

Satellite-Derived Secchi Depth for Improvement of Habitat Modelling in Coastal Areas

AquaBiota Report 2012-02

Authors: Karl Florén¹, Petra Philipson², Niklas Strömbeck³, Antonia Nyström Sandman¹, Martin Isaeus¹ & Nicklas Wijkmark¹

¹ AquaBiota Water Research

² Brockmann Geomatics Sweden AB

³ Strömbeck Consulting AB



AquaBiota
WATER RESEARCH

STOCKHOLM, 25 JUNE 2012

Client:

Performed by AquaBiota Water Research, Brockman Geomatics Sweden AB and Strömbeck consulting AB for Swedish National Space Board and Swedish Environmental Protection Agency.

Authors:

Karl Florén (karl.floren@aquabiota.se)
Petra Philipson (petra.philipson@brockmann-geomatics.se)
Niklas Strömbeck (niklas@strombeckconsulting.se)
Antonia Nyström Sandman (antonia.sandman@aquabiota.se)
Martin Isaeus (martin.isaeus@aquabiota.se)
Nicklas Wijkmark (nicklas.wijkmark@aquabiota.se)

Contact information:

AquaBiota Water Research AB
Adress: Löjtnantsgatan 25, SE-115 50 Stockholm, Sweden
Tel: +46 8 522 302 40
www.aquabiota.se

Quality control/assurance:

Göran Sundblad (goran.sundblad@aquabiota.se)

Distribution:

Free

Internet version:

Downloadable at www.aquabiota.se

AquaBiota Report 2012:02

ISBN: 978-91-85975-18-1

ISSN: 1654-7225

© AquaBiota Water Research 2012



BROCKMANN GEOMATICS
SWEDEN AB



CONTEXT

Introduction	6
1 Objectives	6
2 Field data	7
2.1 Data set 1 – Luode Consulting & Strömbeck Consulting	7
2.2 Data set 2 – AquaBiota	8
3 Image data	9
3.1 Landsat and MERIS	9
3.2 Radiometric resolution	11
3.3 Image pre-processing	12
4 Optical modelling	12
4.1 Algorithm development	12
4.2 Algorithm evaluation	13
5 Analysis & Results	14
5.1 Data extraction	14
5.2 Regression analysis	14
5.3 Secchi depth transfer algorithms and resulting maps	15
5.3.1 Field based Secchi map 100818	15
5.3.2 Model based Secchi map 100818	17
5.3.3 Model based Secchi map 110603	18
6 Evaluation	20
6.1.1 Field based Secchi map 100818	21
6.1.2 Model based Secchi map 100818	22
7 Conclusions from optical modelling	23
8 Habitat modeling	24

8.1	Method	24
8.1.1	The modeling process.....	24
8.1.2	Biological data.....	28
8.1.3	Species modelled.....	30
8.1.4	Manipulation of datasets for the modelling process.....	34
8.1.5	Environmental variables	35
9	Results	43
10	Conclusions habitat modelling.....	50
11	References.....	51
12	Appendix 1	54

INTRODUCTION

Satellite based water quality estimations are usually focused on sensors with a spatial resolution between 300-1000 meters, high sensitivity and spectral properties developed for water targets (SeaWiFS, MODIS and MERIS). However, 300 meters spatial resolution is not considered sufficient for all applications and for all areas, e.g. Lake Mälaren and the coastal zone with narrow bays and a very complex morphology. On the other hand, the spectral and radiometric resolution is insufficient in higher spatial resolution sensors like Landsat TM, ETM+ and SPOT.

Since the launch of Landsat and SPOT in the 70ties, attempts have been made to derive water quality information from these sensors (i.e. 20-30 meters). In most of these investigations, the analysis has been based on satellite data and a limited amount of field data, which might be successful if the time lap between image and field data is small, but usually results are not repeatable and methods less dependent of field data is preferable. We have evaluated another approach and combined the lower resolution MERIS FR data with 30-meters resolution Landsat TM and ETM+ data in order to make use of the best properties of each sensor to generate a high resolution Secchi depth map.

As plants and algae are dependent on light availability, decreased Secchi depth and increased sedimentation due to increased primary production is likely to influence the distribution and abundance of benthic vegetation (Kautsky et al. 1986; Eriksson & Johansson 2003; Eriksson & Johansson 2005). Also, benthic animals, such as blue mussels (*Mytilus edulis* L.) are dependent on particulate matter in the water column as a food source (c.f. Tedengren & Kautsky 1986). Therefore, the Secchi depth, since it is correlated to turbidity (i.e. particles in the water), might also be important for explaining the distribution of benthic animals.

1 OBJECTIVES

The first objective of this study has been to evaluate if Landsat and MERIS FR data can be correlated, which indicates that a combination of the two data types can render information on water quality with higher spatial resolution than initially available from MERIS FR.

Secondly, we have developed, applied and evaluated an algorithm to derive Secchi depth from MERIS FR data and then transferred the result to Landsat data generating a Secchi depth maps with 30 meters resolution.

The final objective is to evaluate if habitat modelling results can be improved using the derived Secchi map as an additional input layer to the model.

2 FIELD DATA

2.1 Data set 1 – Luode Consulting & Strömbeck Consulting

During one day, the 18:th of August 2010, field measurements in flow-through mode (Lindfors et al. 2004) were made in the study area. Briefly described, a pump system, on a boat moving at a typical speed of 10-15 knots, was used to pump water up from about 1-m depth into a series of instruments connected by pressure tubes for measurements of water quality. These measurements included the inherent optical properties spectral absorption $a(\lambda)$ and beam attenuation coefficient $c(\lambda)$ at nine discrete wavelength bands in the visible domain using a WET Labs ac-9 spectrometer, hyperspectral measurements of absorbance (*Abs*, 200-730 nm) with a scan spectrolyser and temperature, salinity, turbidity with a YSI 6660V2 multiparameter sonde. At the same time, the position using a Garmin 72 DGPS receiver was recorded. Data from the ac-9, the YSI 6600V2 and the Garmin DGPS were recorded at a 1-sec interval, while data from the spectrolyser was recorded at a 1-min interval.

At approximately every 10 minutes, the boat was halted to a complete stop for stationary measurements. At each station the Secchi disk depth (*SDD*) was measured with a 25-cm diameter white disk. Measurements were done off the shady side of the boat as the average of two separate readings. No water telescope was used. Altogether this was repeated 41 times during the day and the stations have been plotted in the Landsat image in Figure 1.

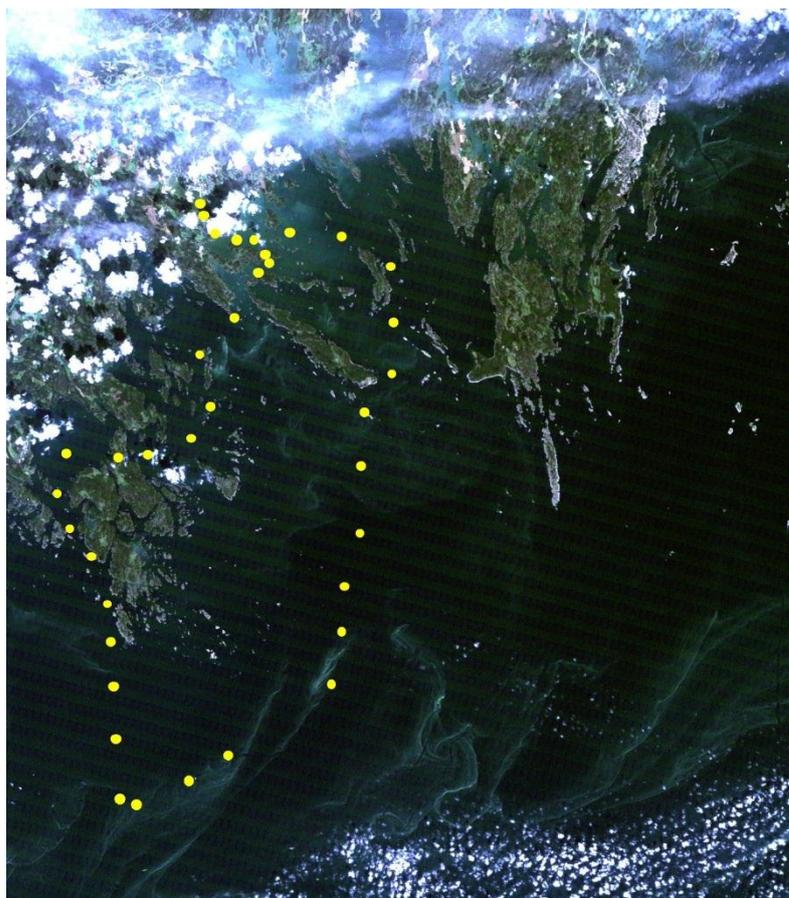


Figure 1. Landsat TM 100818 including field sampling stations.

After the field work was completed, data from the ac-9, the YSI 6600V2 and the Garmin DGPS was merged into the same data file and corrected for time lag induced by the flow-through system in order to match the DGPS positions. This shift typically consisted of 10-10 seconds. Thus, spectro::lyser data with a 1-min time stamp was not shifted but used as is. During this post processing, it turned out that the a measurements of the ac-9 were unusable due to optical filter delamination (WET Labs, pers. communication). Therefore, as proxy data from the spectro::lyser were instead used. As the spectro::lyser is measuring the physical quantity of *Abs* and not *a*, a conventional conversion was done using the relation $a = 2.303 Abs$. Furthermore, as the design of the spectro::lyser not is as optically correct as is the ac-9, a scattering correction based on the s::can standard method was applied.

2.2 Data set 2 – AquaBiota

Between August 4 and 28 2010, 1006 stations throughout the study area were surveyed with a drop-video system (described further under biological data). During the time period August 11-27 2010, Secchi depth was measured in 62 of the drop-video locations using a Secchi disk. 26 of those were measured on August 18, the same day as the field measurements described in section 3.1.

The Secchi depth measurements were used to evaluate the derived Secchi maps, and the drop-video data will be used for the habitat modelling.

3 IMAGE DATA

3.1 Landsat and MERIS

During spring 2010 a list of overpass dates for Landsat TM over Östergötland was made. The focus was on Landsat TM as it was anticipated that the non-functional Scan Line Correction (SLC) in Landsat ETM+ would be a problem. Comparing MERIS and Landsat TM, TM is the limiting sensor with respect to the temporal resolution. In best case it is possible to get an image over our area once a week, compared to 4-5 days per week for MERIS. Based on the available dates for overpass it was decided to focus field work around the 18th of August.

Four images of relatively good quality were collected between June – August 2010: 100615, 100624, 100710 and 100818. The weather was very good on the 24th of June and almost no clouds at all in our area of interest. The other three images are partly cloudy, but was still considered useful for the analysis. The focus of our work so far has been on the data from the 18th of August as field data was collected on that day.

During the analysis of the TM data from 18th of August, it was realised that the existing noise and artefacts in the data could be a problem in the final step of the evaluation, as the patterns most likely would be transferred to the model results (Figure 8). A cloud free ETM+ image from the 3rd June 2011 was therefore included in the analysis to get a comparison with another sensor. However, on 31st of May 2003 the Landsat ETM sensor had a failure of the Scan Line Corrector (SLC). Since that time all Landsat ETM images have had wedge-shaped gaps on both sides of each scene, resulting in approximately 22% data loss. As this image was collected after the SLC failure, a gap fill technique has been applied to the data to make it complete.

MERIS FR data from the 18th of August 2010 and 3rd of June 2011 have been downloaded, processed and used as base image in the analysis together with Landsat TM and ETM+ data from the corresponding dates.

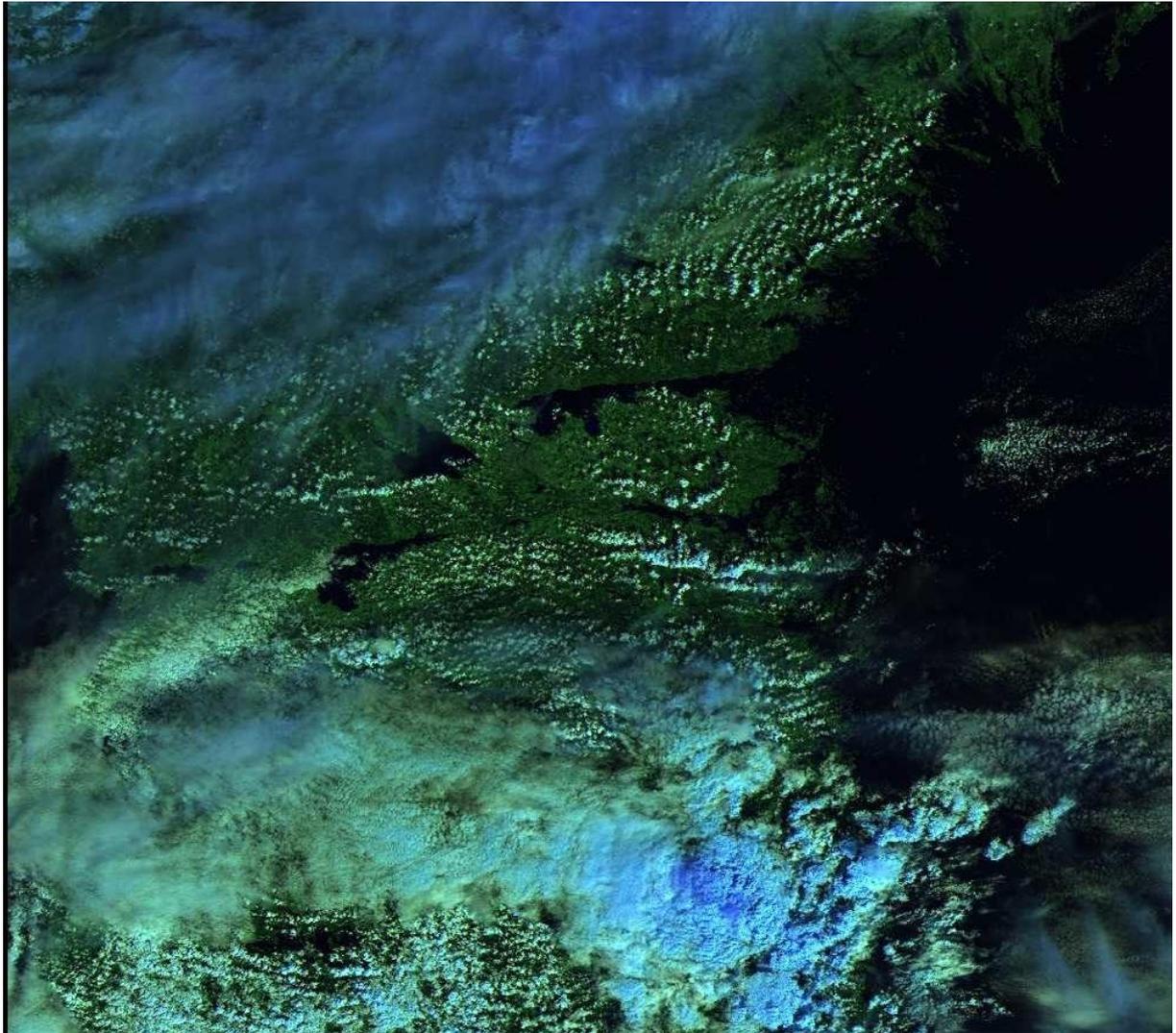


Figure 2. Landsat TM data 100818.

