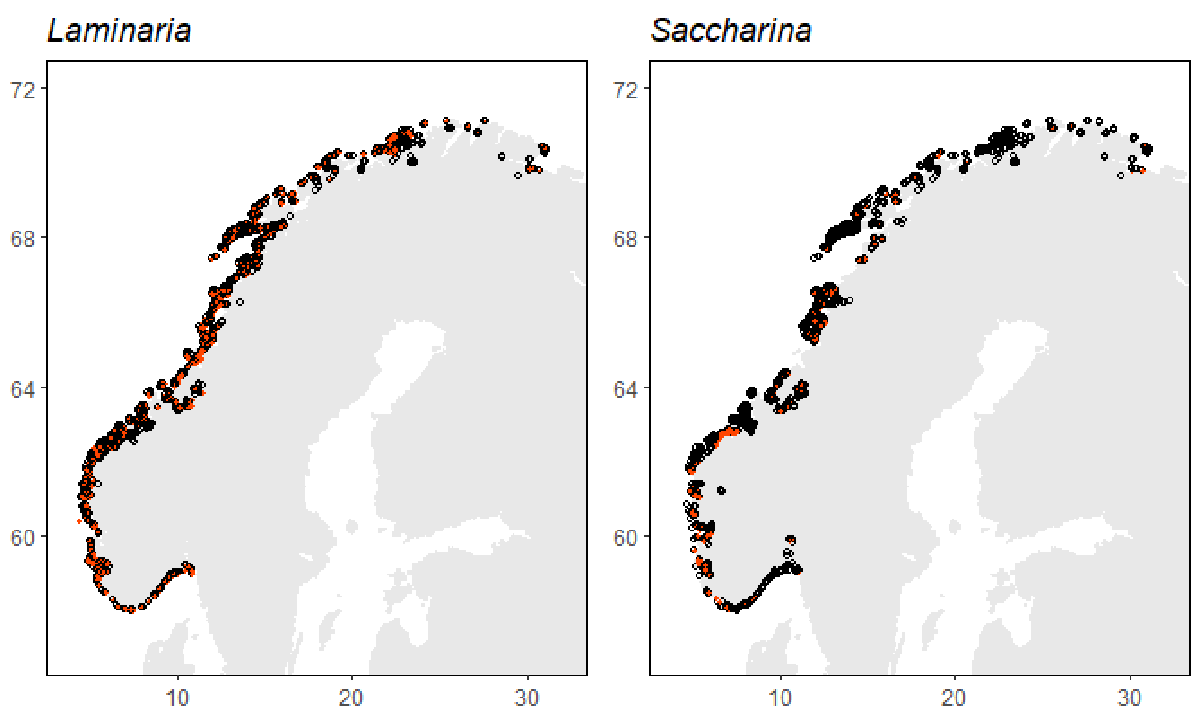
Supplementary Material

Contents

1 Methods: Kelp data

1.1 Norway

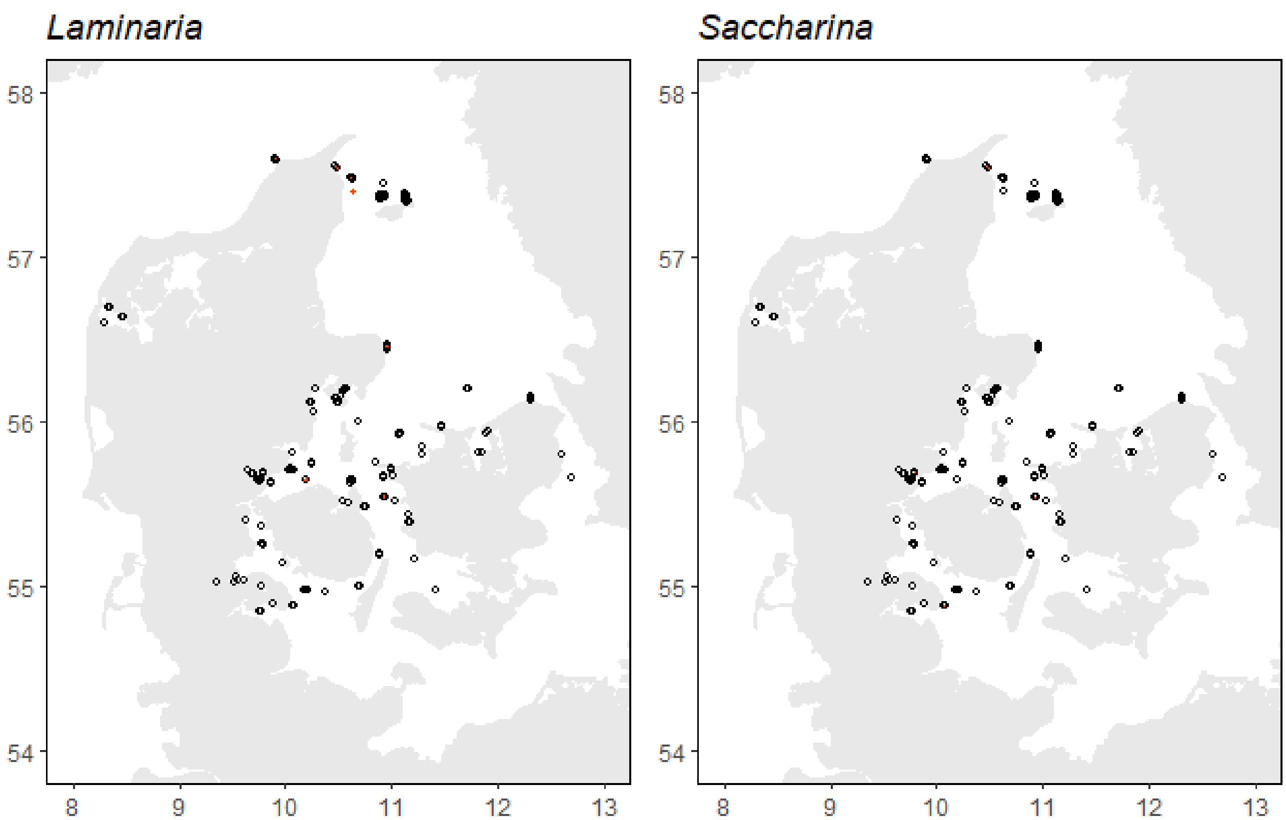
Data for Norway were mostly sampled through the National Program for Mapping Biodiversity - Coast from 2007 to 2018 (Bekkby et al., 2013). The field design was developed with the aim to detect *Laminaria hyperborea* forests, but *Saccharina latissima* was also recorded when found. It should therefore be noted that the design was not developed to capture *S. latissima*'s distribution within its full environmental space. In the field, the abundance of the two species was quantified as absent, single individuals, scarce, moderately dense or dense kelp forest, using underwater camera covering an area of approx. 1 m^2. We defined presence of kelp forest as moderately dense or dense forest (i.e. estimated to be >=5 individuals per m^2 for *Laminaria* and >=7 individuals per m^2 for *Saccharina*, cf. Gundersen et al., 2021). Since data collection followed a pre-defined sampling design aiming to get representative data within areas with likely occurrence of *L. hyperborea*, absences were already included. Data from kelp - sea urchin studies funded by the Norwegian Environment Agency and the Norwegian Research Council were also included. The total number of observations therefore differed between species (*Laminaria*: 7094 absences, 4797 presences; *Saccharina*: 10283 absences, 766 presences) (Supplementary Figure 1). We did not obtain quantitative data (abundances or coverage) of kelp around Svalbard.



