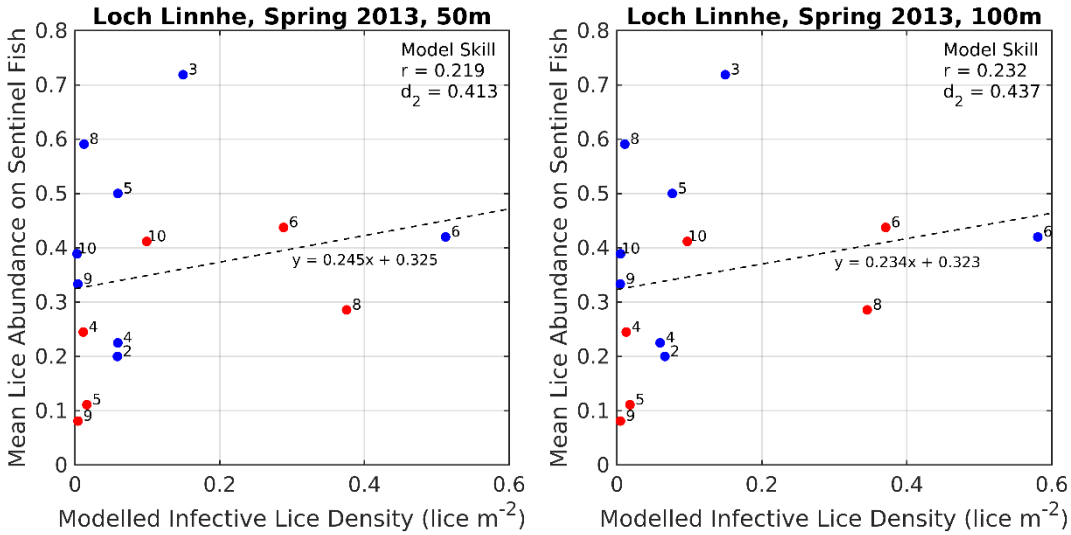


Figure S3.10: Comparison of modelled mean infective lice density with mean lice abundance on sentinel fish from the UnPTRACK PTM for Spring 2013, using 50m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

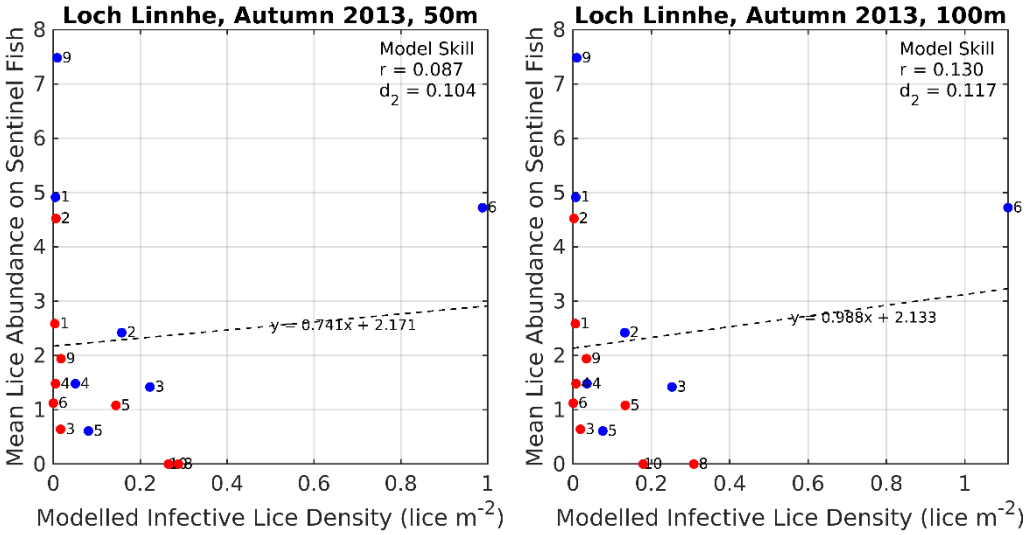


Figure S3.11: Comparison of modelled mean infective lice density with mean lice abundance on sentinel fish from the UnPTRACK PTM for Autumn 2013, using 50m and 100m radius circles to calculate modelled lice density for Deployment 1 (●) and Deployment 2 (●).