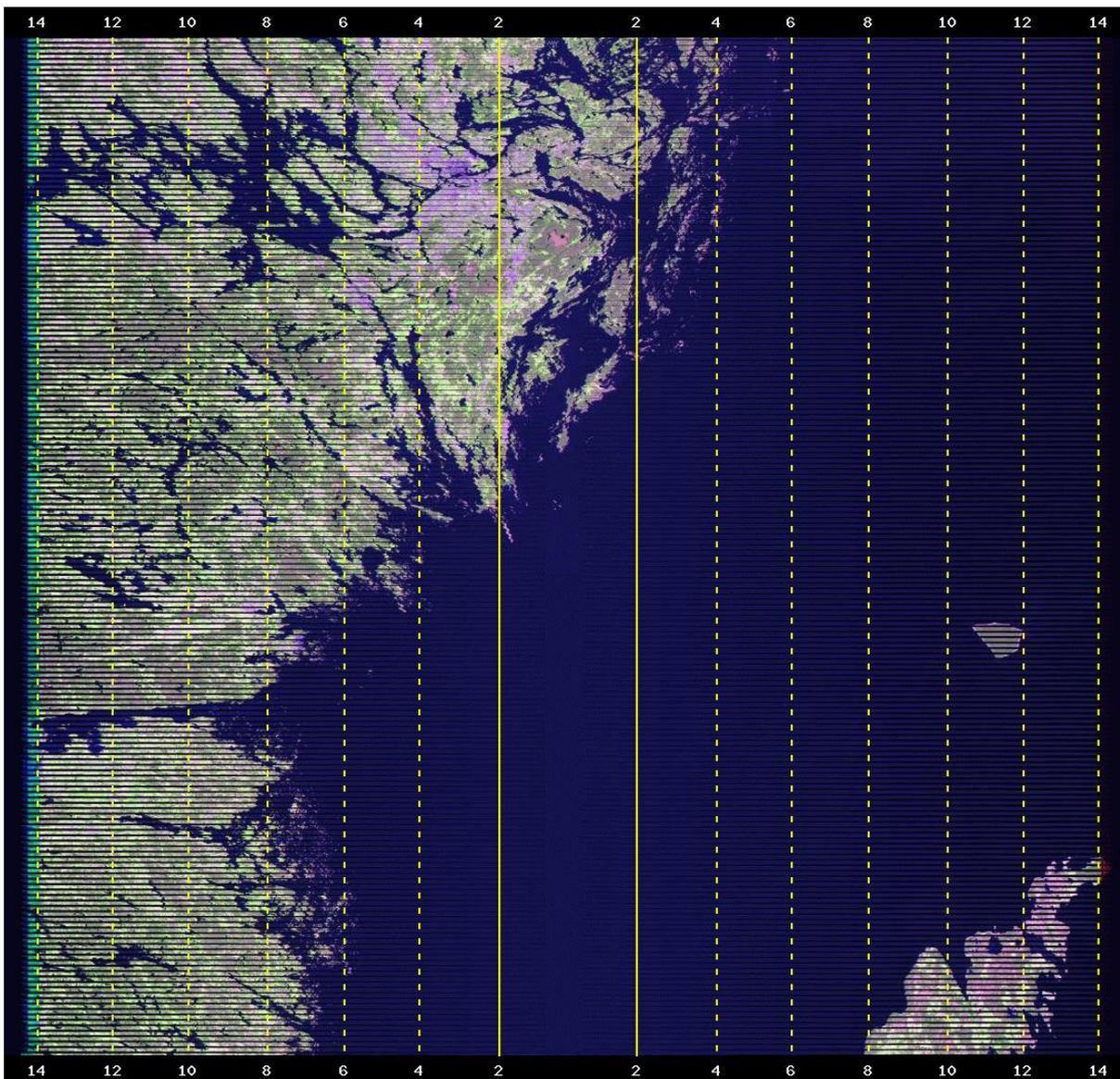


Figure 3. Landsat ETM+ data 110603.Quicklook including data loss due to SLC.

3.2 Radiometric resolution

The problem with low radiometric resolution was mentioned earlier and can be exemplified as follows: The bit-resolution of Landsat data is 8-bits. This means that 256 digital levels are available. However, only a small part of these are used to display the dynamic of the water. The table below gives an example of the variation in the Landsat TM image from 18th of August and represents approximately 95% of the pixels masked as water. In Band 3, only 8 grey levels are used to represent the variation our area, so it is understandable that regression analysis based a few point measurements can fail. See also Figure 6.

Table 1. Available DN levels for water in Landsat TM data (100818).

	DN _{Band1}	DN _{Band2}	DN _{Band3}	DN _{Band4}	DN _{Band5}	DN _{Band7}
--	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------

MIN	45	16	11	6	3	2
MAX	56	23	18	12	12	8
DNs	12	8	8	7	10	7

This will impair the resolution of generated Secchi depth maps, making them discrete rather than continuous. The properties of Landsat ETM+ are better and for band 3 around 20 grey levels are used to represent the variation our area. See also Chapter 6.3.1-2

3.3 Image pre-processing

All images have been geometrically corrected (RT90) before the analysis. Both the TM and ETM+ image was corrected using orthophoto images as base reference (App. 15 meters accuracy). The MERIS FR data from 100818 was corrected using the corresponding TM data as base reference. The MERIS data from 110603 was georeferenced earlier using orthophoto images as base reference. The derived geometric accuracy for both MERIS images was app. 150 meters.

Calibration and atmospheric correction of the Landsat TM data was not considered necessary as “true” reflectance values were not needed in this application. The MERIS FR data used in the analysis corresponds to radiometrically calibrated and atmospherically corrected reflectances (L2) in eight bands and a number of products including chlorophyll, CDOM and TSM. These reflectances and products are a result of FUB-processing of the L1 MERIS FR data.

4 OPTICAL MODELLING

4.1 Algorithm development

For each of the 41 stations described in chapter 3.1, a correlating median value was calculated for each of the parameters, based on the subjectively chosen universally most stable 40-s interval during the stop of the boat. This data was then used to I); calculate a transect of Secchi disk depth, and II); to make an in-situ optical model relating over-water radiance reflectance in MERIS bands to Secchi disk depth.

- I. Good correlation was found both between $c(442)$ and SDD ($r^2 = 0.96$) and turbidity and SDD ($r^2 = 0.97$). However, as the turbidimeter of the YSI 6600V2 is less sensitive than the ac-9, the latter was chosen as most suitable and thus used to calculate a transect of SDD with 1-sec resolution.
- II. The optical properties of a and c were first used to calculate the scattering coefficient b by the conventional formulation $b = c - a$. From b , the backscattering coefficient b_b was calculated by multiplication of a backscattering ratio of 2.1% which is a conventional number for turbid coastal waters (Petzold 1972, Kirk 1994). a and b_b were then linearly interpolated from the nine discrete

wavelengths to hyperspectral data with 1-nm resolution over the range of 400 to 750 nm. Using a and b_b , the semi-analytical model expression of deep water by (Lee et al. 1998) was used for calculating hyperspectral over-water radiance reflectance $R_r(0+)$. The model coefficients μ_0, μ_1 and μ_2 were in an earlier work tuned against the fully physical model Hydrolight© using data from Swedish natural waters and relevant sun elevation and wind speed (SNSB 188/05 and corresponding, Niklas Strömbeck). The hyperspectral radiance reflectance was then integrated into MERIS bands using the MERIS-specific response functions. From these bands, three bands, closely correlating to the spectral positions of the three Landsat-5 TM-bands were selected; $B3$ centered at 490 nm, $B5$ at 560 nm and $B7$ at 665 nm. The three bands were combined two and two in all possible combinations and plotted against measured SDD . The best fit ($r^2 = 0.90$) was found for:

$$SDD = 6.55726 * e^{(-0.8737 * (MER7/MER3))}$$

4.2 Algorithm evaluation

The developed algorithm has been applied to both MERIS images with reasonable Secchi maps as a result. The Secchi depth map derived from the MERIS data collected 110603 is displayed in Figure 4. However, the statistical evaluation based on the independent field data set collected by AquaBiota is focused on the final 30 meter Landsat product. The result of that analysis can be found in chapter 7.

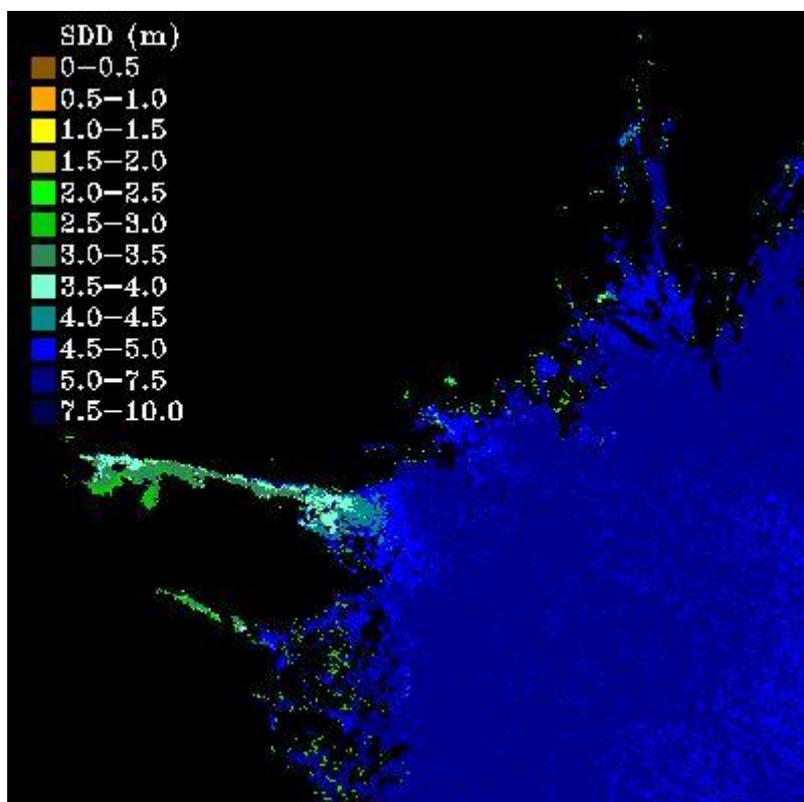


Figure 4. Secchi depth map from MERIS 110603.

5 ANALYSIS & RESULTS

The image analysis consists of data extraction from corresponding areas in TM/ETM+ and MERIS data, regression analysis, development of Secchi depth transfer algorithms and generation of Secchi depth maps.

5.1 Data extraction

Image data from 150 manually defined regions (averages of 50-1200 pixels) were extracted from both images collected 100818. The data was analysed to see if any correlation could be found between MERIS bands and/or water quality products and Landsat TM bands. At a later stage approximately 200 single pixels were extracted and added to the data set, but the new set did not change the relation significantly. The new data mostly represented low Secchi depth areas, which were sparsely represented by the first 150 areas. Additionally, image data corresponding to all field stations were extracted from Landsat data. Both single pixel values and 3x3 averages centered on the pixel containing the sampling station were extracted.

The single pixel approach, for calibration of images, was also used for the data collected 110603. Data corresponding to approximately 500 pixels spanning the whole available dynamic range was extracted for further analysis. No field data is available from 2011.

5.2 Regression analysis

Initially, before the development of the SDD algorithm was finalized, an investigation was made based on field data set 1 and Landsat TM data. 35 stations were used in the analysis. A strong correlation was found between field data and TM3 (Figure 5), which implies that the data should be useful for this application.

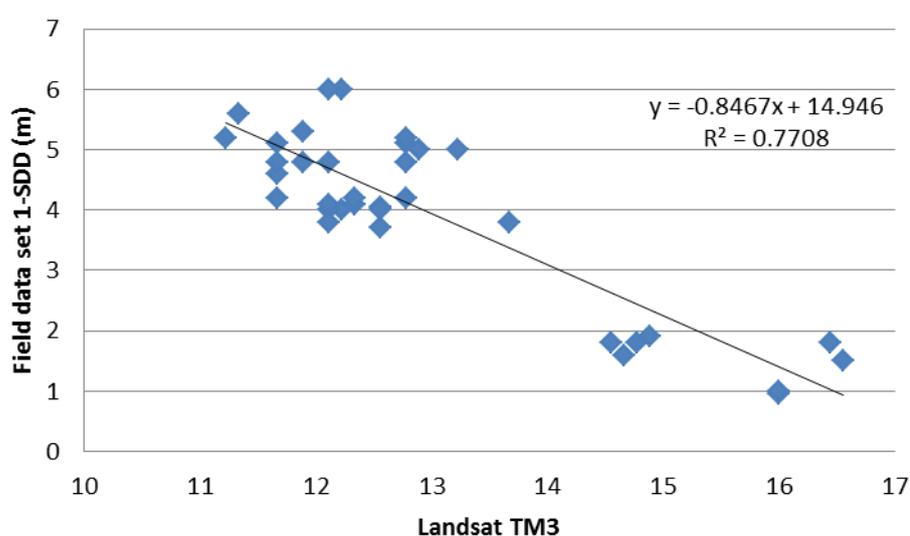


Figure 5. Landsat TM values plotted against MERIS based Secchi depth.

At a later stage, the relation between MERIS and Landsat images were investigated. For both image pairs, correlation could be found between TM/ETM+ bands and MERIS bands and also between TM/ETM+ bands and MERIS products. Additionally, the algorithm derived from the optical modelling (Ch. 5) was applied to the extracted image data and analysed together with Landsat band. The best correlation, 0.62 and 0.6 respectively, was derived from a combination of MERIS-SDD and Landsat band 3 for both image pairs. The extracted and analysed data from 100818 can be seen in Figure 6. Figure 6 also illustrates the limitations of the radiometric resolution of Landsat TM.

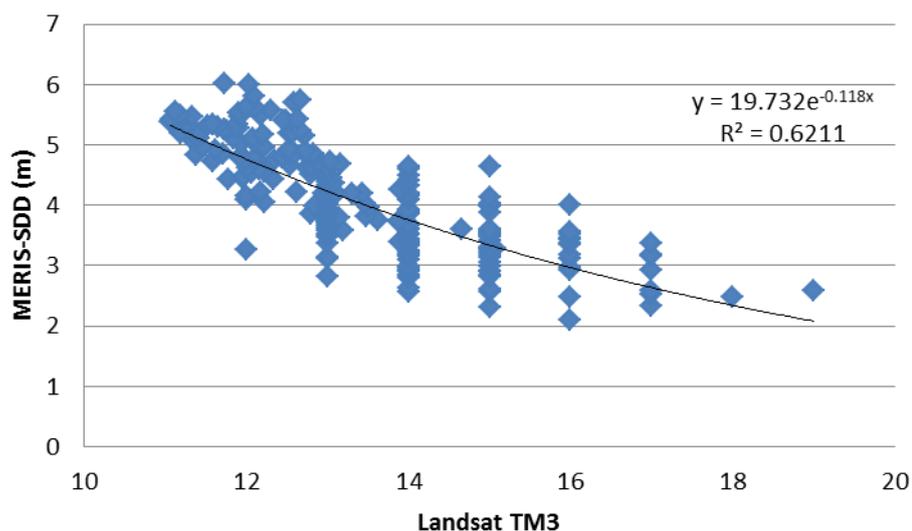


Figure 6. Landsat TM values plotted against MERIS based Secchi depth.

5.3 Secchi depth transfer algorithms and resulting maps

5.3.1 Field based Secchi map 100818

As described above, good correlation (0.77) was found between field data (Set 1) and Landsat TM band 3. The following algorithm for calculation of Secchi depth from TM data was established:

$$SDD = -0.8467*TM3+14.946$$

The algorithm was applied to TM data and the resulting Secchi depth map can be seen in Figure 7 and 8. Figure 8 show a part of Figure 7 around Askö in more detail.