Supplementary Figure 1: Presence/absence data for kelp forests (dense or moderately dense forest) in Norway, including *Laminaria* (*L. hyperborea*, left) and *Saccharina* (*S. latissima*, right). Orange: Presence, Black: Absence

1.2 Denmark

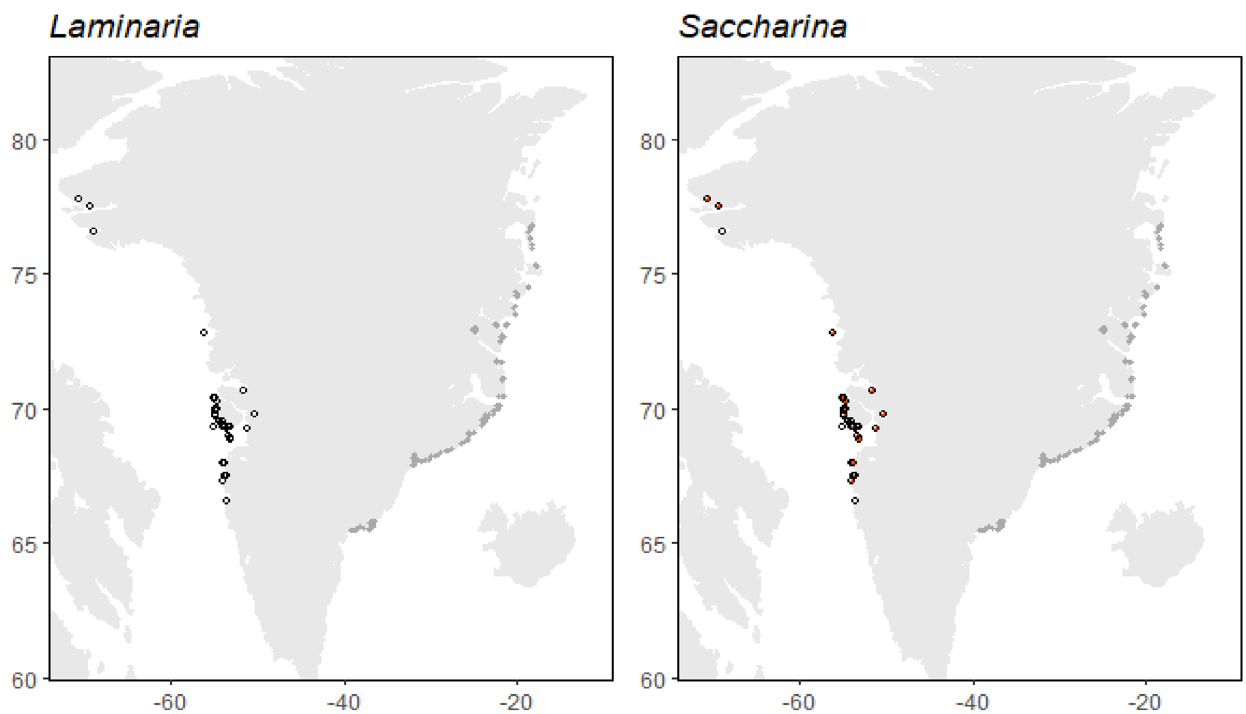
The Danish kelp dataset was collected as part of the Danish National Aquatic Monitoring and Assessment Program and associated activities (2011-2016) according to defined methods for coastal hard bottom vegetation (Høgslund et al. 2014) and vegetation in stone reefs (Dahl & Lundsteen 2018). Specifically, the field work was performed by divers following specific depth intervals along a transect (0-1 m, 1-2 m, 2-4 m, 4-6 m, 6-8 m, 8-10 m and 10-12 m). The divers recorded % cover of hard substrate, and % cover of macroalgae within the hard-bottom areas, in three points of around 25 m^2 each within each depth interval. The data were then averaged per depth interval, and the total cover of macroalgae was defined as % cover of macroalgae multiplied with % cover of suitable hard substrate. The resulting dataset consisted of coverage of 221 different species, including *Laminaria* (encompassing *L. hyperborea*, *L. digitata* and *Laminaria* sp.) and *S. latissima*. At ~4% of the sampling locations (defined by longitude, latitude and depth), repeated measures were given for the same species. In these cases, we used the highest record. There were never records of different species in the same genus at the same location. ≥50% cover was used as a threshold value for kelp forest. We selected combinations of longitude, latitude and depth without any records of the species/genus in question as absences, resulting in 2294 absences and 107 presences for *Laminaria* and 2338 absences and 13 presences for *Saccharina* (Supplementary Figure 2).