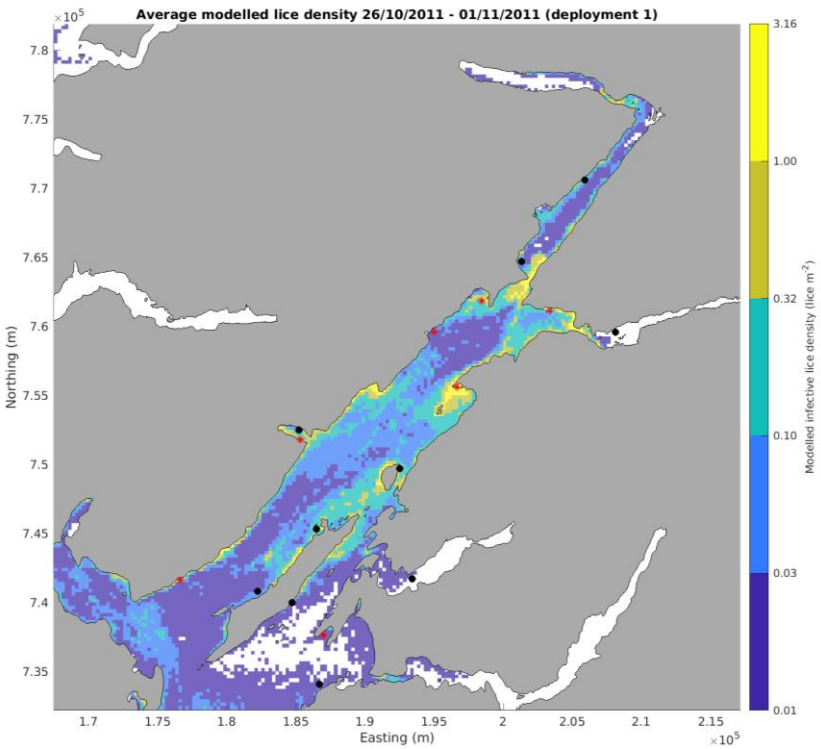


Figure S3.12: Average modelled lice density from the Biotracker PTM for the duration of the first deployment of sentinel cages in Autumn 2011

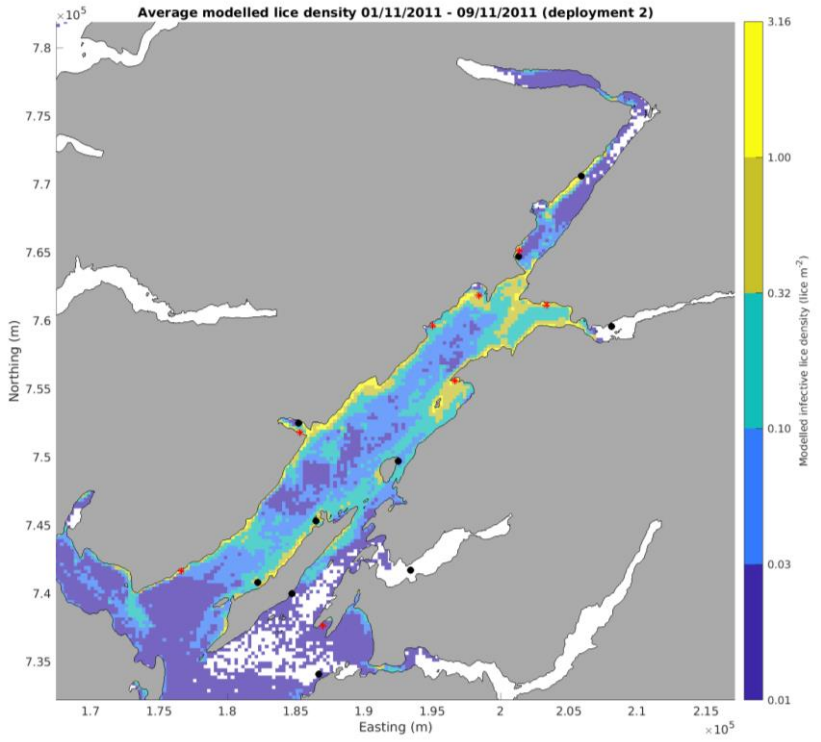


Figure S3.13: Average modelled lice density from the Biotracker PTM for the duration of the second deployment of sentinel cages in Autumn 2011.