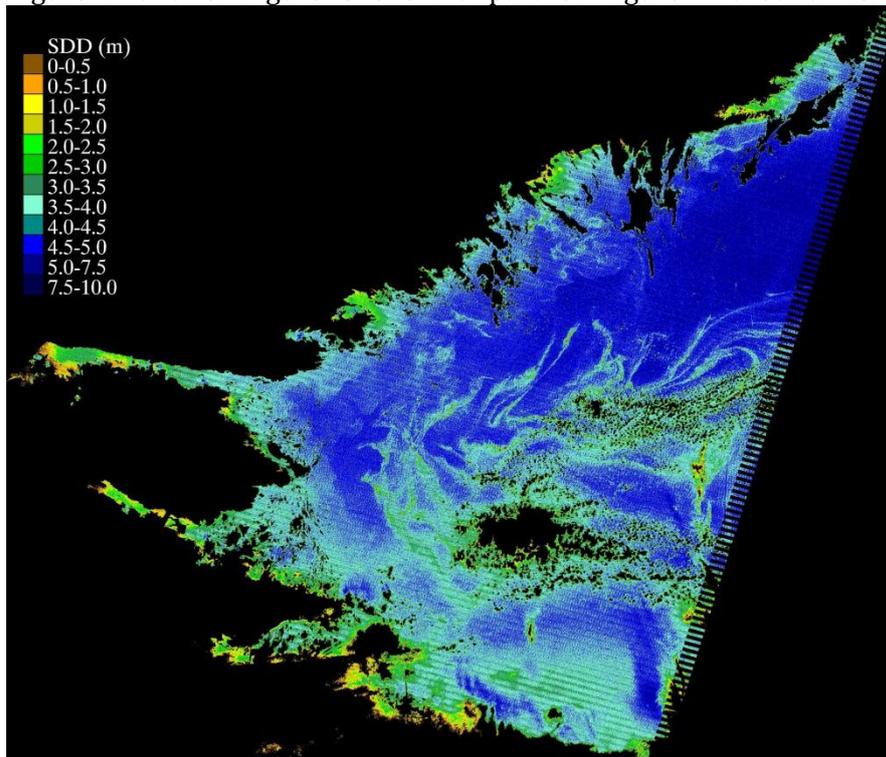


Figure 7. Field based Secchi depth map.

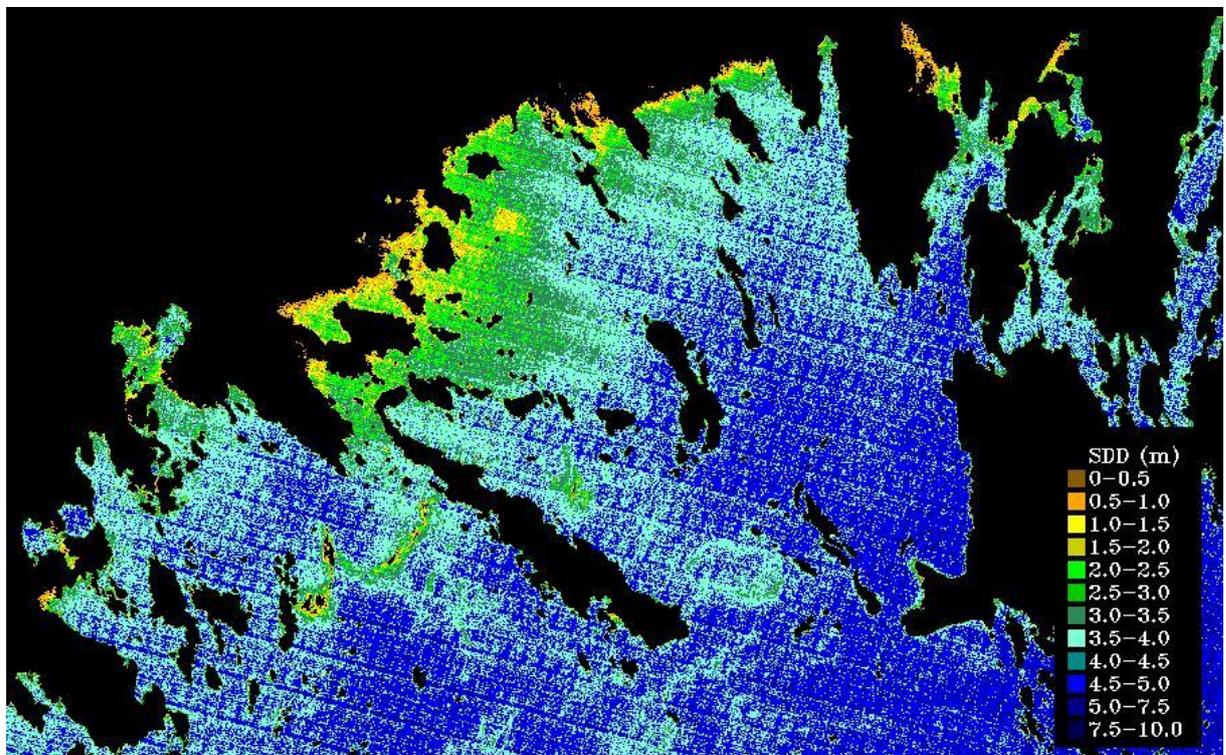


Figure 8. Field based Secchi depth map, Askö.

Due to the limitation based on the radiometric resolution the depth resolution l is 0.85 meters in this map. As one example, the 3-4 meters interval is represented by 3,06 and 3,91

5.3.2 Model based Secchi map 100818

The correlation coefficient between MERIS SDD and Landsat TM band 3 was 0.62 and the following algorithm for calculation of Secchi depth from TM data was established:

$$SDD = 19.732 * e^{(-0.118 * TM3)}$$

The algorithm was applied to TM data and the resulting Secchi depth map can be seen in Figure 9 and 10. Figure 10 show a part of Figure 9 around Askö in more detail.

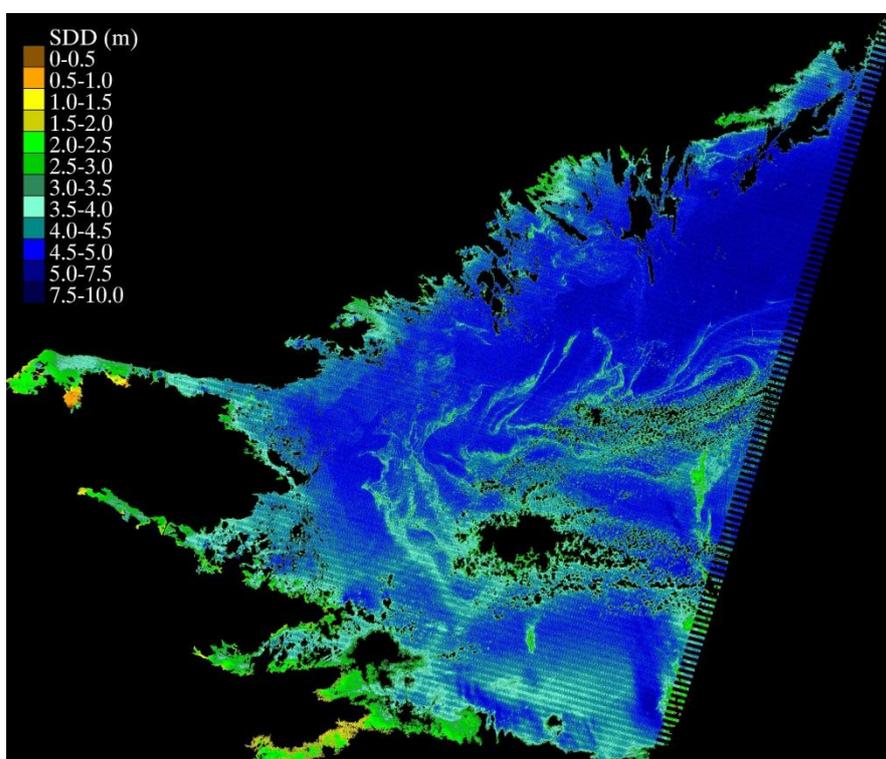


Figure 9. Model based Secchi depth map.

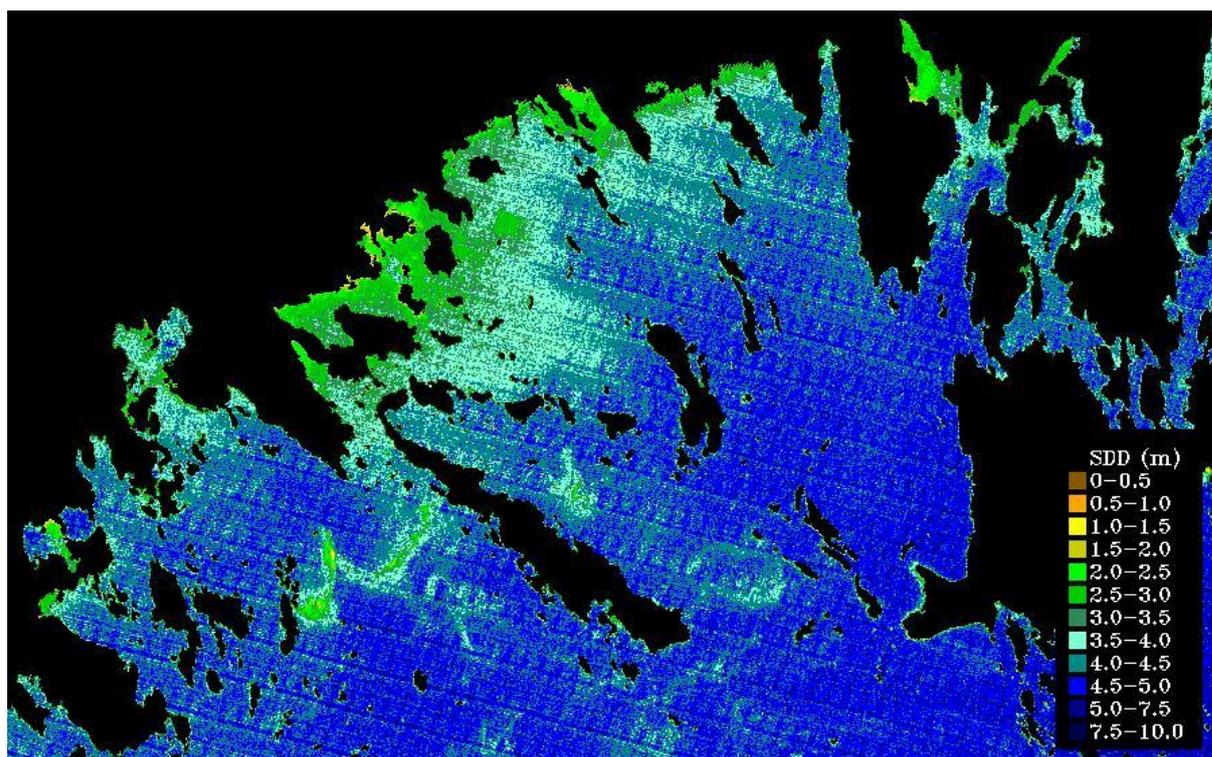


Figure 10. Model based Secchi depth map, Askö.

Due to the limitation based on the radiometric resolution, and with respect to the exponential shape of the algorithm, the depth resolution in the 1-6 meters interval is between 0.2 and 0.8 meters in this map. As one example, the 3-4 meters interval is represented by 3,300847 and 3,713453.

5.3.3 Model based Secchi map 110603

The correlation coefficient between MERIS SDD and Landsat ETM+ band 3 was 0.60 and the following algorithm for calculation of Secchi depth from TM data was established:

$$SDD = 15.814 * e^{(-0.049 * ETM3)}$$

The algorithm was applied to TM data and the resulting Secchi depth map can be seen in Figure 11 and 12. Figure 12 show a part of Figure 1 around Askö in more detail. The lower Secchi depths in the north east corner of the image are not correct. It is caused by haze and not differences in water quality.

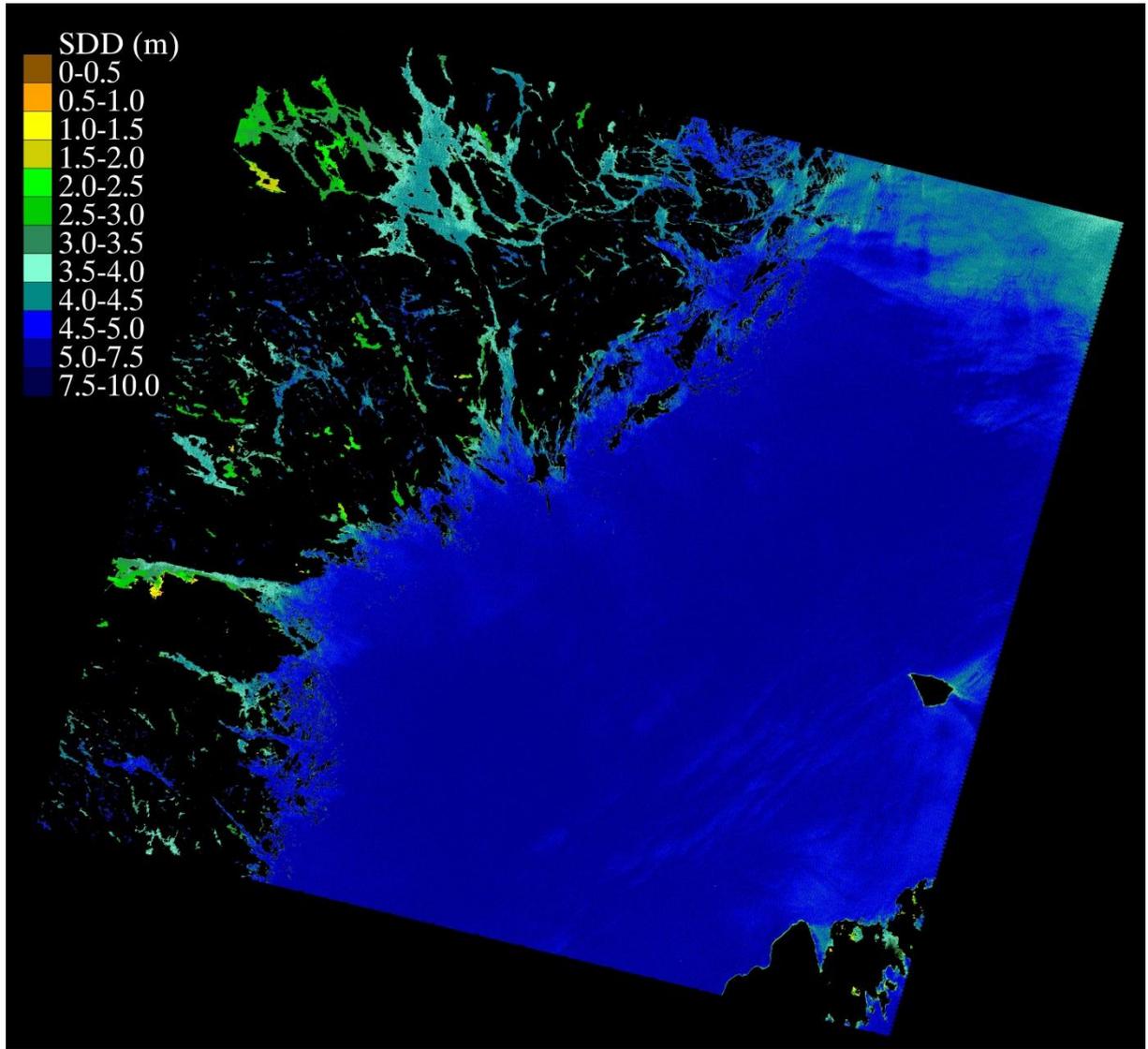


Figure 11. Model based Secchi depth map.