Supplementary Figure 2: Presence/absence data for kelp forests (dense or moderately dense forest) in Denmark, including *Laminaria* (*L. hyperborea* and *L. digitata*, left) and *Saccharina* (*S. latissima*, right). Orange: Presence, Black: Absence

1.3 Greenland

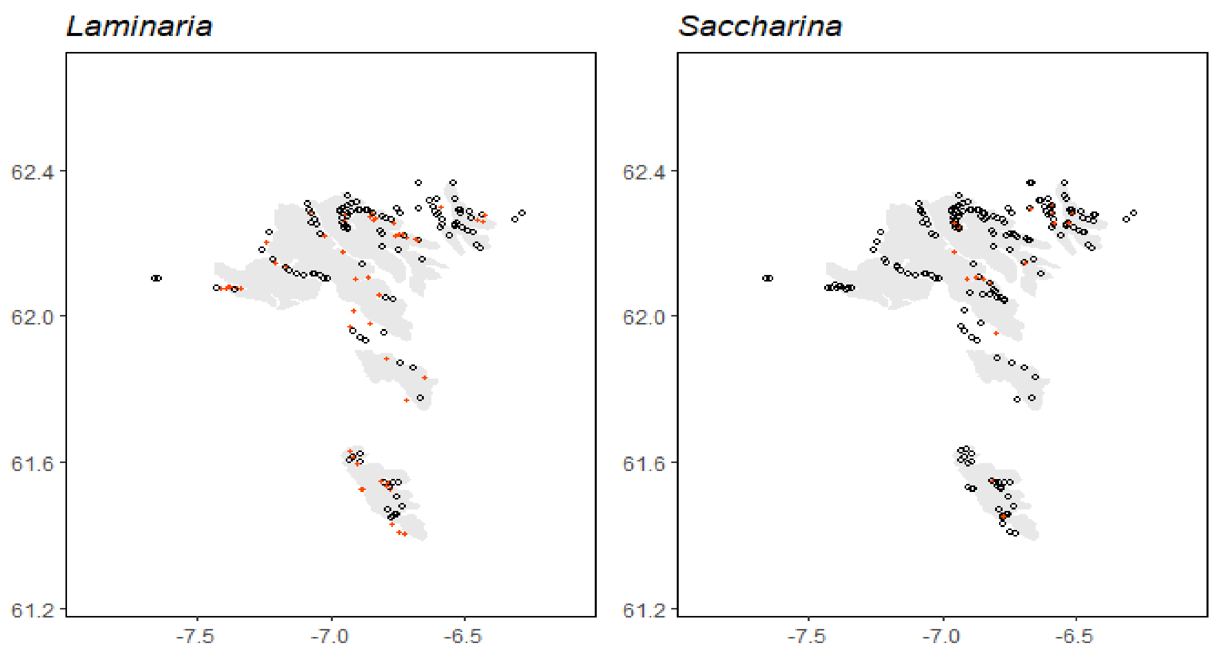
Data from Greenland were available from three different datasets (Supplementary Figure 3). All datasets contained % cover of different kelp and rockweed genera, and >=50% cover of either *Laminaria* or *Saccharina* species was considered as presence of kelp forest, since these genera are represented in all Nordic countries. But note that other species (e.g. *Alaria esculenta* and *Agarum clathatum*) are important contributors to Greenland's kelp forests, possibly leading to an underestimation of kelp forest cover for Greenland using this approach. First, data were available from underwater video transects conducted in 2009 and 2010 on the west coast of Greenland (Krause-Jensen et al., 2012). The transects spanned from the surface and down to max. 43 m, and included percent cover of *Saccharina* along the transects (*S. latissima* or *S. longicruris*, Krause-Jensen et al., 2012), in total 690 recorded absences and 98 presences of kelp forest (>=50% cover). Secondly, data were available from video transects and sampling by SCUBA divers in the outer Disko Bay area in 2009 (Hansen et al., 2013; Krause-Jensen et al., 2019). The video transect data were percent cover of *Saccharina* sp. along transects spanning from 4 m to a maximum depth of 64 m. The SCUBA data were collected at fixed depths of 0, 5, 10, 15 and 20 m along the transects (complementing the video records but not overlapping), and contained percent cover of *S. latissima*, *S. longicruris*, *S. nigripes* (combined as *Saccharina*) and *Laminaria solidungula* (always at < 50 % cover). In total, this dataset resulted in 437 absences and 30 presences of *Saccharina* kelp forest. Thirdly, data on percent cover of *S. latissima* and *Laminaria solidungula* at each 1 m depth down to 50 m were available from underwater video transects conducted at the east coast of Greenland in 2016 and 2017 (Wegeberg et al., unpublished data).



Supplementary Figure 3: Presence/absence data for kelp forests (dense or moderately dense forest) in Greenland, including *Laminaria* (*L. solidungula*, left) and *Saccharina* (*S. latissima*, *S. longicruris* and *S. nigripes*, right). Orange: Presence, Black: Absence. The data for the east coast of Greenland are unpublished, therefore only locations of data are depicted (grey).