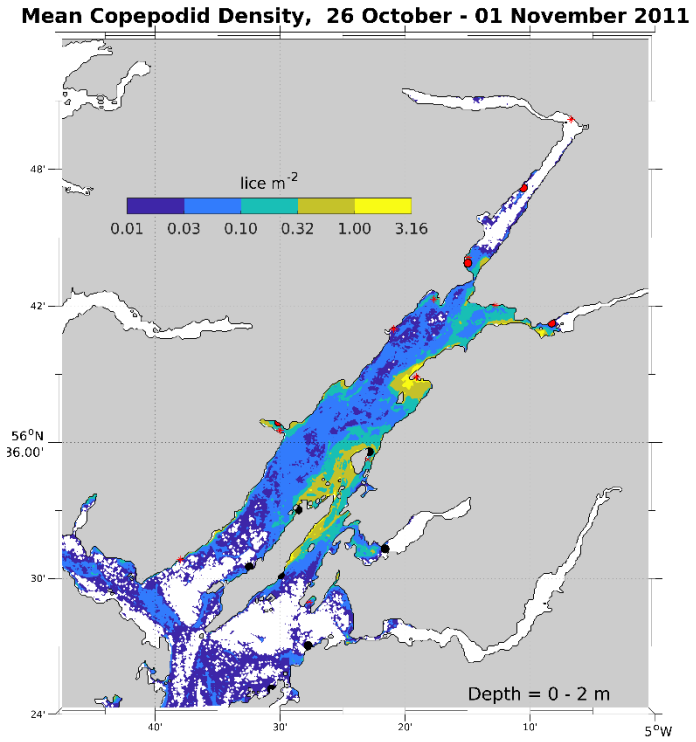


Figure S3.14: Average modelled lice density from the UnPTRACK PTM for the duration of the first deployment of sentinel cages in Autumn 2011

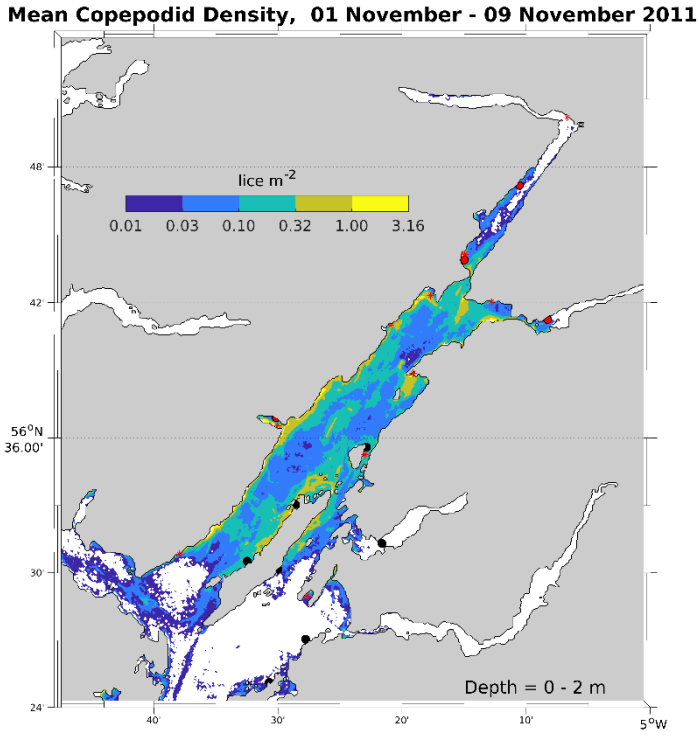
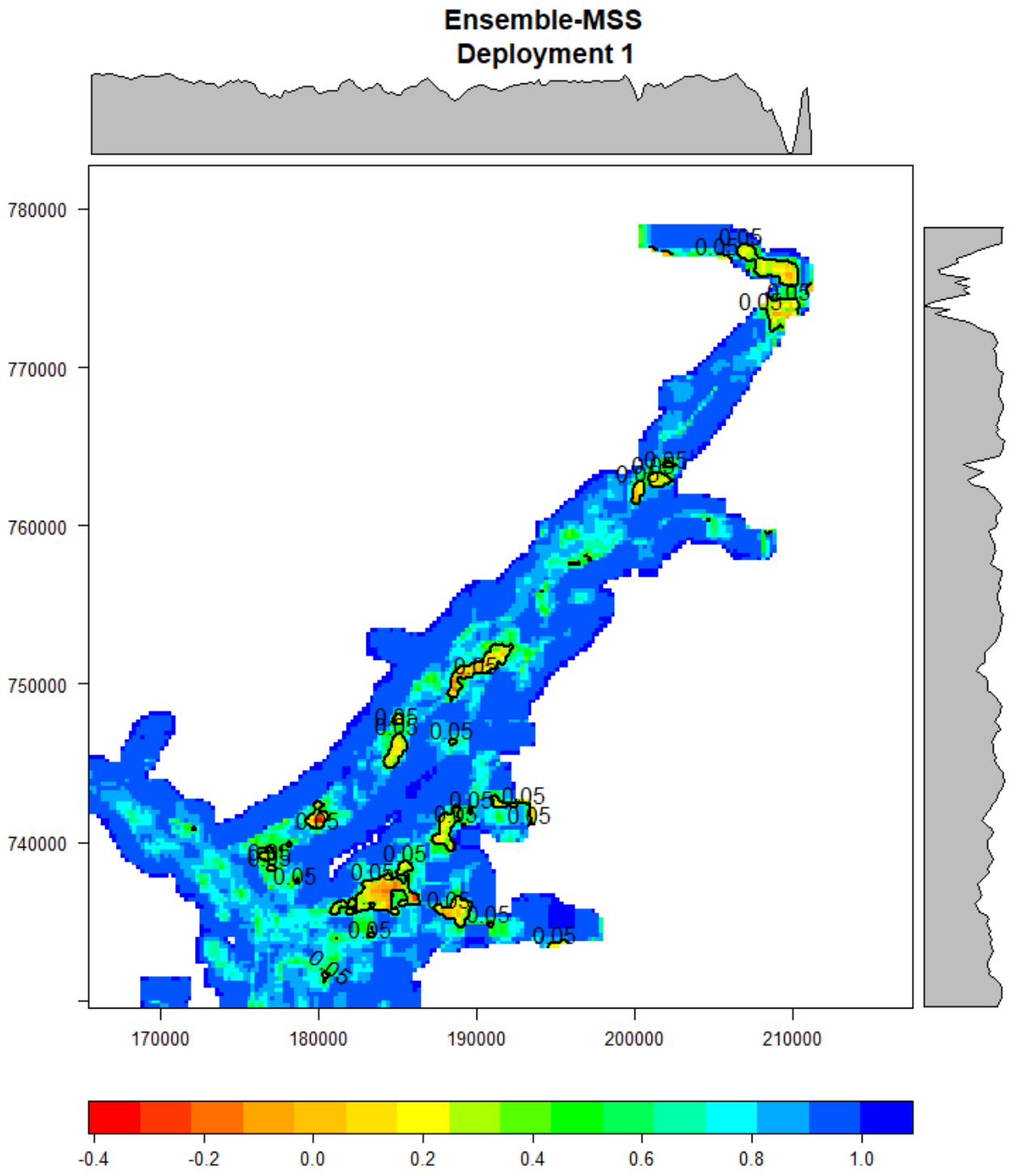