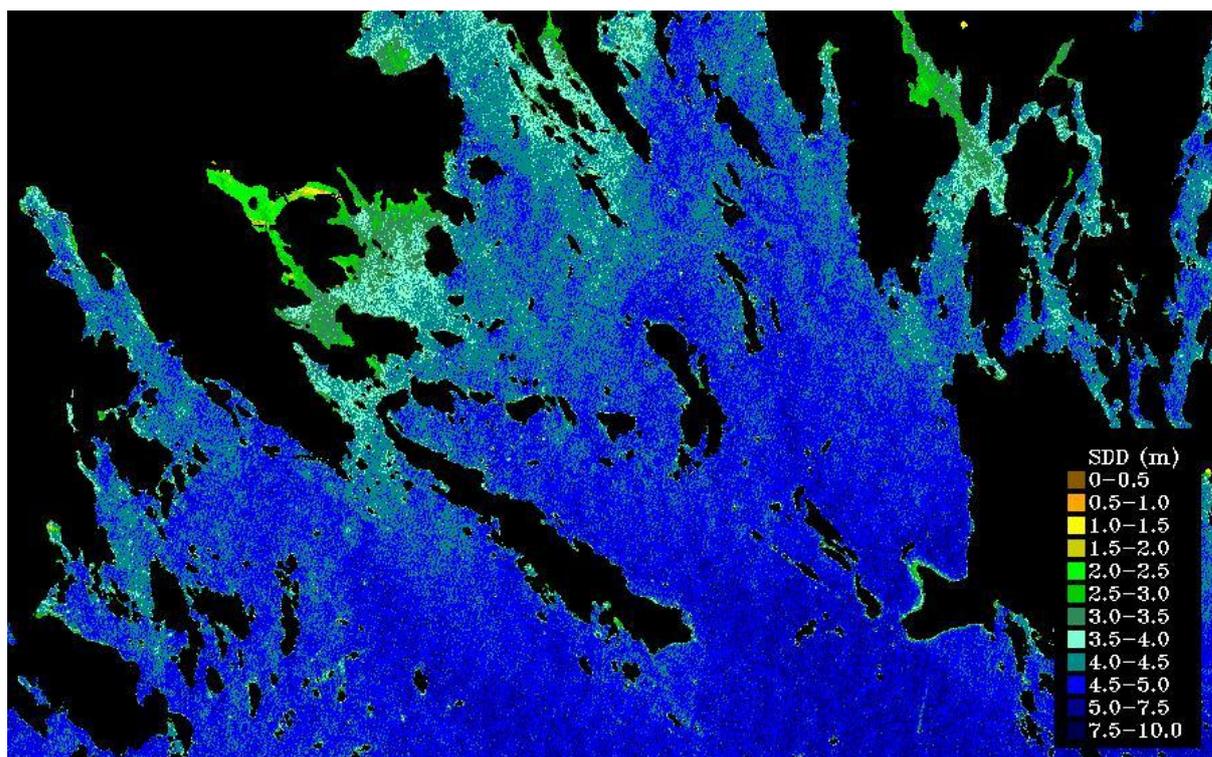


Figure 12. Model based Secchi depth map, Askö.

Due to the limitation based on the radiometric resolution (but still an improvement compared to TM), and with respect to the exponential shape of the algorithm, the depth resolution in the 1-6 meters interval is between 0.06 and 0.3 meters in this map. As one example, the 3-4 meters interval is represented by 3.100832, 3.286878, 3.410909, 3.596956, 3.783002, 3.969048.

ETM+ band 3 do not have the striping problem as TM5, but instead, Coherent Noise (CN) can be seen as a repeating pattern in the data. CN is most visible over dark homogenous regions. CN can arise from many electrical systems on board the satellite, including the power supply, the detector circuitry, and every electrical system in between. The pattern is transferred to the Secchi map, but ETM+ might still be a better alternative compared to TM.

6 EVALUATION

All three maps in chapter 6 have been evaluated using the field data (Set 2) collected by AquaBiota.

6.1.1 Field based Secchi map 100818

As described in chapter 3.2 above, the field data was collected during 17 days. In Figure 13 all available observations have been included and compared to TM results. The correlation coefficient is low ($R^2=0.1339$), RMSE = 1.06 and MAE =0.89.

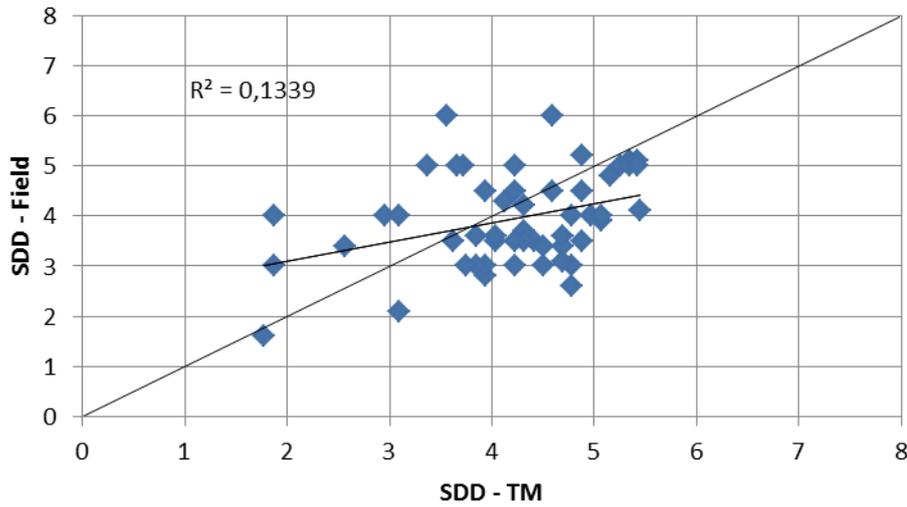


Figure 13. SDD evaluation based on all sampling dates, 0811-0827.

In Figure 14 only observations from the 18th of August has been included. The correlation coefficient is much higher, but the image based Secchi depths are in general a bit higher compared to the field observations. RMSE = 1.03 and MAE =0.90.

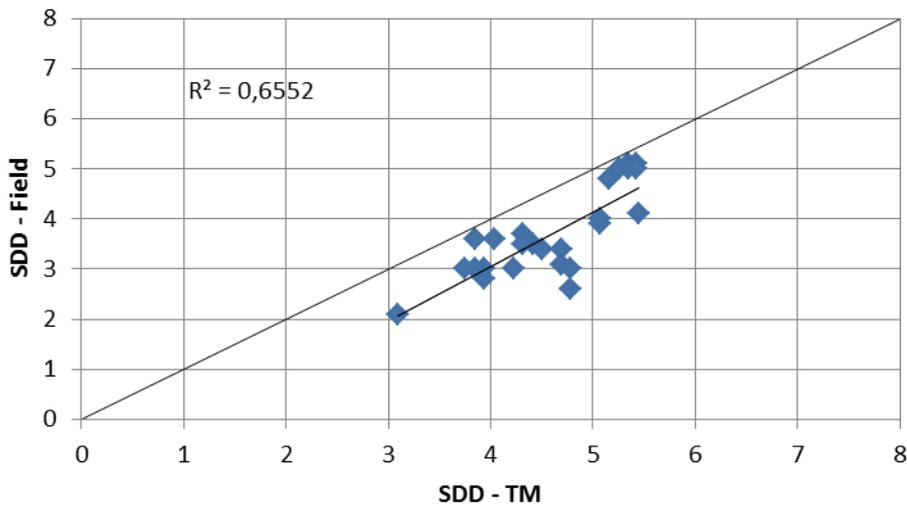


Figure 14. SDD evaluation based on sampling date 0818.

6.1.2 Model based Secchi map 100818

As described in chapter 3.2 above, the field data was collected during 17 days. In Figure 15 all available observations have been included and compared to TM results. The correlation coefficient is low ($R^2=0.14$), RMSE = 1.02 and MAE = 0.86.

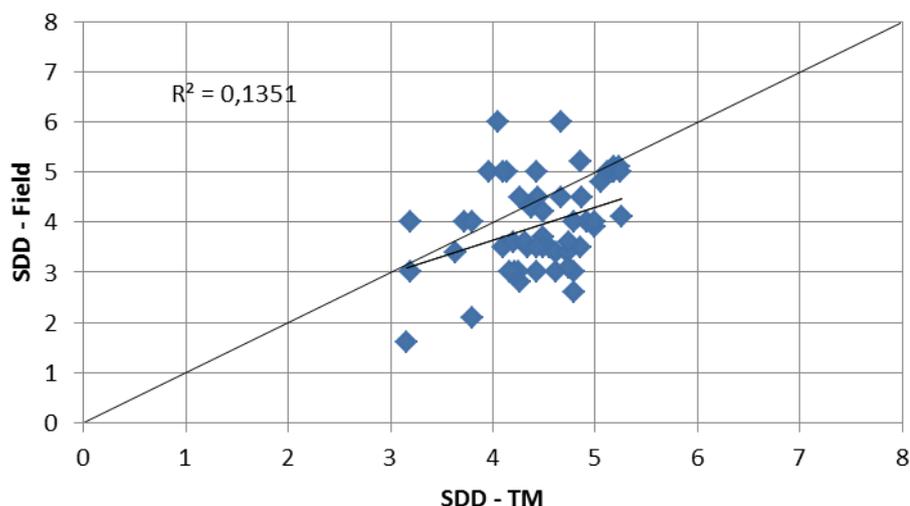


Figure 15. SDD evaluation based on all sampling dates, 0811-0827.

In Figure 16 only observations from the 18th of August has been included. The correlation coefficient is much higher ($R^2=0.67$), but the image based Secchi depths are in general a bit higher compared to the field observations. RMSE = 1.15 and MAE = 0.99.

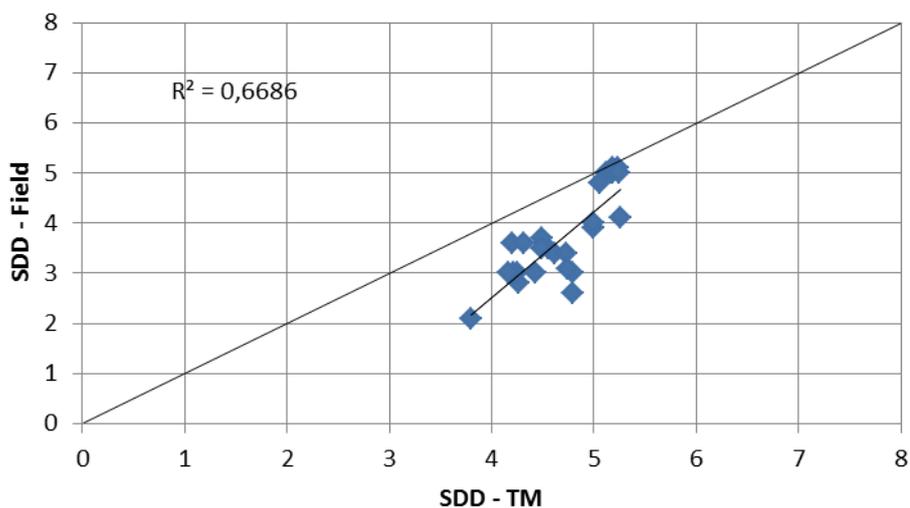


Figure 16. SDD evaluation based on sampling date 0818.

7 CONCLUSIONS FROM OPTICAL MODELLING

The results are very positive and show the possibility to derive high resolution Secchi depth maps in the coastal zone. The innovation is that Landsat data, which hardly would produce such results separately, can generate representative maps in combination with MERIS fr data.

The evaluation shows that the correlation between independent field data and the generated maps is good if the time lap between data collections is limited, but that there is an offset in absolute level. It should be noted though, that when the TM based map seen in figure 9-10 is “evaluated” using field data set 1 (Luode & Strömbeck), the same offset cannot be seen, which at least indicates that the Secchi depth transfer routine seems to work well.

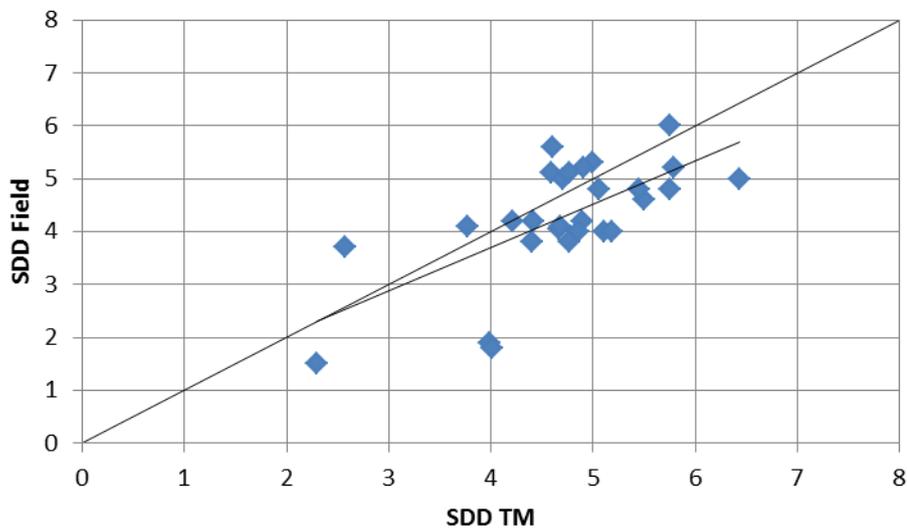


Figure 17. SDD comparison between field data set 1 and the model based Secchi map.

8 HABITAT MODELING

As habitat structure is strongly correlated to environmental variables, abiotic factors such as salinity (Remane & Schlieper 1971; H. Kautsky 1995), light (Krause-Jensen et al. 2007), wave exposure (Kiirikki 1996; Eriksson & Bergström 2005; Isæus 2004) and substrate (Kautsky & van der Maarel 1990) can be used to predict the distribution of phytobenthic communities. Today predictive habitat distribution models are a widely used tool in both ecology and in nature conservation and habitat management. By calculating statistical relationships between species and environmental variables species distribution maps can be generated. However, in order to apply the distribution models in geographical space, full coverage maps of the environmental variables are needed. Satellite derived measurements may provide a quick and cost-efficient collection over large geographical areas.

Light availability is one factor that determines the distribution and abundance of benthic vegetation, and therefore information on the spatial variability of Secchi depth most likely should increase the predictive capacity of species distribution models. As Secchi depth is related to eutrophication, our assumption was that we would get more fine-tuned information on environmental differences e.g. between areas that are otherwise similar, but with different inflow regimes. We have modeled the occurrence of different species of vegetation as well as blue mussels (*Mytilus edulis* L.) as a response to Secchi depth in combination with other environmental variables, such as wave exposure, depth, seafloor slope and salinity. By making two models for each response, one with and one without Secchi depth and keeping the other variables constant we have been able to evaluate the contribution of this variable.

8.1 Method

8.1.1 The modeling process

Modeling can include everything from simple causalities to advanced statistical calculations. In this context the purpose is spatial predictions based on statistical modeling by relating the distribution of species to relevant environmental variables. From empirical data the spatial distribution of a response variable is calculated. The response variable can be a single species of algae, a type of substrate or a certain habitat. The modeling process is shown in Figure 18.

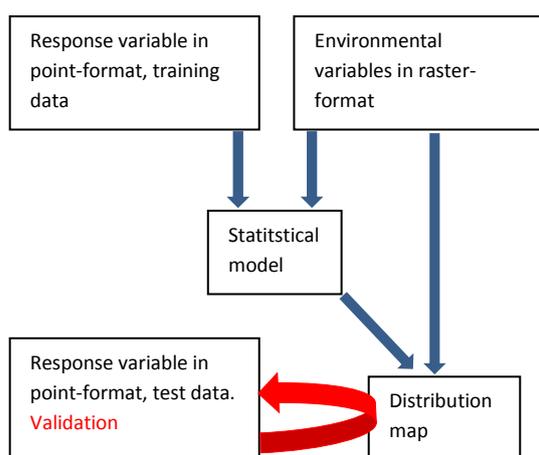


Figure 18. The modeling process.