1.4 The Faroe Islands

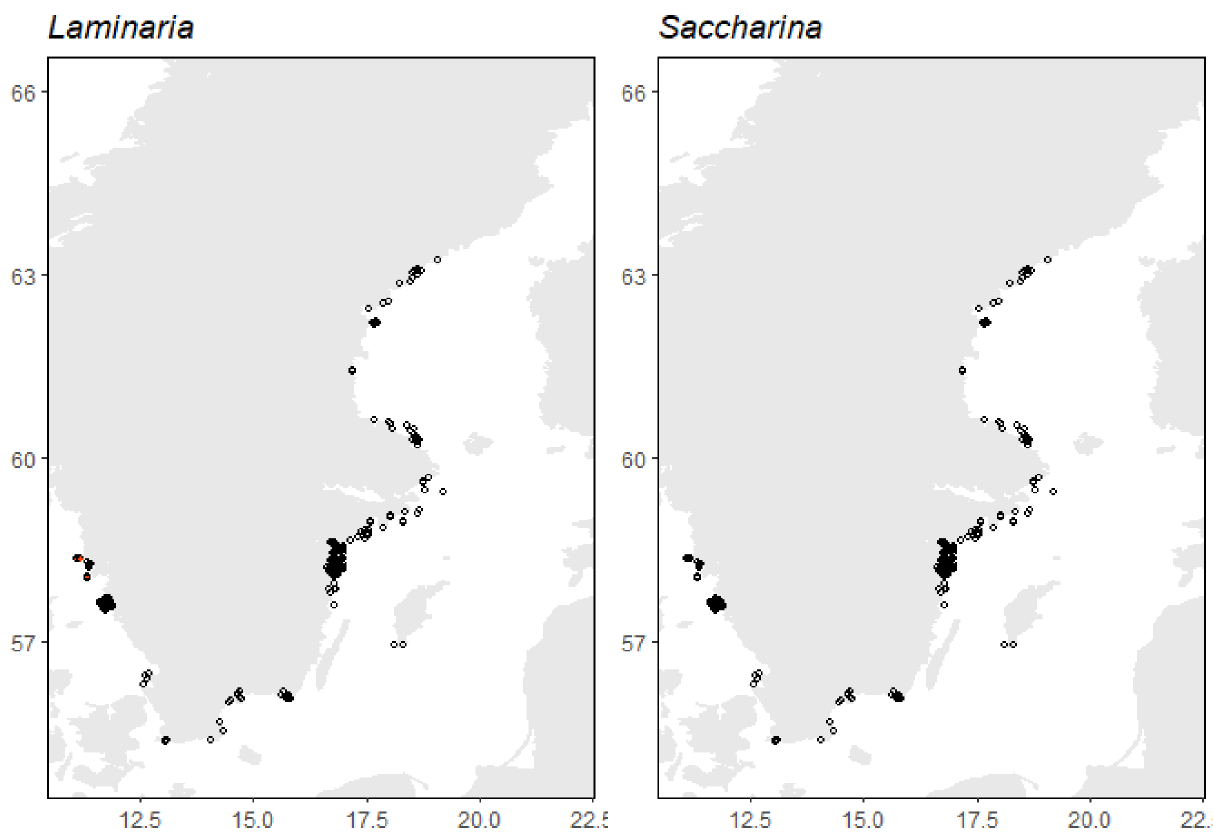
Data from the Faroe Island were available from the BIOFAR program (1996-1997). Separate datasets were provided for the littoral and sublittoral zones, with different classification schemes for species cover. The littoral dataset included records of *Laminaria digitata* and *Saccharina latissima*, among other hard bottom species, and species cover was recorded as >90%, 50-90%, 20-50%, 5-20%, <5%, single plants, or scarcer. We used ≥50% cover as threshold for kelp forest. The sublittoral dataset additionally included observations of *L. hyperborea*, and used the following qualitative categories: "few individuals," "quite common" (<50%) and "very common," where we defined presence of kelp forest as the third category. Littoral field data were collected by visual inspection along 8 m long transects from the upper limit of the dominant species and down to the lowest water level, divided into 10 depth intervals (Bruntse et al. 1999). The abundances of each of 23 dominant species were recorded at the depth interval where the species was most abundant. We therefore lacked information about sampling depth for the littoral, and set this fixed to 0 m (the tidal amplitude was mostly between 0.4 and 2.5 m, Bruntse et al., 1999). The sublittoral dataset was collected by SCUBA diving and provided observations per 5 m depth interval for all species. In contrast to the datasets from Sweden and Denmark, the same species was never recorded more than once at a station, and we therefore assumed that records of *L. hyperborea* and *L. digitata* at the same longitude/latitude in the sublittoral dataset were in fact co-occurring and not records from different, nearby locations. Therefore, to combine data of *L. hyperborea* and *L. digitata* observed in the same station into one record of presence/absence of *Laminaria* forest, we used the highest category (for example, if *L. hyperborea* was recorded as "few individuals" and *L. digitata* as "very common," the station was classified as *Laminaria* forest). Combining the littoral and sublittoral datasets resulted in 174 absences and 123 presences of *Laminaria* and 270 absences and 25 presences of *Saccharina* (Supplementary Figure 4).



Supplementary Figure 4: Presence/absence data for kelp forests (dense or moderately dense forest) in the Faroe Islands, including *Laminaria* (*L. hyperborea* and *L. digitata*, left) and *Saccharina* (*S. latissima*, right). Orange: Presence, Black: Absence