Figure S3.15: Average modelled lice density from the UnPTRACK PTM for the duration of the second deployment of sentinel cages in Autumn 2011.

10. Appendix 4 –Geographically Consistency Figures (Large Version)



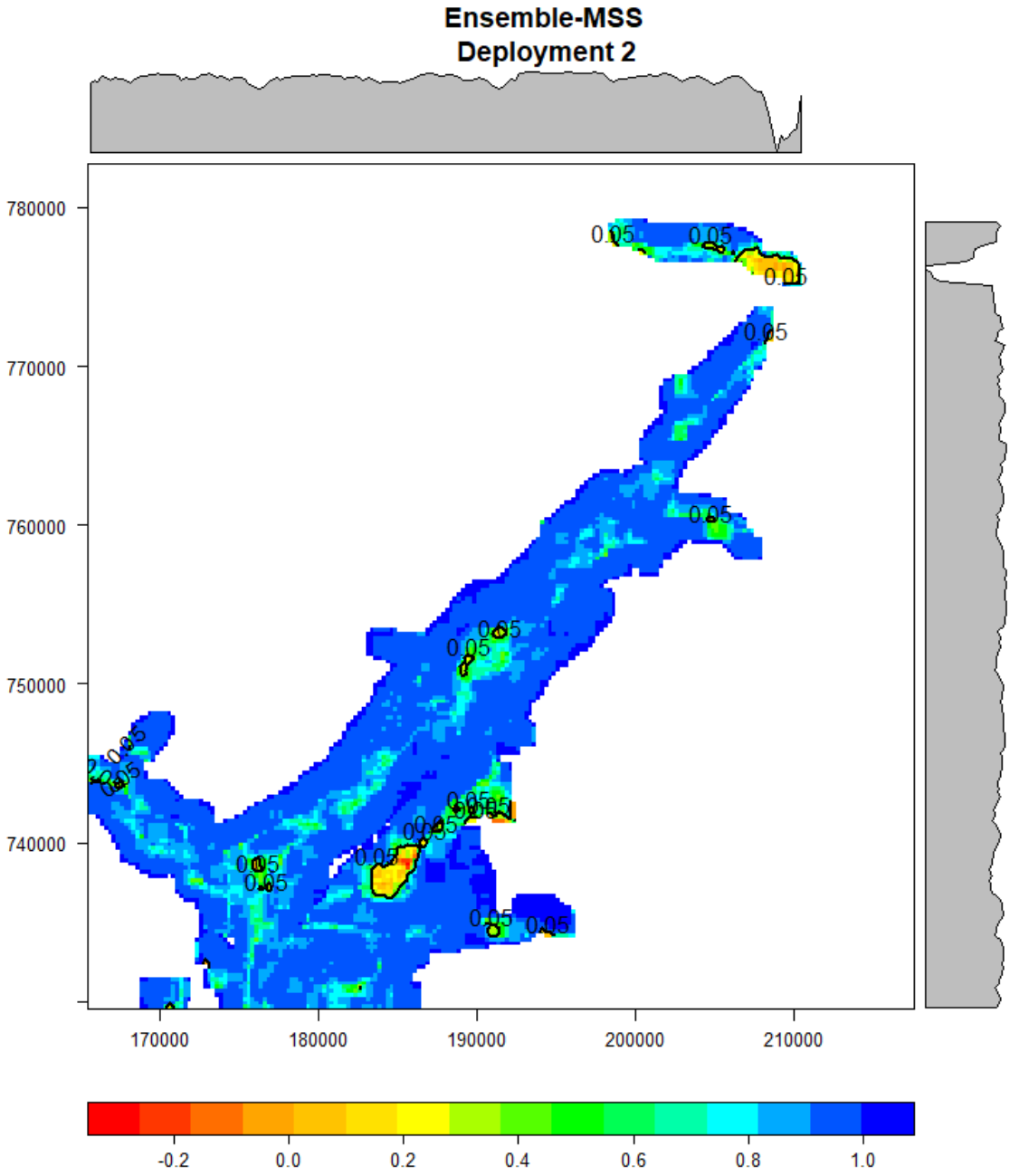
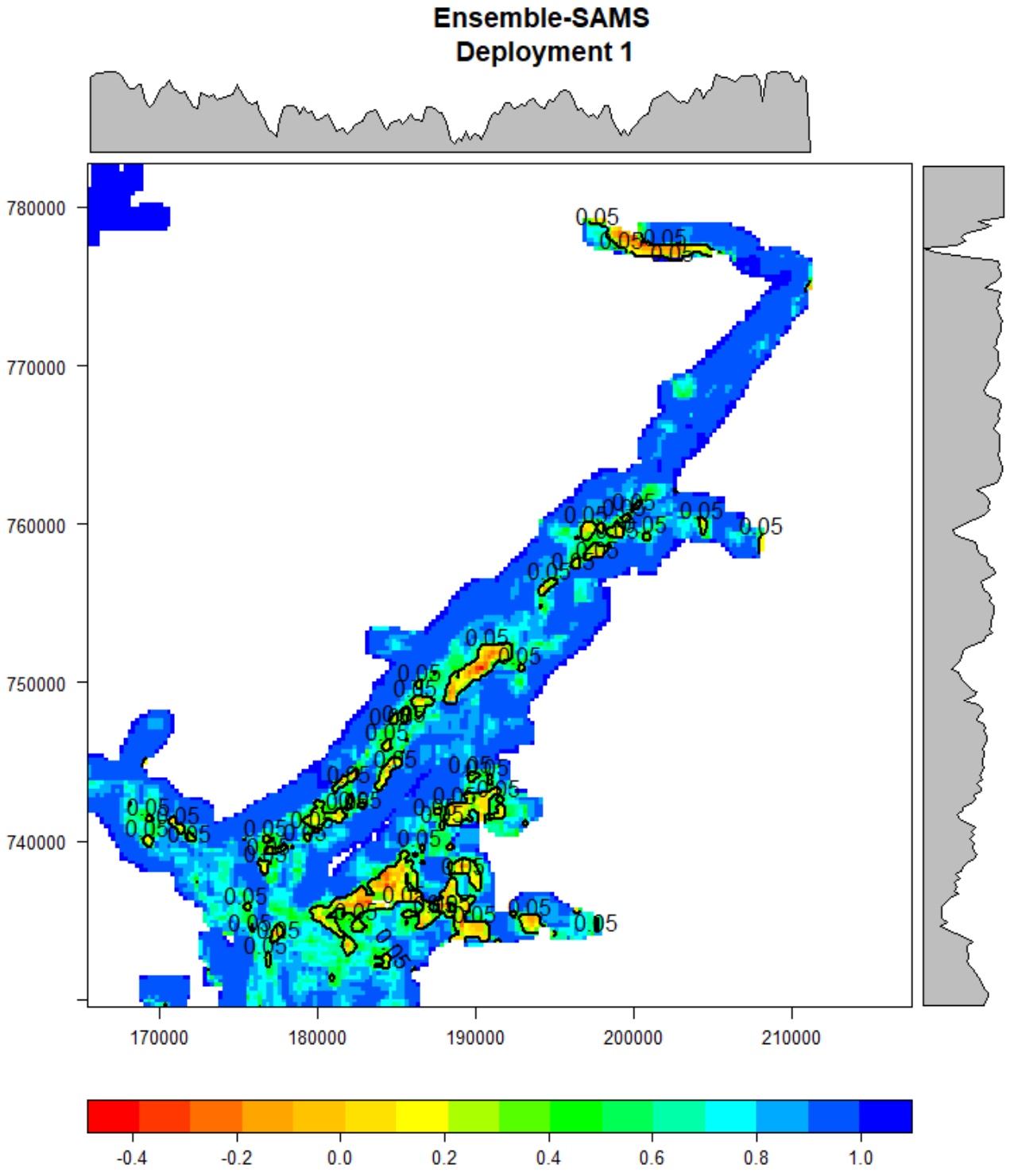