8.1.1.1 Step 1 – Model

8.1.1.1.1 Development of model

In the first step the relationships between the value of the response variable (presence or absence) and the values of the environmental variables (predictors) are statistically calculated. Predictors like depth and substrate can be measured simultaneously with the response variable. Other environmental variables such as wave exposure and bottom topography, that are hard to measure in the field, are extracted from the raster layers in GIS. In the modeling process a measure of each predictor’s importance is generated. Depending on the response variable different predictors may be important in explaining the distribution of that response variable.

In this study GAM (Generalised Additive Models) were used in the statistical analysis. This modeling technique is widely used and has been performing well in many studies in the coastal areas of the Baltic Sea (Sundblad et al, 2001; Snickars et al, 2010; Florin et al, 2009). The MGCV package in R was utilized for relating the distribution of species to the environmental predictors. The presence or absence of the response was described by smooth functions of the predictors using penalized regression splines (Wood & Augustin, 2002). The smoothness/wiggleness of the response is determined by how many degrees of freedom the smooth function is given. Here we used 4 degrees of freedom as a maximum starting point as it allows a suitable tradeoff between model fit and ecological relevance. Penalization of the response is regulated by a ‘gamma value’ which is set at 1.4 as default. We have found that this default value performs well and it was therefore used (Bonus funded project PREHAB, www.prehab.gu.se). The gamma value regulates to what extent the model fitting process should try to reduce the response to a horizontal line, which then functions as a semi-automatic model selection

process. Predictors with poor fit to the response will be penalized to zero and removed from the model, while predictors that fit the data well will be kept. An illustration of two different gamma values is given in Fig 19.

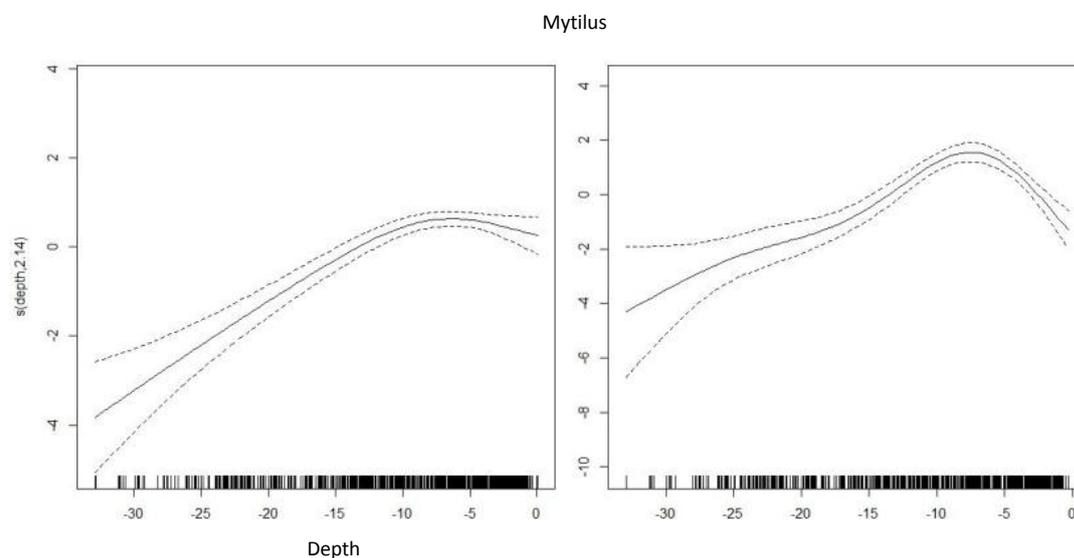


Figure 19. The response of Blue mussel (*Mytilus edulis*) to depth with the same initial degrees of freedom but penalized in different ways. Both response curves had 4 degrees of freedom initially but these were reduced to 2.14 in the left panel and XX in the right panel. The left panel was penalized using a gamma value of 10 and the right curve a gamma value of 1,4.

8.1.1.1.2 Collinearity

A potential problem using regression models is the existence of correlation between environmental variables (Zuur et al 2009). An example in the Baltic could be salinity and distance to mainland since rivers from the mainland supply the sea with fresh water. Collinearity can mask important relationships between variables and confuse the regression process. Before starting the modelling process all environmental variables were checked for collinearity. The statistical dependence between the predictors was calculated using Spearman's rank correlation coefficient.

A statistical tool called VIF (variance inflation factor) was also used for pinpointing variables with a potential correlation problem.

8.1.1.1.3 Model performance

Before a model can be accepted its quality and stability must be evaluated. Measurements of how much of the variation in the training dataset that can be explained by the model are generated in the modeling process. Poor model performance is often due the lack of important predictors, a low number of observations or overfitting. An overfitted model will adapt, not only to the variation in the predictors, but also to other factors or to chance. Poor models should not be used for making predictions. A measure called deviance explained by model has been used for model evaluation. By looking at Chi square values the importance of the predictors was analysed. For every model Chi square values for each predictor were divided with the total Chi square value in that model. This gives a measure of the predictors' influence to the model.

8.1.1.2 Step 2 – Prediction

In the second step the model is used together with raster layers of all environmental variables explaining the variation to make a prediction. The model is run for each raster pixel which contains values for all predictors influencing the model. The resulting raster layer (prediction) shows the expected distribution of the response. Flaws in the raster layers describing the environmental variables will be transferred to the prediction. An important predictor will transfer more of its flaws to the prediction and the quality of this predictor is therefore very important.

8.1.1.2.1 Validation of prediction

The quality of the prediction should be assessed using external data that has not been used to build the model. In this project all predictions have been validated with external data where depth values were extracted from the depth raster layer. This means validation of the final prediction map. Predictions resulting in probability of occurrence are usually validated with a measure called AUC (Area Under Curve, Fielding & Bell, 1997). An AUC value of 1 means that the model has classified all presences and absences correctly. An AUC value of 0.5 indicates a prediction no better than chance. An AUC value of 0.8 means that a random sample from the presence stations will, in 80 % of cases, have a higher prediction value than a random sample from the absence stations. According to a recommendation by Hosmer & Lemeshow (2000) the following terms are used in this report.

AUC-value	Quality
0.9-1	Excellent
0.8-0.9	Good
0.7-0.8	Intermediate
0.5-0.7	Poor

AUC is a combination of 2 measures, sensitivity and specificity. Sensitivity is a measure of the models ability to classify a response's true presence whereas specificity measures the ability to classify a response's true absence.

The AUC-value is a good indicator of the overall quality of the prediction as long as the validation stations have a good spread in the area of interest. However there could be local defects in the map not picked up by the AUC-value. Therefore all maps should also be judged by experts with good local knowledge.

8.1.2 Biological data

The drop-video data was collected using an underwater drop-video system (Sea-Trak 1.06 from SeaViewer Cameras Inc.) connected to a GPS receiver (Garmin GPS 60). Bottom conditions and benthic communities were inspected using the underwater video. When the recording was started, a waypoint was taken with the GPS receiver and the depth from the echo sounder was noted in the field protocol. An area of approximately 5x5 meters was surveyed. If the depth varied within the survey area the average depth was estimated. The observation was interpreted directly onboard the vessel by watching the screen and filling in the protocol. If there are any uncertainties, the recorded videos can be used in a later stage to ensure the quality of the observations by discussing with colleagues and/or other experts. Videos were recorded as digital .avi files to portable DVR, together with the GPS position at each station. In total 1006 drop-video stations were visited (Figure 20).

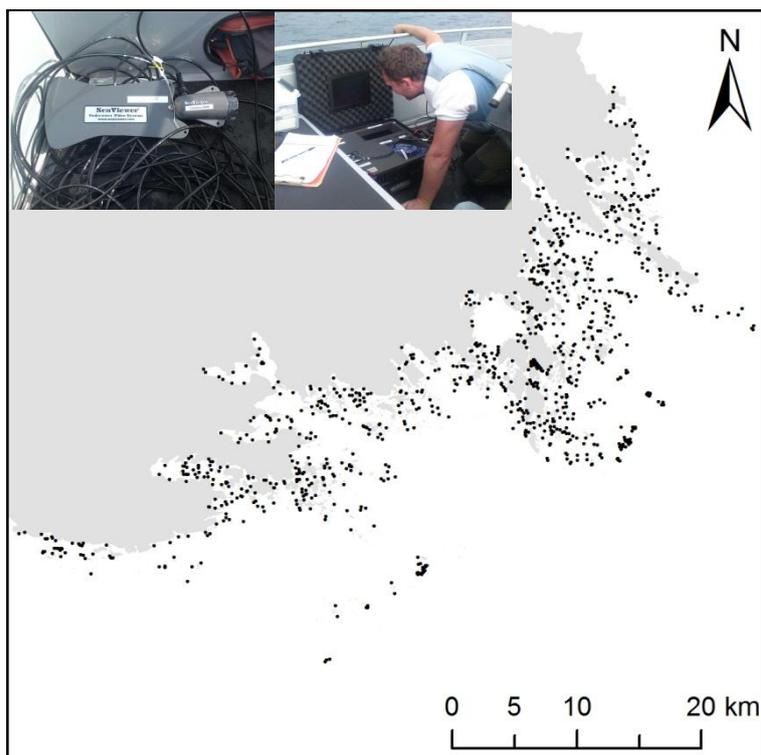


Figure 20. Positions for the 1006 drop-video stations used in the project. Observations are made from a screen in the boat. Camera is equipped with a fin for improved stability in the water. Photographer: Julia Carlström, AquaBiota Water Research.

8.1.2.1 Substrate

Substrate was divided into 9 different grain size classes (Table 6). The coverage (%) of the different classes was estimated for the surveyed area. The total coverage of substrate cannot exceed 100%, therefore the classes with highest coverage were sometimes adjusted downwards correspondingly.

Table 2. Substrate classification in underwater drop-video assessment.

Class	Grain size (mm)
Bedrock	Indefinable boulder
Large boulder	> 600
Boulder	200 - 600
Large stone	60 - 200
Stone	20 - 60
Gravel	2 - 20
Sand	0.06 - 2
Soft (silt & clay)	< 0.06
Clay (firm)	< 0.06

8.1.2.2 Benthic organisms

Estimates on abundance of taxonomic groups of benthic organisms were made in accordance to the standard method for monitoring of the phytobenthic of the Swedish Baltic coast (Kautsky 1992; Blomqvist 2007). The coverage by sessile species of algae, blue mussels and barnacles were estimated during observations in the field using 7 different coverage classes: 1, 5, 10, 25, 50, 75 and 100%. The ambition was always to identify down to species level if possible. If not, the organism was identified as accurately as possible, i.e. to genus, functional group etc. Abundance of mobile organisms (e.g. *Sadura entomon*, Mysidae) was estimated on a three-grade relative scale: 1) occasional, 2) common, 3) many.

The duration of the recorded video clips from different stations varied to due substrate and species composition. For example, a station with 100% sand and no species of algae is swiftly investigated, whereas a station with boulders and rock with high algae cover and diversity needs more time for identification.

8.1.3 Species modelled

The abundance of 9 species/groups of species was analysed. Four species of algae, one group of algae, three species of phanerogams and one animal species (Table 3).

Table 3. Species modelled.

Response	Taxonomic unit	Description
<i>Chorda filum</i>	Macroalgae	Brownalgae growing in shallow areas (1-7 m in study area)
<i>Coccotylus/Phyllophora</i>	Macroalgae	Redalage growing from 2 to 25 m in the study area
<i>Fucus vesiculosus</i>	Macroalgae	Brownalgae growing in shallow areas (0-10 m in study area)
Filamentous red algae	Macroalgae	Includes Polysiphonia sp., Rhodomela confervoides and Ceramium sp.)
<i>Furcellaria lumbricalis</i>	Macroalgae	Redalgae growing from 1 to 25 m in the study area
<i>Myriophyllum spicatum</i>	Phanerogam	Reaching over 1 m in height
<i>Potamogeton perfoliatus</i>	Phanerogam	Reaching up to 3 m in height
<i>Stuckenia pectinata</i>	Phanerogam	Reaching over 1.2 m in height
<i>Mytilus edulis</i>	Mollusc	Bivalve found at all depths (0-30 m) in the study area

Sea lace (*Chorda filum*) is a brownalgae with long cord-like fronds growing in the sublittoral zone down to 7 m (Tolstoy & Österlund, 2003). It is mostly found in sheltered areas attached to stones or shells. *Coccotylus truncatus* is a redalgae that can be found on depths between 2 and 25 m in the Baltic Sea (Tolstoy & Österlund, 2003). It grows up to 10 cm in height in the study area. The species cannot be separated from *Phyllophora pseudoceranoides* in a drop-video survey, which is the reason they are modeled as one. With its three dimensional structure bladderwrack (*Fucus vesiculosus*) is the main structure forming algae in shallow areas in the Swedish part of the Baltic Sea. It grows on hard substrates like bedrock, rock and stone. The distribution of Bladderwrack is limited by substrate and the influx of light. It can be found down to 12 m depth where conditions are good. Along with baltic agar (*Furcellaria lumbricalis*) the filamentous red algae dominate the biomass on hard substrates between 5 and 25 m depth in the southern part of the Baltic Sea (AquaBiota surveys in Östergötland, Södermanland, Stockholm and Blekinge). This group of algae consists mainly of *Polysiphonia fucoides*.

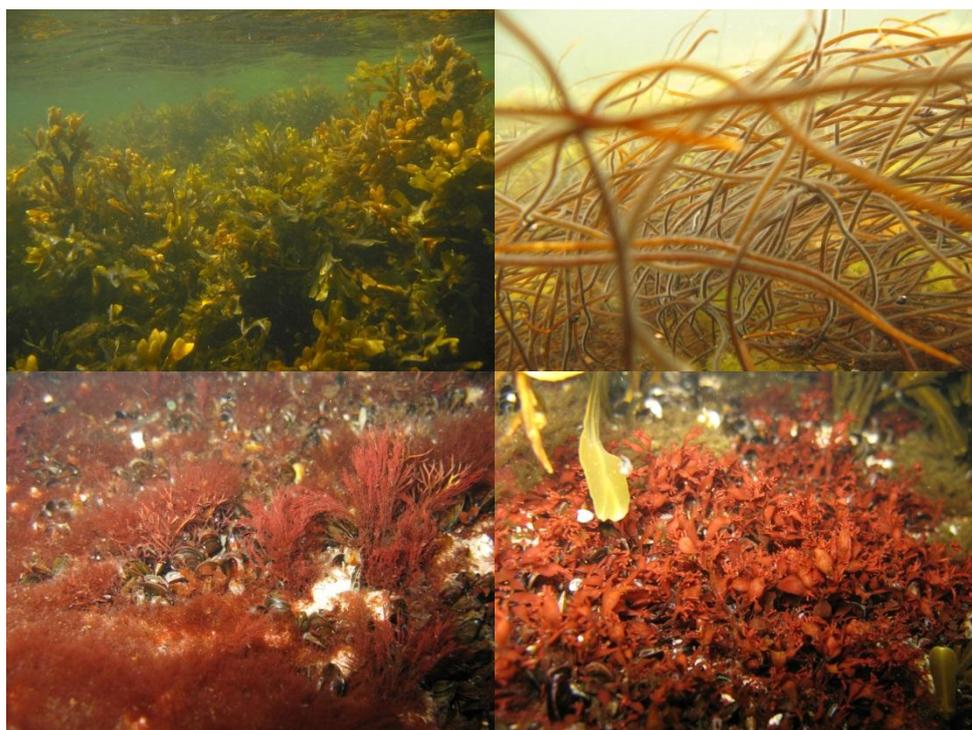


Figure 21. Top left: bladderwrack (Photographer: Martin Iseaus, AquaBiota water research). Top right: sea lace (Photographer: Karl Florén, AquaBiota water research). Bottom left: baltic agar and the filamentous red algae *Polysiphonia fucoides* (Photographer: Karl Florén, AquaBiota water research). Bottom right: The red algae *Coccotylus truncatus* (Photographer: Martin Isaeus, AquaBiota water research).