1.5 Sweden

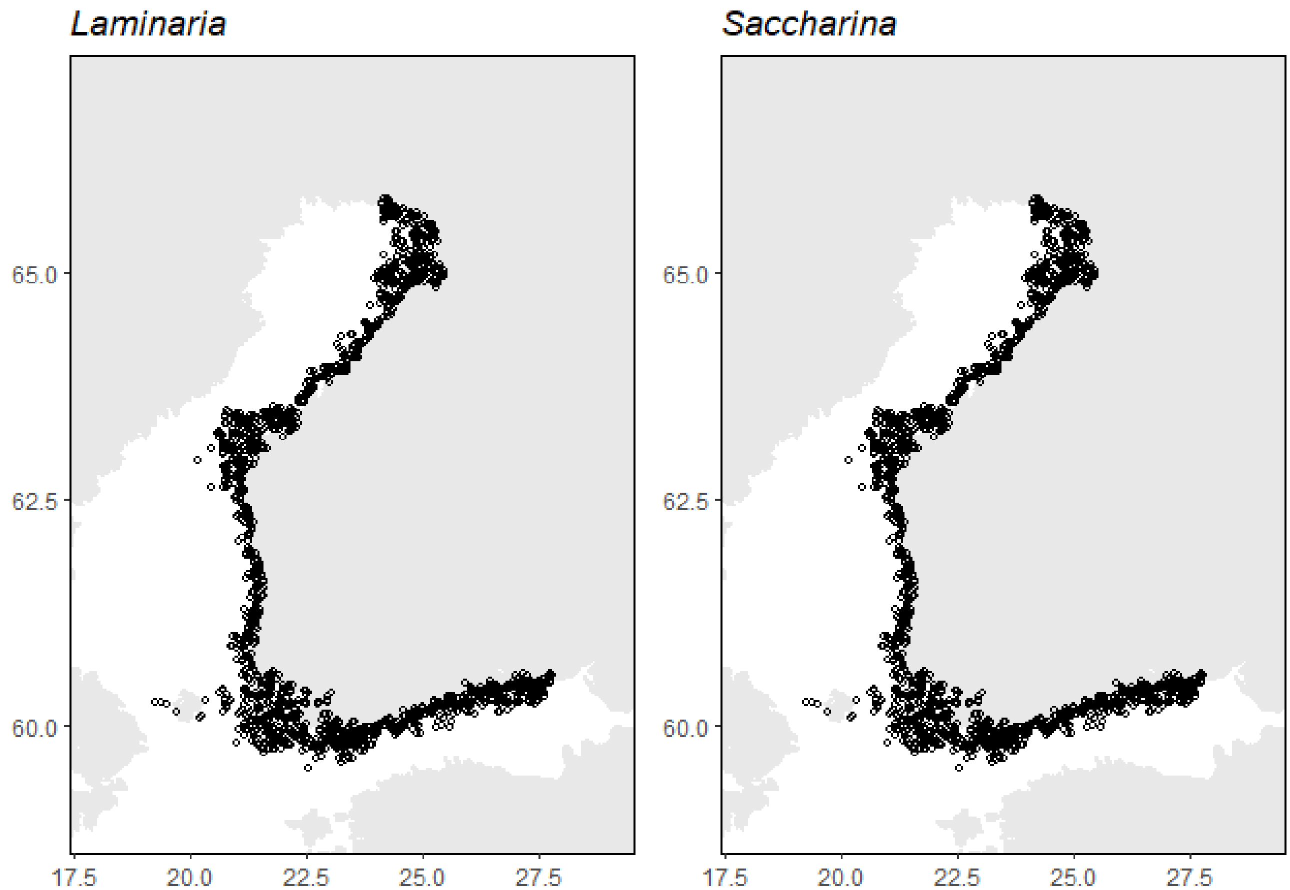
Data for Sweden collected between 2001 and 2015 were available from the Swedish Meteorological and Hydrological Institute (SMHI, downloaded from sharkweb.smhi.se). Data were collected through the National monitoring program for macroalgae, where a SCUBA diver photographs a sample area covering 0.25 m^2 at fixed depths along transects from the surface down to 20 m (max. 32 m), and species cover is determined from the images (Norling & Karlsson 2016). Different quantitative measures were available in the dataset, but only data given as percent cover of the seafloor were used for the analyses, as these encompassed a larger area. The dataset included observations of *Laminaria hyperborea*, *Laminaria digitata* and *Laminaria* sp. that were combined as *Laminaria*, and *Saccharina latissima*. A number of duplicated records were removed from the dataset, and we removed records with missing information about sampling depth as we could not determine if these cases were replicates from the same location or observation along a transect. As for Denmark, when multiple observations of the same species were given at the same station (longitude, latitude, depth), we used the highest recorded coverage, and if different *Laminaria* species occurred at the same station, the coverages were summed (the latter occurred in ~3% of the records). We converted percent cover to presence/absence of kelp forest using a threshold at ≥50% for kelp forest. We set locations (unique combinations of longitude, latitude and depth) with records of other genera but not the ones in question as absences, resulting in 1366 absences and 133 presences for *Laminaria* and 1385 absences and 105 presences for *Saccharina* (Supplementary Figure 5).



Supplementary Figure 5: Presence/absence data for kelp forests (dense or moderately dense forest) in Sweden, including *Laminaria* (*L. hyperborea* and *L. digitata*, left) and *Saccharina* (*S. latissima*, right). Orange: Presence, Black: Absence