Figure S4.1: Deployment 1 showing MSS FISCM model geographical consistency with mean ensemble for deployment 1 and 2. The pearson's correlation provides an estimate of the strength of the linear relationship between two variables. Values of p = > 0.05 are generally considered significant, these areas are outlined in black. These figures are larger version of those shown in section 4.2.


Ensemble-SAMS
Deployment 1
780000
770000
760000
750000
740000
170000
180000
190000
200000
210000
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0

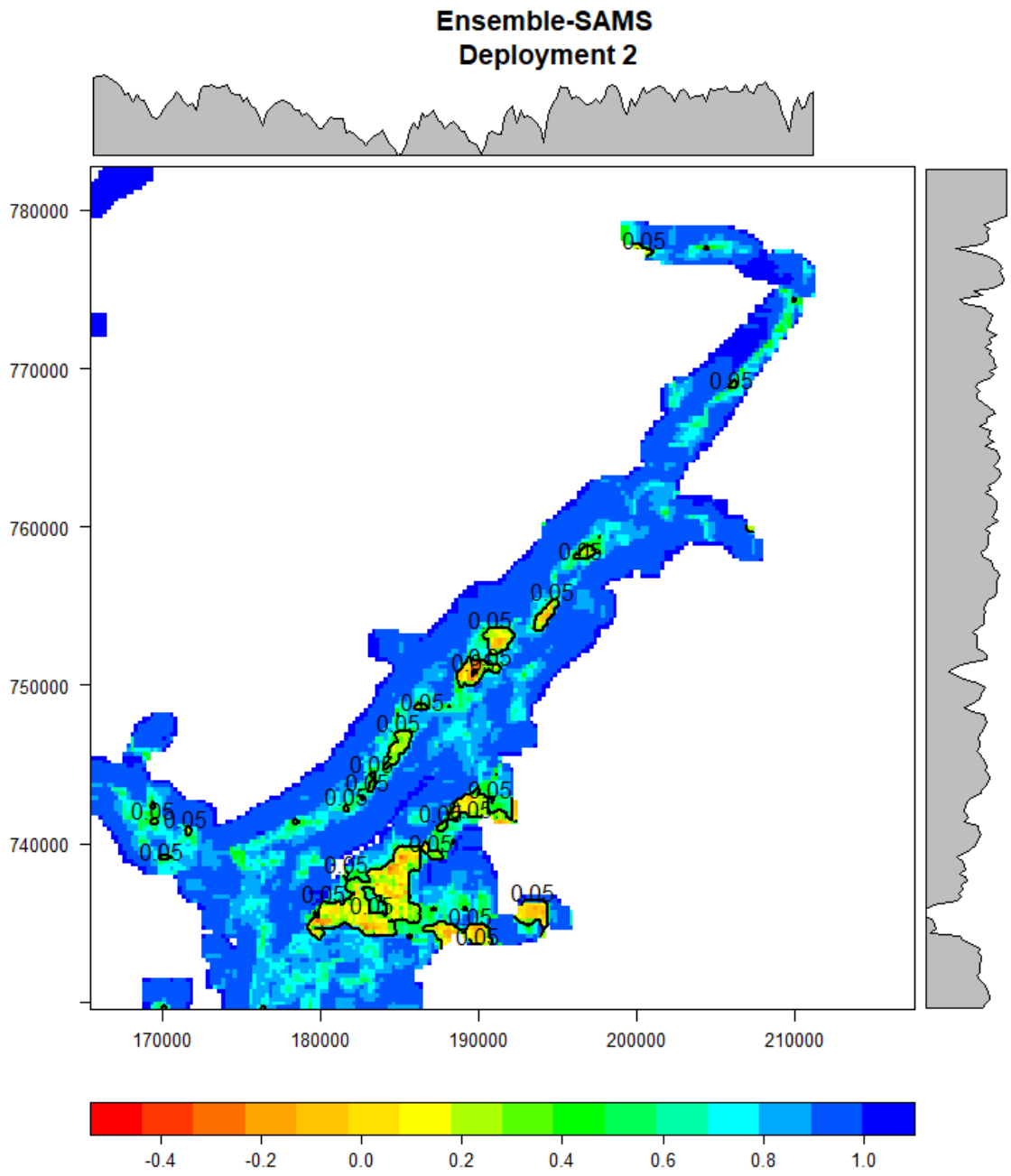
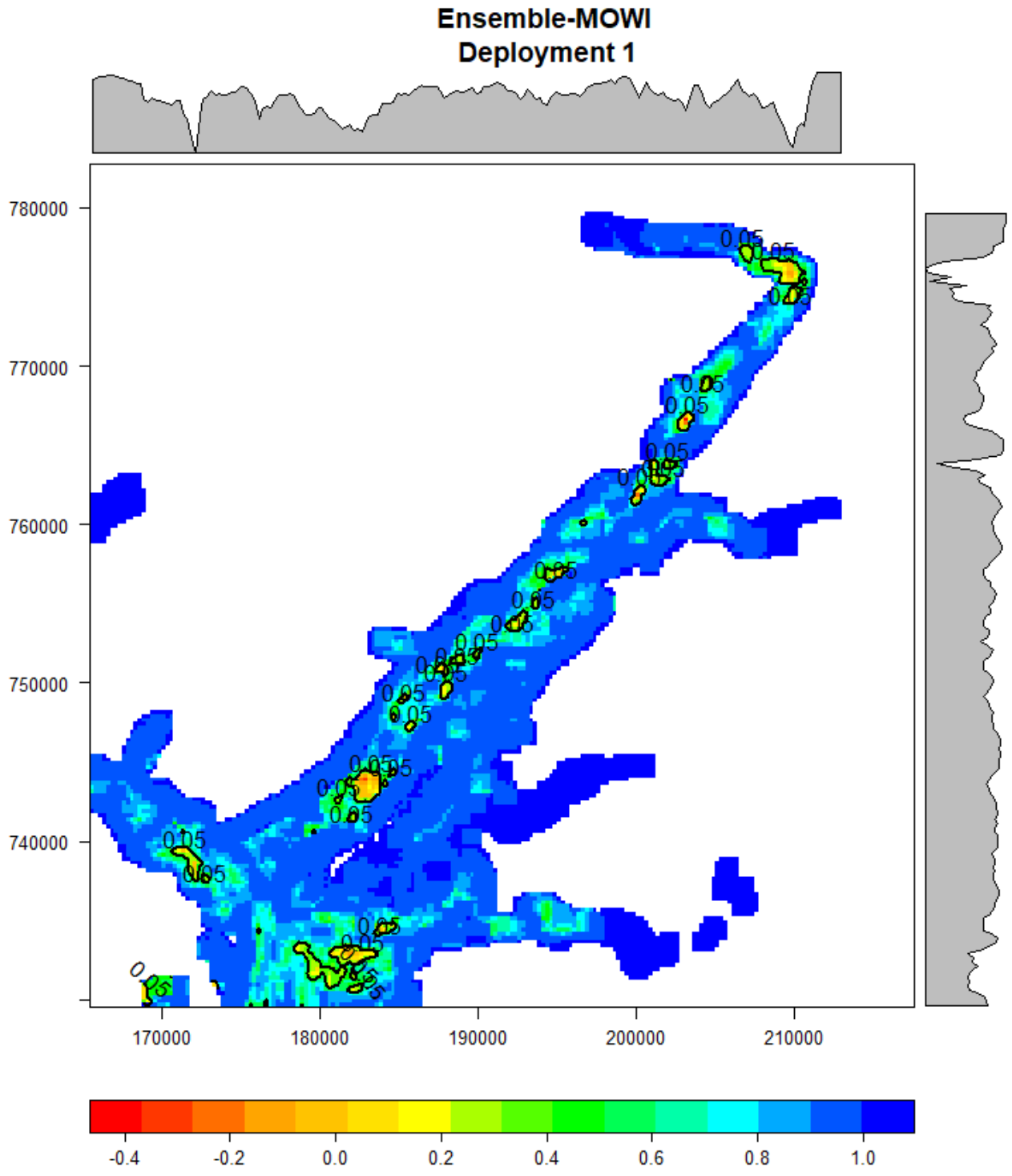


Figure S4.2: Deployment 1 showing SAMS BioTracker model geographical consistency with mean ensemble for deployment 1 and 2.. The pearson's correlation provides an estimate of the strength of the linear relationship between two variables. Values of p = > 0.05 are generally considered significant, these areas are outlined in black. These figures are larger version of those shown in section 4.2.