Spiked water milfoil (*Myriophyllum spicatum*) grows on soft substrates like all the other phanerogams. The species is often found in sheltered bays between 0.2 and 6 m depth (Wallentinus, 1979). Claspingleaf (*Potamogeton perfoliatus*) can reach several meters in height and is found in sheltered environments between 0.5 and 6.5 m depth (Wallentinus, 1979). The species is generally prevalent in both fresh and brackish water. Fennel pondweed (*Stuckenia pectinata*) was the dominant phanerogam in the study area. The species can be found from 0.3 to 10 m depth (Wallentinus, 1979). The high morphological complexity of both spiked water milfoil and fennel pondweed benefits macroinvertebrate abundance (Hansen et al, 2010).

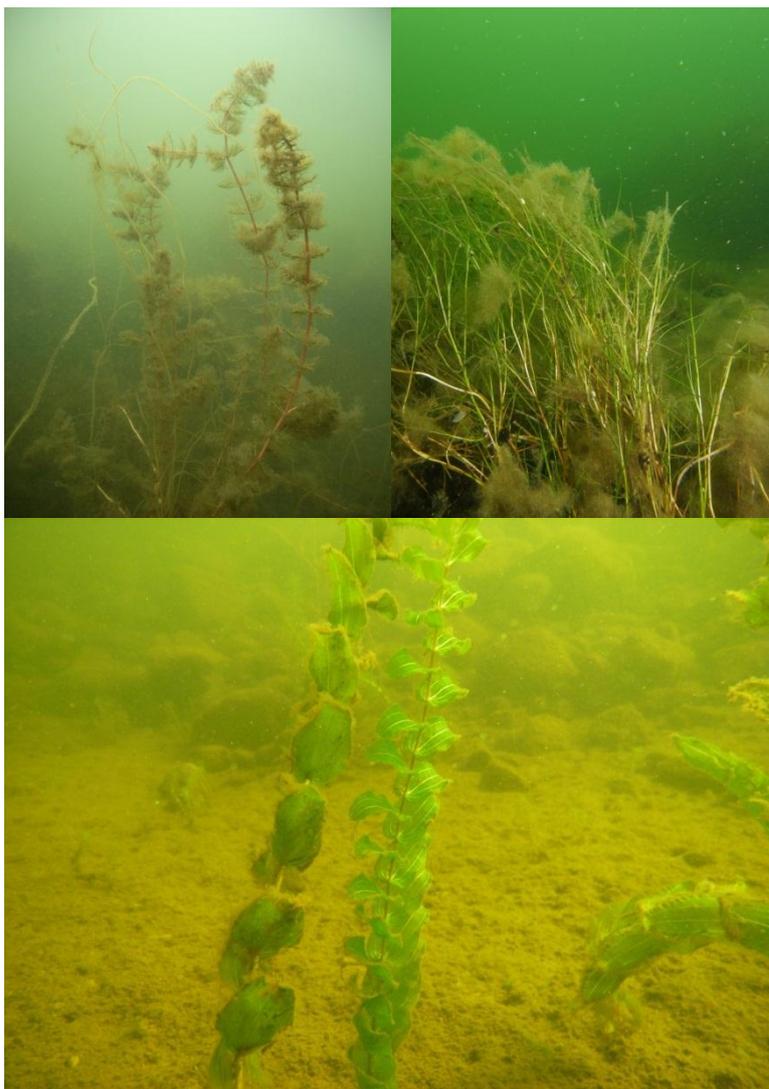


Figure 22. Top left: spiked water milfoil (Photographer: Jonas Edlund). Top right: fennel pondweed (Photographer Mathias H. Andersson). Bottom: clasp leaf pondweed (Photographer: Karl Florén AquaBiota water research).

In the Baltic Sea the blue mussel (*Mytilus edulis*) reaches up to 3 cm in length. The species attaches to hard substrates but can also be found on soft/sandy bottoms attached to shells. The species can cover large areas on hard substrates in exposed environments and can constitute an important food source for benthic feeding organisms (Lappalainen et al, 2005).



Figure 23. Blue mussels (Photographer Mathias H. Andersson).

8.1.4 Manipulation of datasets for the modelling process

The datasets used to build models and validate predictions were different depending on the range of the response variables presences along two important environmental gradients. With the 1008 drop-video stations as base each dataset was extracted according to the species distribution range along the depth and wave exposure gradients (illustrated in Figure 24). Reducing sample size is generally considered a bad idea in statistics. However, building a model where a very large proportion of the gradient are sure absences may give (falsely) inflated performance values since the model will then be very accurate simply because of the long redundant gradient covered. Hence, by excluding stations far outside a species observed presence range, although we may end up with lower model performance, the prediction may still improve as the resolution in the response within the relevant presence range will be higher and likely more reliable. That is, people do not need a model saying there are no bladderwrack at 100 m depth since it is quite obvious. For some species the full dataset (1006 stations) was of course needed. 20 % of the stations from the extracted dataset were used for validation of the predictions.

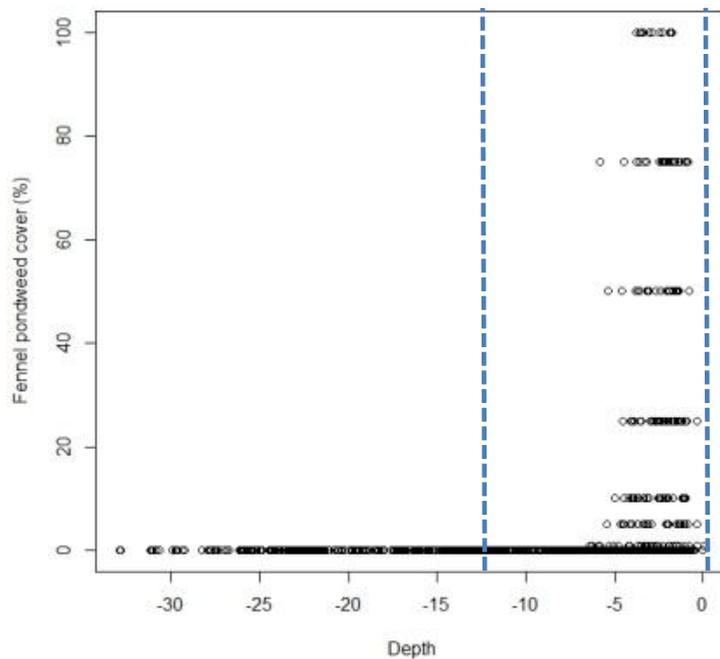


Figure 24. The modeling dataset for fennel pondweed (between the dotted lines). Stations deeper than 12.5 meters were excluded from this dataset. A model was also built from the whole dataset but with a less satisfactory result both in model performance and prediction accuracy.

8.1.5 Environmental variables

The environmental variables used in the statistical modelling and spatial predictions were developed as raster layers (10 m spatial resolution) in ArcGis 10. Maps of all the environmental variables are shown in Figure 25 – 32.

8.1.5.1 Secchi depth

The Secchi depth map developed for this project resampled to 10 m resolution

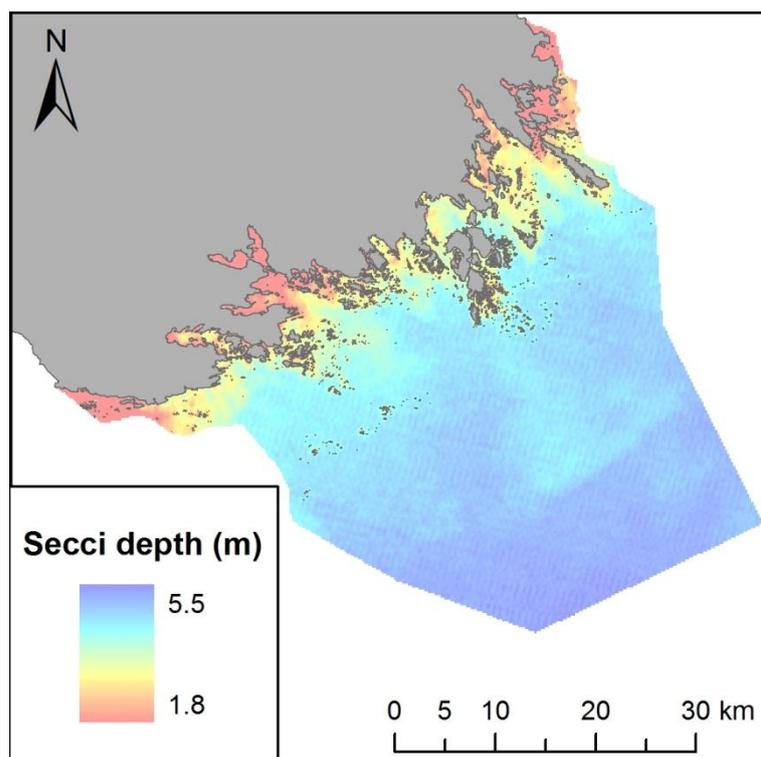


Figure 25. Secchi depth in Södermanlands marine area, spatial resolution 10 m.

8.1.5.2 Depth

The Swedish Maritime Administration supplied digitalised depth data in GIS point shape files for the whole study area. Depth information was usually dense in waterways and other frequently travelled waters but scarce in other parts of the study area. By interpolation gaps between known depths were assigned values resulting in a raster layer. The interpolation method used was ordinary kriging which is a geostatistical interpolation technique. Unlike many other interpolation techniques kriging uses, not only nearby neighbours for calculations, but also the spatial relationship of the whole dataset.

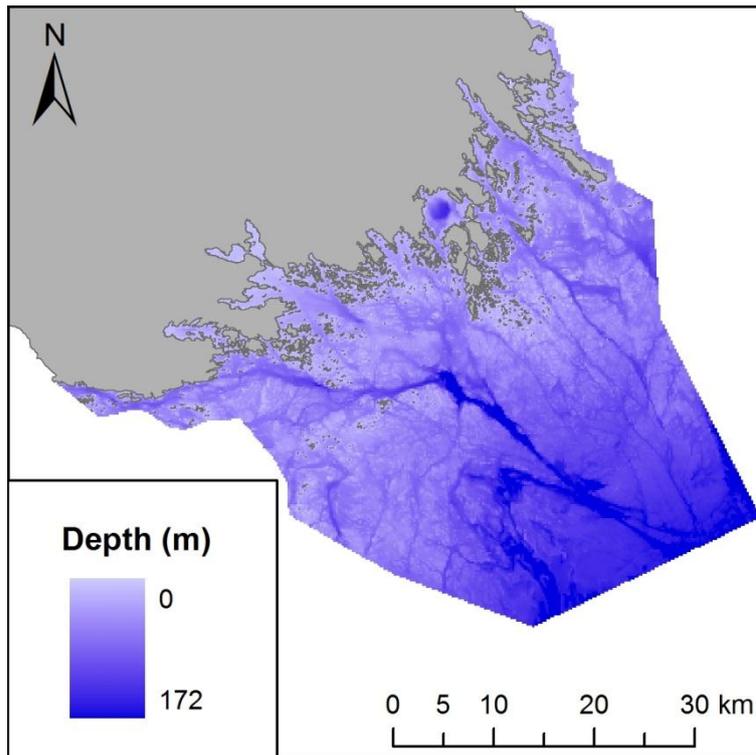


Figure 26. Depth in Södermanlands marine area, spatial resolution 200 m. The map is available in 10 m resolution in a classified attachment.

8.1.5.3 Slope

Slope was estimated by calculating the maximum rate of change of depth from one cell to its eight neighbours (in degrees) using the ArcGis 9.3.1 "Slope" function. The resulting slope raster layer ranged from 0 - 54° in the study area.

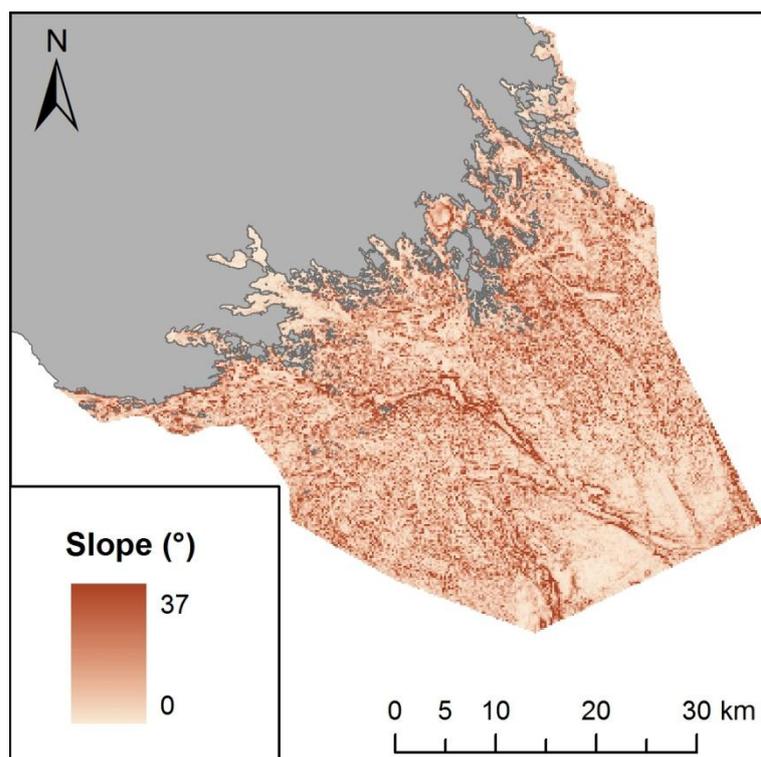


Figure 27. Slope in Södermanlands marine area, spatial resolution 200 m. By reducing resolution maximum slope changed from 54° to 37°. The map is available in 10 m resolution in a classified attachment.

8.1.5.4 Curvature

Curvature was calculated in two steps. In the first step the mean depth in a moving neighbourhood, defined by circle with a 300 m radius was calculated. In the second step this value was subtracted from the depth value in each raster cell. A positive raster value indicates a shoal and a negative value indicates that the raster cell is within a basin. Curvature within the study area ranged from -44 – 41 m.

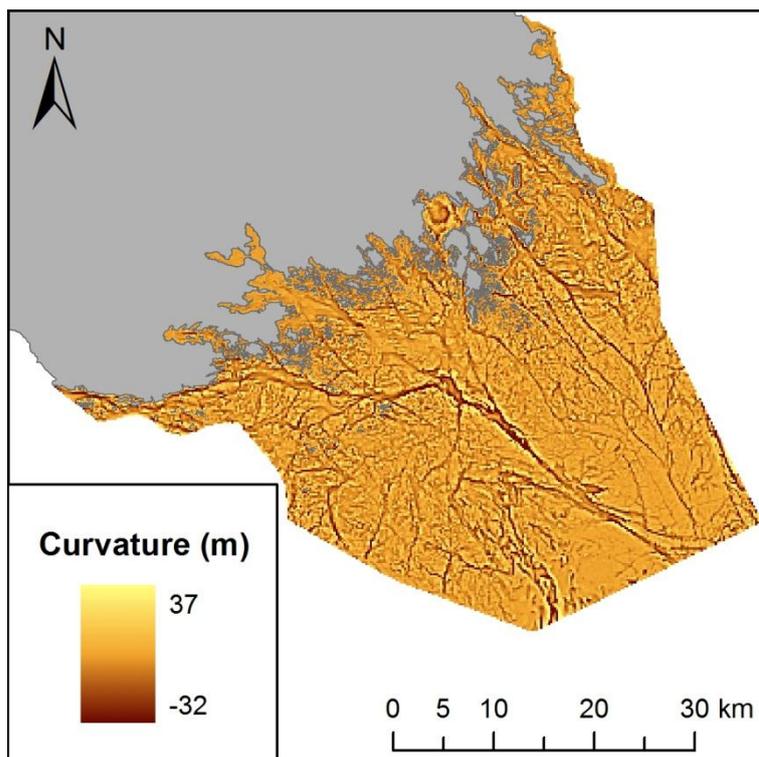


Figure 28. Curvature in Södermanlands marine area, spatial resolution 200 m. By reducing resolution curvature range changed from -44 – 41 m to -32 – 37 m. The map is available in 10 m resolution in a classified attachment.