1.6 Finland

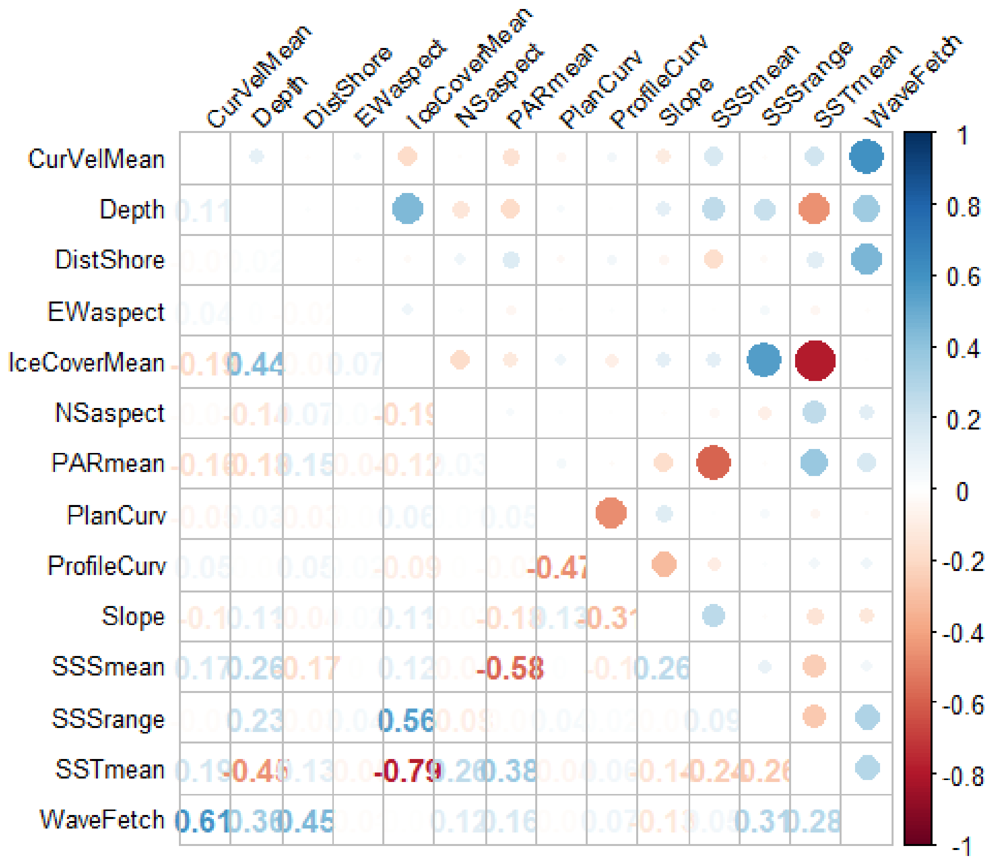
The inner Baltic Sea basins along the Finnish coast are not within the natural distribution range of kelp as the salinity is too low. To include confirmed absences of kelp in this region in the dataset, we used observations of rockweed (*Fucus*) and seagrass (*Zostera marina*) from the Finnish Inventory Programme for the Underwater Marine Environment (VELMU). Since the original dataset was very large, a random sample of 10% (n = 1 707) of the unique locations in the Finnish dataset was used as absences of *Laminaria* and *Saccharina* (Supplementary Figure 6).



Supplementary Figure 6: Kelp data for Finland (only absences)

2 Methods: Correlations between explanatory variables

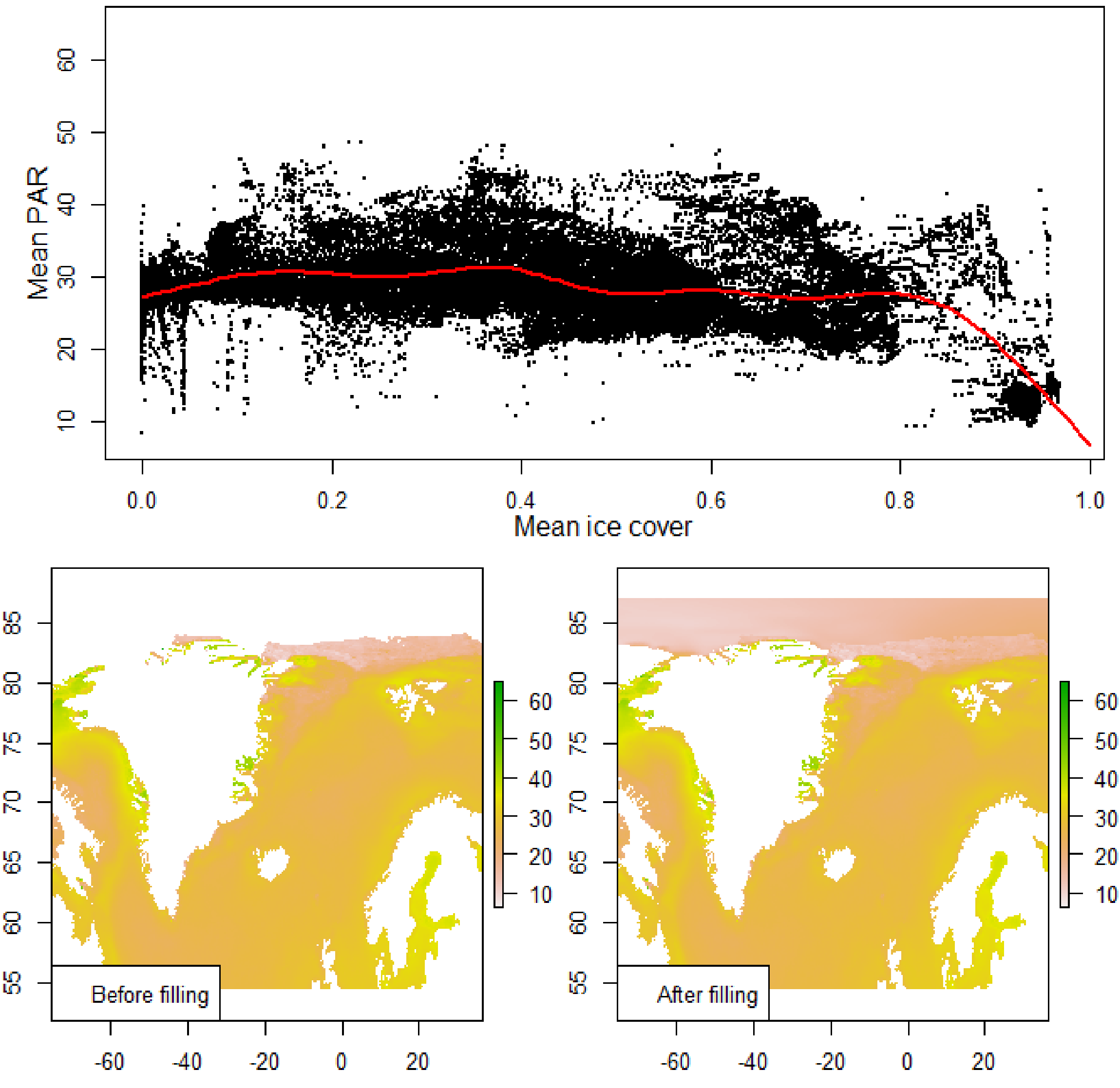
Supplementary Figure 7 shows the pairwise correlation coefficients for all explanatory variables included in the Boosted Regression Tree (BRT) models before model simplification. Correlations were calculated using the environmental data at the locations in the *Laminaria* dataset.