Ensemble-MOWI
Deployment 1
780000
770000
760000
750000
740000
170000
180000
190000
200000
210000
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0

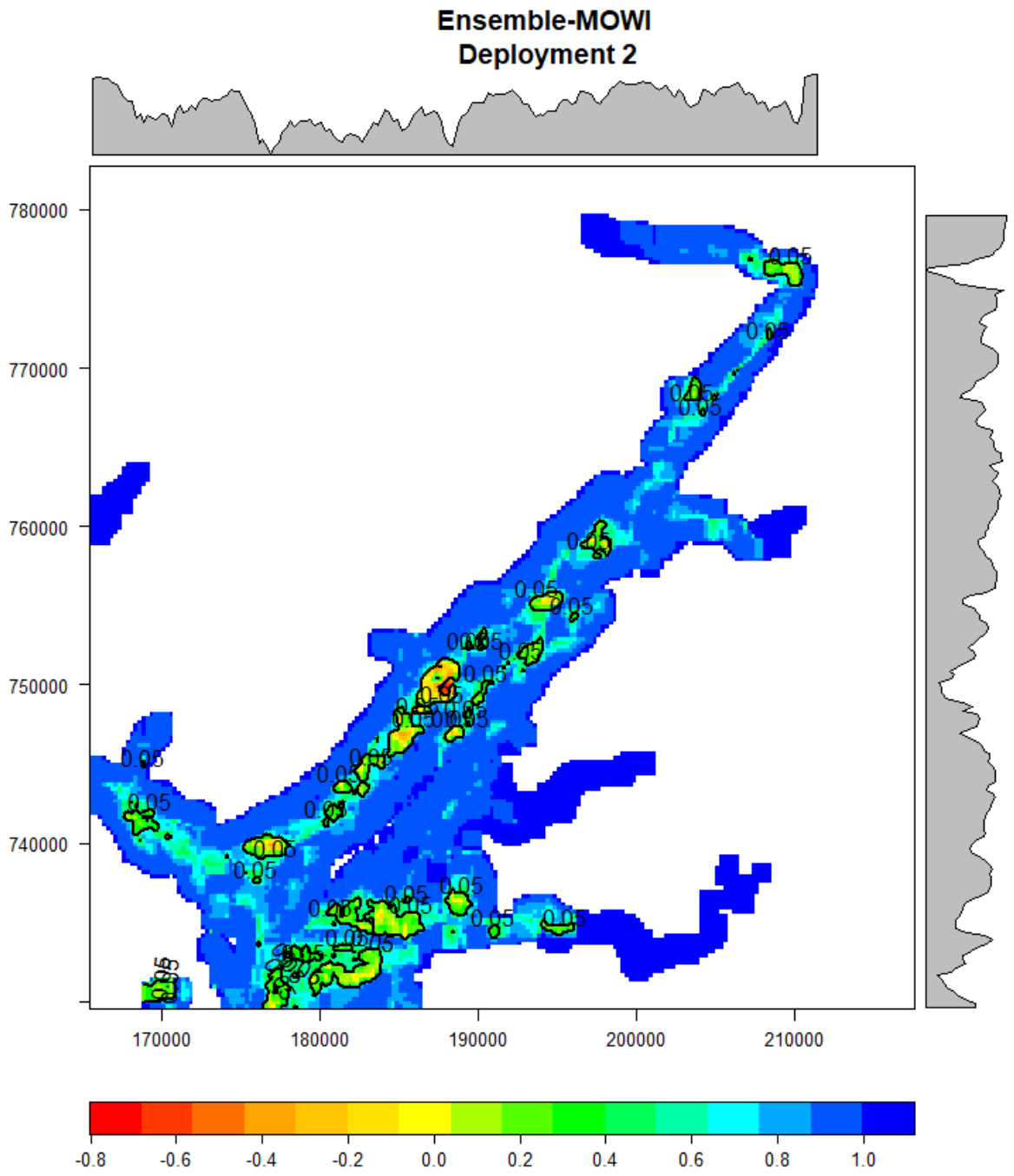


Figure S4.2: Deployment 1 showing MOWI's UnPTRACK model geographical consistency with mean ensemble for deployment 1 and 2.. The pearson's correlation provides an estimate of the strength of the linear relationship between two variables. Values of p = > 0.05 are generally considered significant, these areas are outlined in black. These figures are larger version of those shown in section 4.2.