8.1.5.5 Wave exposure

The wave exposure was calculated using (SWM) Simplified Wave Model (Isæus 2004). The model uses data on fetch (distance to nearest shore), wind strength and wind direction. The structure of the Södermanland archipelago provides a gradient of exposure level ranging from ultra-sheltered to exposed.

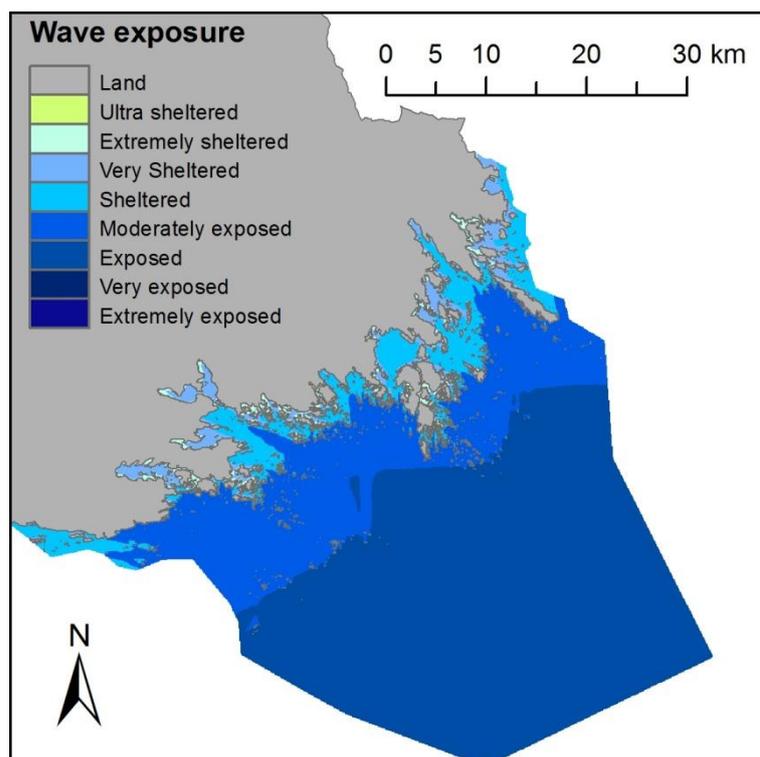


Figure 29. Wave exposure calculated with SWM (Simplified Wave Model) in Södermanland marine area, spatial resolution 10 m.

8.1.5.6 Salinity

The salinity used in the secchi-modeling describes the minimum salinity at the surface. It is a composite of values from two different types of coastal hydrodynamic models; the three-dimensional, large-scale Arc3D model with 0.5' x 0.5' grid resolution (Engqvist and Andrejev, 2003) and the coastal one-dimensional coupled basin model CouBa (Engqvist, 2008). To the largest possible degree data were used from the Arc3D-model, but in the near-coastal regions not covered by Arc3D, data from CouBa were used. The minimum salinity was calculated as the 10-percentile of all values monthly values over a period of twelve years; from January 1993 to December 2004 (no newer data were available from these two models). Minimum Salinity at surface within the study area ranged from 0 to 6.6 psu.

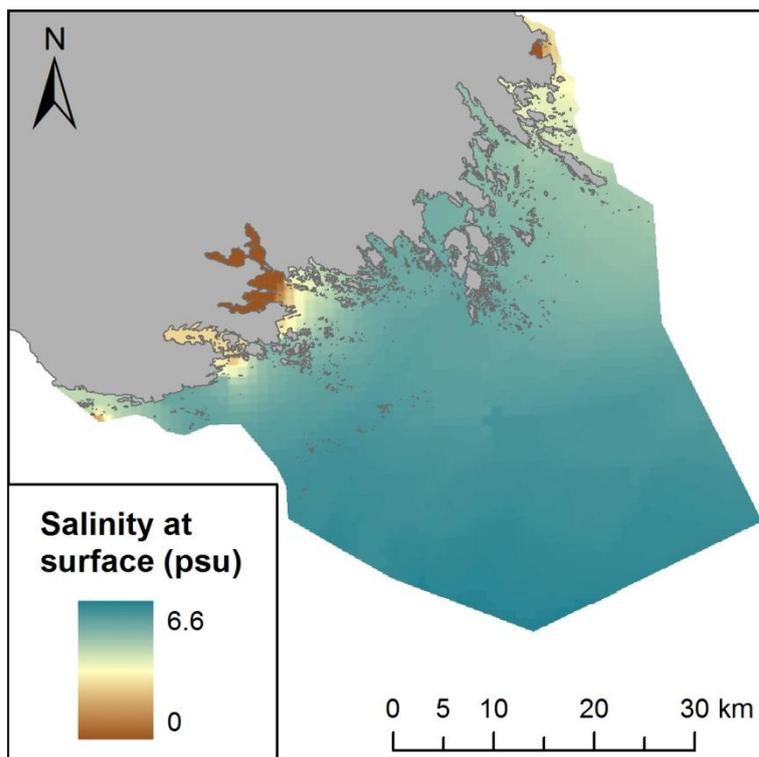


Figure 30. Minimum salinity at surface in Södermanland marine area, spatial resolution 10 m.

8.1.5.7 Temperature

The used temperature layer describes the mean temperature at bottom level. The raster was compiled from model data from the same sources as salinity (see the above section), with input of bottom depth from the raster layer produced in this project. The mean temperature was calculated based on monthly values over the twelve-year period January 1993 to December 2004. Mean Temperature at bottom level within the study area ranged from 2.1 to 8.7 °C .

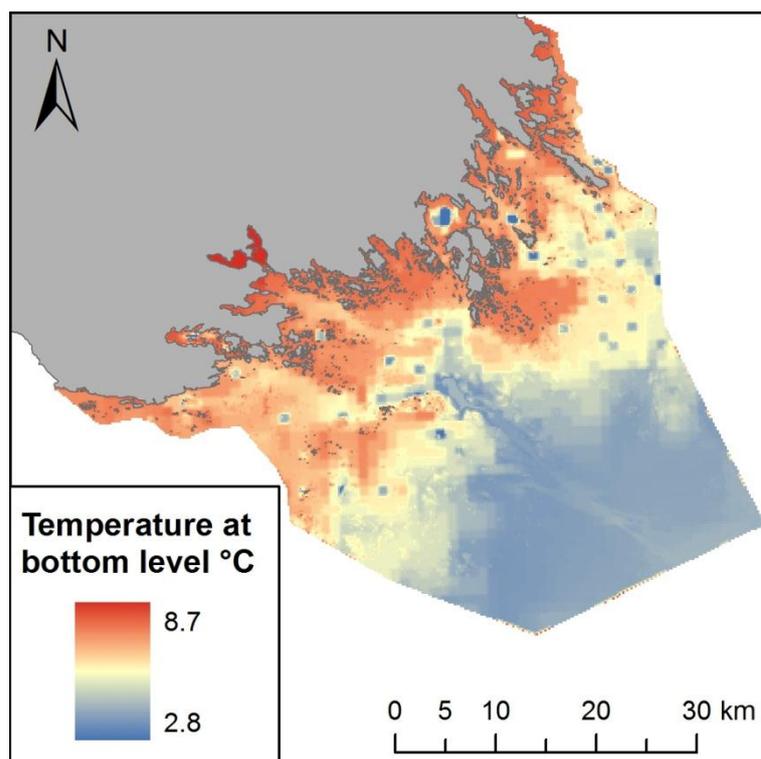


Figure 31. Mean temperature at bottom level in Södermanland marine area, spatial resolution 200 m.

8.1.5.8 Density of potentially polluted areas (DPPA)

Between 2006 and 2009 a survey was conducted with the aim to identify polluted areas in the county of Södermanland (Länsstyrelsen i Södermanland, 2009). Based on the results from that project a second survey was conducted between 2009 and 2010 (Länsstyrelsen i Södermanland, 2010). In these two surveys over 2000 objects were identified as potentially polluted. From this data a raster layer was developed in GIS by a simple point density analysis. Each raster cell was assigned a value calculated from the number of objects within a radius of 10 km from the cell.

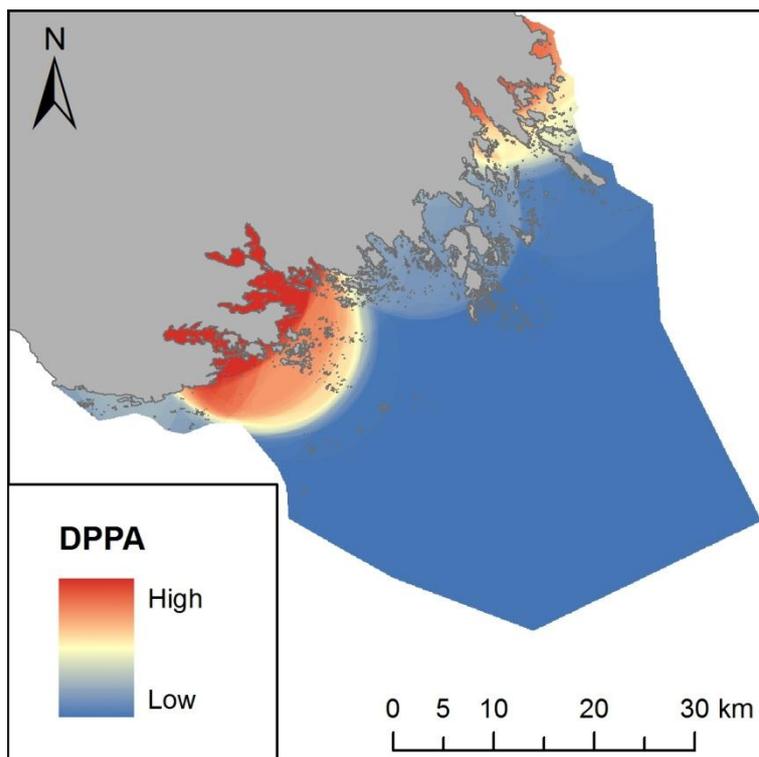


Figure 32. Density of potentially polluted areas (DPPA) in 10 m resolution.

9 RESULTS

The correlation between predictors are shown in Figure 33. Correlation coefficients over 0.7 could distort model estimation. The highest correlation occurred between depth and temperature at sea floor (0.55) and between salinity and DPPA (0.52). Some correlation also occurred between Secchi depth and wave exposure (0.42). After studying results from the correlation analysis along with results from the VIF analysis (Table 4) it was decided that none of the predictors had to be omitted from the analysis.

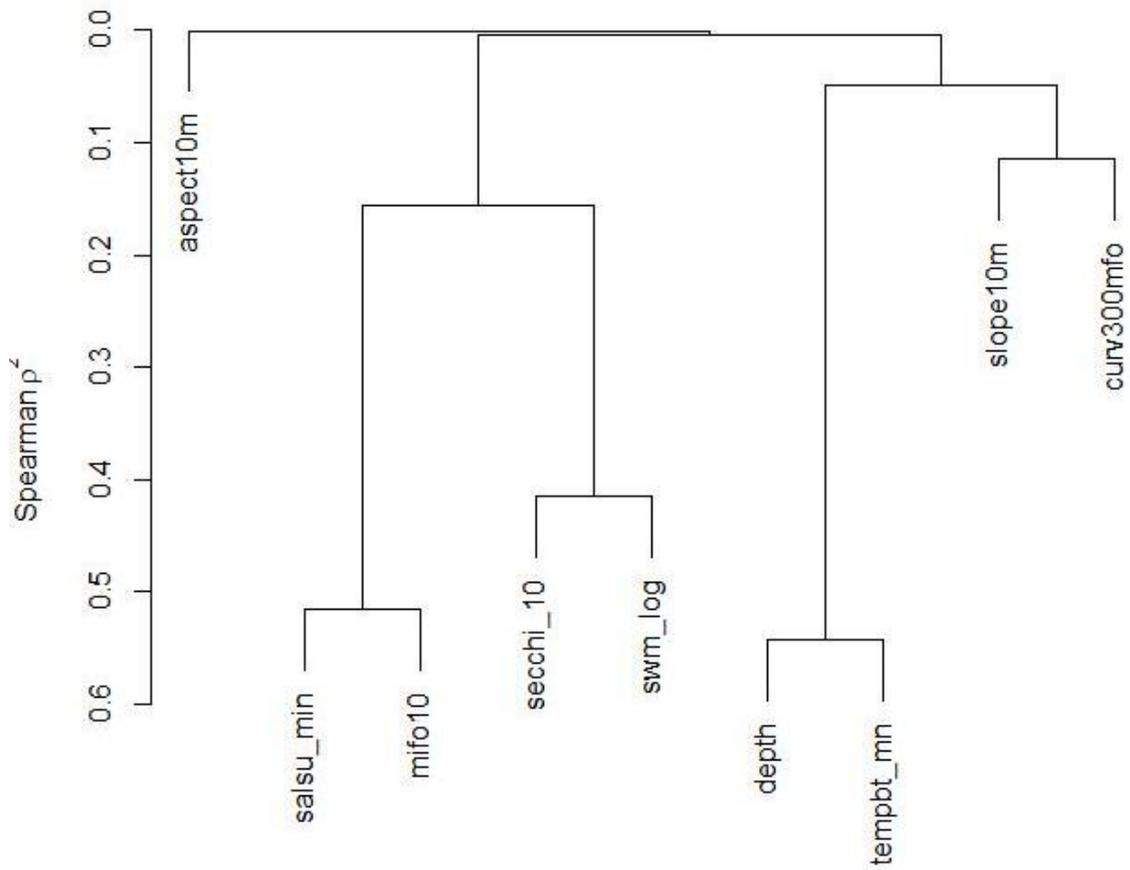


Figure 33. Cluster diagram showing the Spearman correlation between the predictors used in this project. The strongest correlation (coefficient 0.55) was between depth and temperature at seafloor level.

Table 4. Results from the Variance Inflation Factor (VIF) analysis. According to Zuur et al (2009) predictors with values above 3 should be removed.

Variable	VIF
Depth	2.55
Secchi	2.15
Salinity	1.91
Temperature	1.87
Slope	1.52
Curvature	1.74
Aspect	1.01
DPPA	1.92
Exposure	1.71

For 6 out of 9 species modeled Secchi depth had little or no importance as predictor. However, for bladderwrack, fennel pondweed and sea lace Secchi depth turned out to be an important predictor. For these species models were also built without Secchi depth to

explore the impact of the variable on model performance and prediction accuracy. Prediction maps were visually compared to identify local differences.

The numerical results are shown in table 5. For bladderwrack and fennel pondweed depth was the most important predictor. For sea lace exposure was the most important predictor. With Secchi depth included the predictor explained around 20 % of the species distribution. Without Secchi depth the importance of salinity and exposure increased indicating some correlation between these three predictors.

Table 5. Predictors importance to model with and without Secchi depth for three different species.

Secchi	Species	n	Relative influence of predictors to model							
			Depth	Secchi	Salinity	Temperature	Slope	Curvature	DPPA	Exposure
No	Bladderwrack	798	0.64	-	0.20	0.00	0.00	0.15	0.00	0.00
Yes	Bladderwrack	798	0.82	0.17	0.00	0.00	0.00	0.00	0.00	0.00
No	Fennel pondweed	556	0.50	-	0.00	0.07	0.00	0.00	0.04	0.38
Yes	Fennel pondweed	556	0.51	0.20	0.00	0.02	0.00	0.00	0.07	0.20
No	Sea lace	735	0.27	-	0.23	0.00	0.00	0.00	-	0.50
Yes	Sea lace	735	0.25	0.18	0.14	0.00	0.00	0.00	-	0.43

The deviance explained by the models was between 0.42 and 0.5. This means that the models explain between 42 and 50 % of the variation in the datasets (Table 6). By including Secchi depth these results improved for all three species. The predictive accuracy (AUC) was excellent or good. By including Secchi depth AUC improved for all species. Sensitivity improved for bladderwrack and fennel pondweed but decreased for sea lace. Specificity improved for all three species.