Supplementary Figure 7: Pearson correlation coefficients between environmental variables included in the model fitting, illustrated with the data for *Laminaria*. See the main article for descriptions of the variables.

3 Methods: Missing values of mean Photosynthetically Active Radiation (PAR)

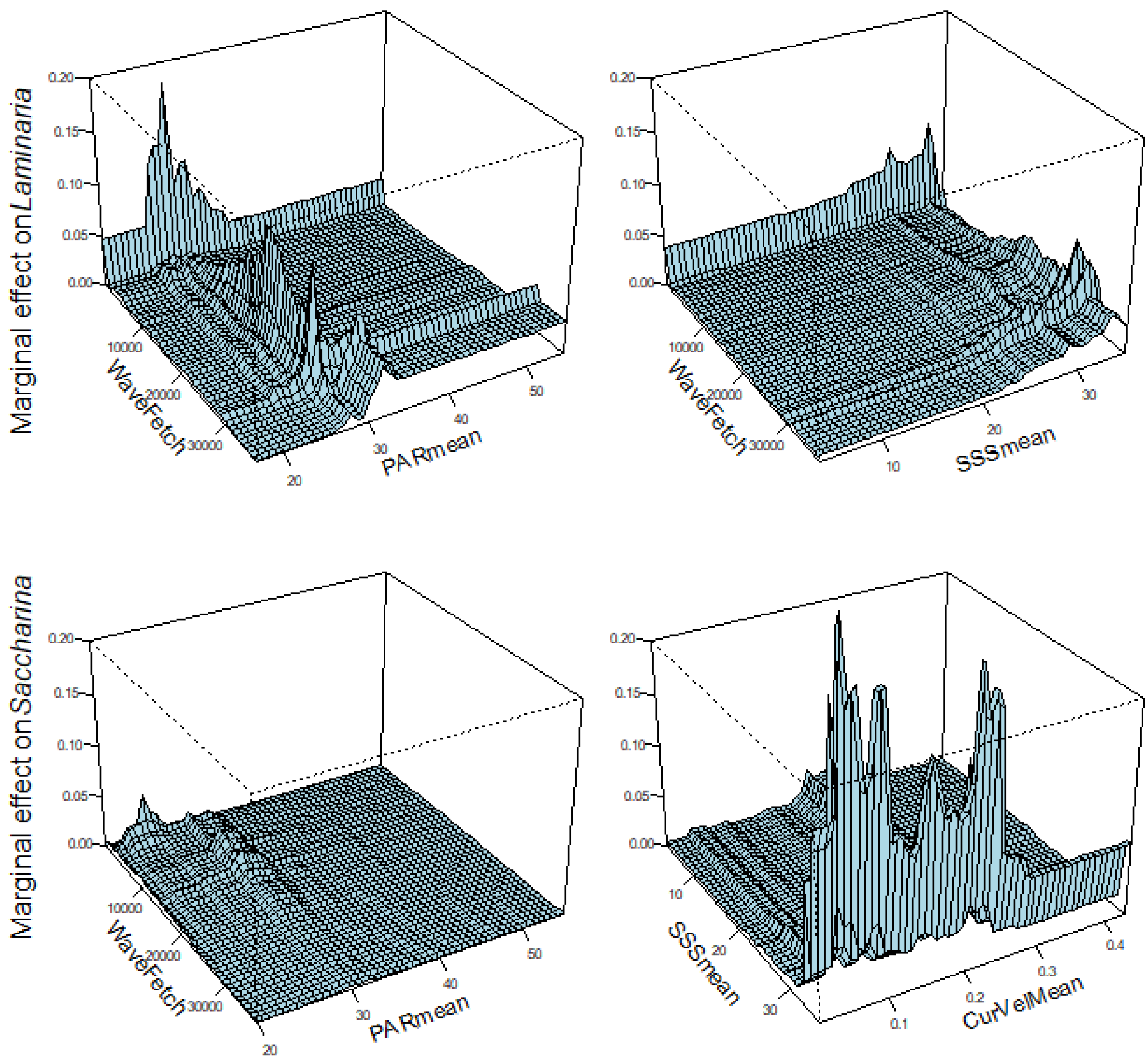
For BRT model predictions, missing values in the predictor variables are assumed to equal the mean of the given variable. To avoid unrealistic model predictions in areas with missing values of mean Photosynthetically Active Radiation (PAR) in the northern part of the study region, we fitted a generalized additive model (GAM) of the additive effect of sea ice concentration on mean PAR (Supplementary Figure 8). Missing values of mean PAR were replaced with the mean PAR predicted as a function of mean sea ice concentration.



Supplementary Figure 8: The upper panel shows the additive effect of mean sea ice concentration on mean PAR in the environmental layers with overlapping data from Bio-ORACLE. Points show the underlying data and the red line shows the additive effect (fitted function). The function was used to fill in values of mean PAR based on sea ice concentration when the former was missing, before running model predictions for the full Nordic region. The middle and lower panels show the PAR mean layer before and after filling of missing data, respectively.

4 Results: Interaction effects

Supplementary Figure 9 shows the partial effects of the two most important pairwise interactions identified in the BRT models. The most important were between mean PAR and wave fetch for both *Laminaria* and *Saccharina*, with the second most important being between mean Sea Surface Salinity (SSS) and wave fetch for *Laminaria* (upper panels), and between mean current velocity and mean SSS for *Saccharina* (lower panels).



Supplementary Figure 9: Three-dimensional partial dependence plots for the two most important pairwise interactions identified in the BRT models for *Laminaria* and *Saccharina*.

5 Results: Model evaluation

For an in-depth evaluation of model performance, we computed confusion matrices and associated statistics using the *caret* package in R (Kuhn, 2020). I.e., we predicted the probability of kelp forest in the locations of actual observations and converted probabilities to presence/absence using the 'threshold' function in the *dismo* package in R. This function determines the cutoff to convert predicted probabilities to presence/absence by comparing observations and predictions using different criteria. We found that maximizing kappa (the proportion of correctly classified units after accounting for the probability of chance agreement) gave the best results across different statistics, and therefore used this criterion when converting predicted probabilities to presence/absence for the full Nordic region. The cutoff probability values were 0.43 and 0.30 for *Laminaria* and *Saccharina*, respectively.

By comparing observed and predicted presence/absence of kelp forest in the locations of observations, we could evaluate how well the models perform in terms of classification of presence (true or false positives) and absence (true or false negatives, Supplementary Table 1). While the comparison is done to observational data used to fit the model and not an independent dataset, only a randomly drawn subset of the data is used in each iteration of the BRT model fitting (70% or 60% of the data at each model iteration for *Laminaria* and *Saccharina* respectively), so we consider comparing model predictions and observations to be reasonable approach. In Supplementary Table 1, 'Accuracy' is the proportion of correctly classified observations (true positives and true negatives), given with 95% confidence interval ('AccuracyLower' and 'AccuracyUpper') and statistically compared to a model always predicting the most common class (absence, 'AccuracyNull'). For both models, the 'AccuracyPValue' indicates that the models perform better than the null model. 'Sensitivity' is the fraction of true positives among all observed positives (predicted presence when presence), 'Specificity' is the fraction of true negatives among all observed negatives (predicted absence when absence), 'Positive prediction value' (or 'Precision') is the fraction of true positives among all predicted positives (correctly predicted presences), and 'Negative prediction value' is the fraction of true negatives among all predicted negatives (correctly predicted presences). 'Kappa' compares the accuracy of the model to a random system.

Most evaluation statistics are high (>0.8), which indicates that the models perform well (Supplementary Table 1). However, the lower sensitivity (0.76) and positive prediction value (0.73) for *Saccharina* indicate that while the model does a good job at predicting absence in locations of actual absences, a substantial fraction of the predicted presences are actual absences (27%) and 24% of actual presences are predicted to be absences. Note that these values are based on comparisons to actual observational data and do not necessarily reflect the accuracy of the predictions to the full Nordic region.

Supplementary Table 1: Confusion matrix (cross-tabulation of observed and predicted classes with associated statistics). For details, see the vignette of Kuhn (2020).

Statistic	*Laminaria*	*Saccharina*
Accuracy	0.96	0.98
Kappa	0.84	0.73
AccuracyLower	0.95	0.98
AccuracyUpper	0.96	0.98
AccuracyNull	0.83	0.96
AccuracyPValue	0.00	0.00
Sensitivity	0.85	0.76
Specificity	0.98	0.99
Pos Pred Value	0.88	0.73
Neg Pred Value	0.97	0.99

6 References

Bekkby, T., Moy, F. E., Olsen, H., Rinde, E., Bodvin, T., Bøe, R., et al. (2013). “The Norwegian Programme for Mapping of Marine Habitats - Providing Knowledge and Maps for ICZMP,” in *Global challenges in integrated coastal zone management*, eds. E. Moksness, E. Dahl, and J. Støttrup (Oxford, UK: John Wiley & Sons, Ltd), 19–30. doi:10.1002/9781118496480.ch2.

Bruntse, G., Lein, T. E., and Nielsen, R. (1999). Response to wave exposure by littoral species in the Faroe Islands. *Fródskaparrit-Annales Societatis Scientiarum Færoensis* 47, 181–198.

Dahl, K., and Lundsteen, S. (2018). Makroalger og hårdbundsfauna på sten- og boblerev. Aarhus Universitet.

Gundersen, H., Rinde, E., Bekkby, T., Hancke, K., Gitmark, J. K., and Christie, H. (2021). Variation in Population Structure and Standing Stocks of Kelp Along Multiple Environmental Gradients and Implications for Ecosystem Services. *Frontiers in Marine Science* 8, 1–19. doi:10.3389/fmars.2021.578629.

Hansen, J. L. S., Hjorth, M., Rasmussen, M. B., Bruhn, A., Christensen, P. B., and Pedersen, P. M. (2013). “Mapping of macroalgae in the Disko West coastal zone,” in *Disko West. A strategic environmental impact assessment of hydrocarbon activities*, eds. D. Boertmann, A. Mosbech, D. Schiedek, and M. Dünweber (Aarhus: Aarhus University), 1–306.

Høgslund, S., Dahl, K., Krause-Jensen, D., Lundsteen, S., Rasmussen, M. B., and Windelin, A. (2014). Makroalger på kystnær hårdbund. Aarhus Universitet.

Krause-Jensen, D., Marbà, N., Olesen, B., Sejr, M. K., Christensen, P. B., Rodrigues, J., et al. (2012). Seasonal sea ice cover as principal driver of spatial and temporal variation in depth extension and annual production of kelp in Greenland. *Global Change Biology* 18, 2981–2994. doi:10.1111/j.1365-2486.2012.02765.x.

Krause-Jensen, D., Sejr, M. K., Bruhn, A., Rasmussen, M. B., Christensen, P. B., Hansen, J. L. S., et al. (2019). Deep penetration of kelps offshore along the west coast of Greenland. *Frontiers in Marine Science* 6, 1–7. doi:10.3389/fmars.2019.00375.

Kuhn, M. (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret.

Norling, K. and Karlsson, J. (2016). *Vegetationsklädda bottnar, västkust*. Version 1:1, 2016-12-07.