Table 6. Model results with and without Secchi depth. The number of stations to build models and validate predictions varied between species. Deviance explained by model (*devexpl*) is a measure of model performance. AUC is a measure of prediction accuracy (quality of map). Sensitivity is the quality of the map in terms of telling where the species is present whereas specificity is the quality of the map in terms of telling where the species is absent.

Secchi	Species	n train	n test	devexpl	AUC	Sensitivity	Specificity
No	Bladderwrack	798	200	0.42	0.91	0.84	0.83
Yes	Bladderwrack	798	200	0.44	0.92	0.87	0.84
No	Fennel pondweed	556	139	0.46	0.90	0.76	0.82
Yes	Fennel pondweed	556	139	0.50	0.92	0.82	0.85
No	Sea lace	735	184	0.45	0.88	0.85	0.83
Yes	Sea lace	735	184	0.49	0.90	0.77	0.85

When visually comparing predictions from models with and without Secchi depth some distinctive local differences were discovered. The bladderwrack model without Secchi predicted high probabilities of presence outside Trosa, north of Öbolandet in the northern part of Södermanland. This area is highly eutrophied due to nutrient rich water from the Trosa river. By adding Secchi depth to this model the probabilities of finding bladderwrack in that area were almost zero (figure 34). None of the six dropvideo stations available from this area had observations of bladderwrack even though depth conditions were right. The waters outside Nyköping are also shallow and have poor visibility but there was little difference between the two predictions (Figure 35). In contrast to the area outside Trosa, this area had almost fresh water conditions which were picked up by the model from the salinity layer.

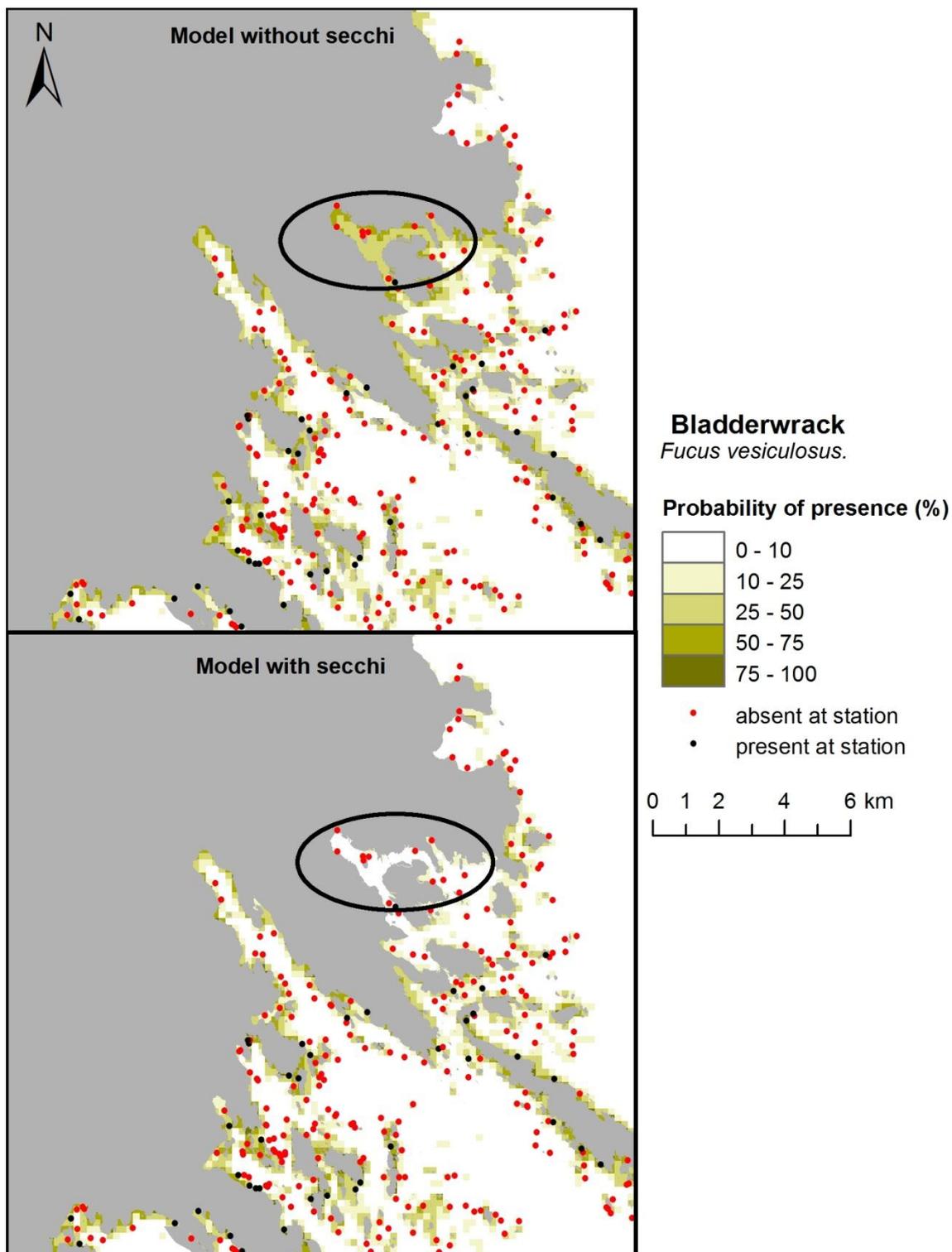


Figure 34. Prediction maps of bladderwrack from two different models, one with and one without secchi depth in 200 m resolution. The ellipse highlights an area where the two predictions differed the most. After approval from the Swedish Maritime Administration the prediction map will be available in higher resolution.

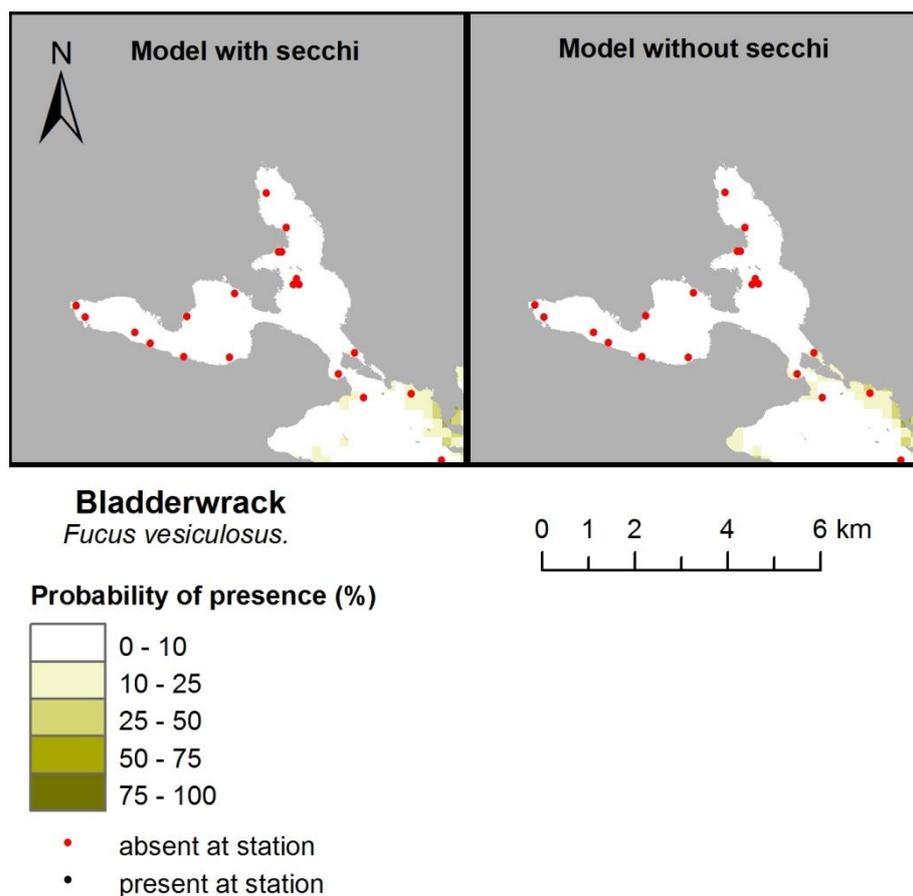


Figure 35. Prediction maps of bladderwrack from two different models, one with and one without secchi depth in 200 m resolution outside Nyköping. The two predictions show similar patterns in this area. The Drop-video stations indicate if bladderwrack was present or not in the survey 2010. After approval from the Swedish Maritime Administration the prediction map will be available in higher resolution.

A comparison of predictions of fennel pondweed showed similar differences as for bladderwrack (Figure 36). The difference was most obvious in the same eutrophied area outside Trosa. The species was not found in this area making the model including secchi depth the most reliable.

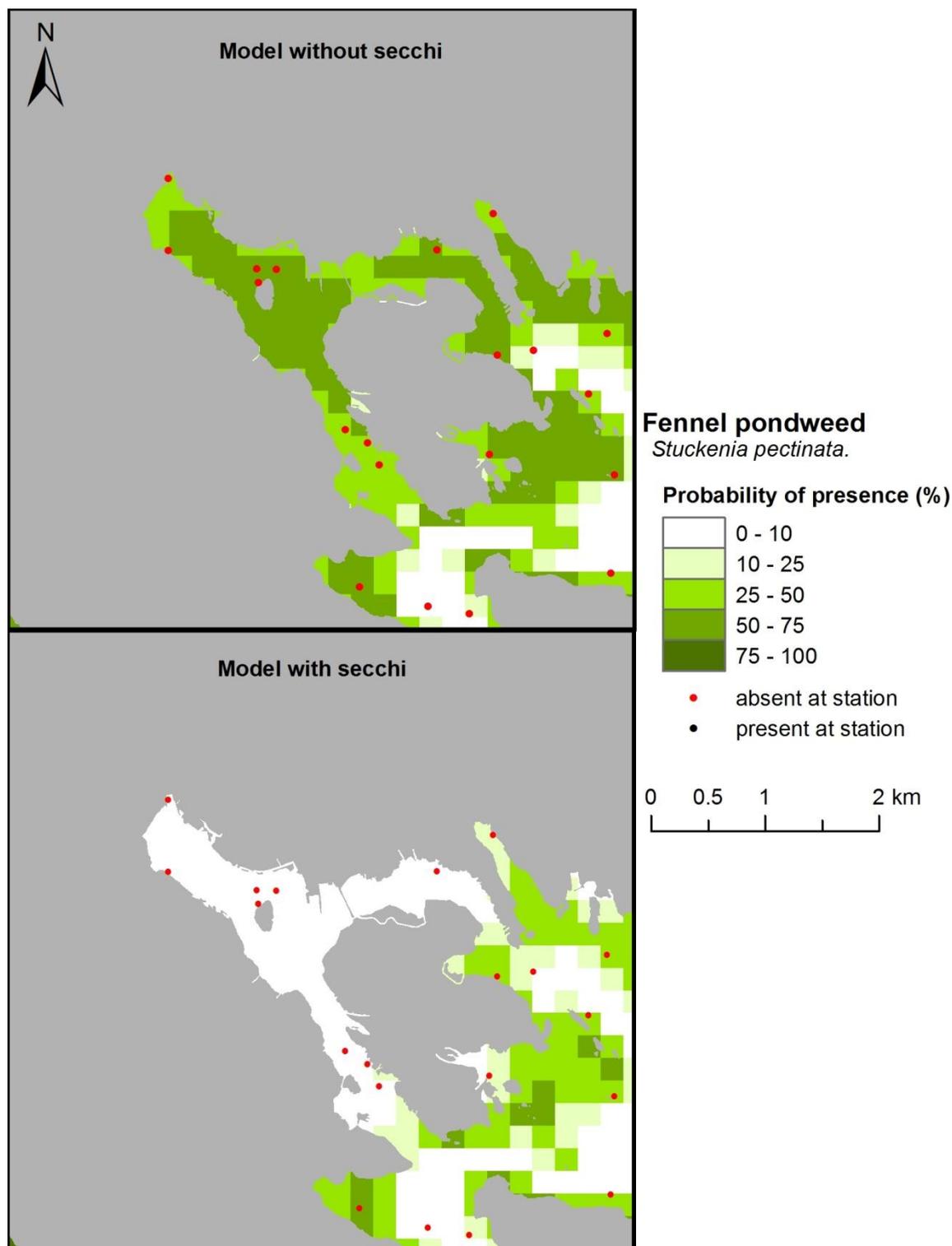


Figure 36. Prediction maps of fennel pondweed from two different models, one with and one without Secchi depth in 200 m resolution. The two predictions show totally different patterns in the eutrophied area outside Trosa. The Drop-video stations indicate if fennel pondweed was present or not in the survey 2010. After approval from the Swedish Maritime Administration the prediction map will be available in higher resolution.

Prediction maps for bladderwrack, fennel pondweed and sea lace, with and without Secchi depth covering the whole study area, are shown in Appendix 1.

10 CONCLUSIONS HABITAT MODELLING

Secchi depth improved model performance and prediction accuracy for three out of nine species. Two of these species (bladderwrack and fennel pondweed) are important fish recruitment habitats in shallow waters (Sandström et al. 2005). The improvements were rather small looking at the evaluation figures of model performance and large scale prediction maps. However, when studying the prediction maps in more detail large improvements were discovered on a more local scale. Secchi depth is a strong indicator on eutrophication, even though also other factors affect the water transparency. Areas affected by heavy nutrient loads from point sources is expected to show most significant effects on local water transparency and therefore also on the distribution of species sensitive to eutrophication. This explains the strong local effect on bladderwrack in the Trosa area, which was nicely predicted when using the Secchi depth map as predictor. Depth and wave exposure can usually explain the large scale patterns of these species distributions with satisfying accuracy, and a small area like the one outside Trosa illustrated in Figure 36 may only have little influence to the overall modeling result. Maps with obvious flaws will be rejected by managers even if the overall quality is excellent. However in areas with nutrient point sources the resulting maps will be incorrect. This is an important reason for having experts to judge the final maps before they come into use for management.

The contribution of a Secchi depth map will depend on the species or habitats modeled, study area and the scale of the maps that are produced. In this project the evaluation of the predictor has been limited to 9 macrophyte species which are detectable with a drop-video camera. The maps produced have covered the whole Södermanland County which includes many types of environments ranging from exposed islets to sheltered bays. Judging from the results the Secchi map has contributed most in the more sheltered areas in the County where depth and exposure conditions usually are less variable. A future study limited to these types of environments and the species found there would probably give similar but more pronounced results. The Secchi depth map will come to further use in Södermanland and Stockholm in a project called MMSS (Marine Modeling in Stockholm and Södermanland) where fish, benthic infauna and additional species of macroalgae will be modeled. The same methodology as in this project will be used to produce Secchi depth maps for modeling of species distributions in the counties of Skåne and Blekinge and in the Finnish archipelagic sea.

The true Secchi depth at a certain location is varying over time and will depend on a number of factors such as wind direction and strength, rainfall, currents, algal blooms etc. Whether a long term average or a snapshot of this variable best explains the distribution of different species is yet to be investigated. Further studies on the temporal scale of Secchi depth is definitely needed to fully understand its impact on marine life.

11 REFERENCES

- Lappalainen, A., Westerbom, M. & Heikinheimo, O. (2005) Roach (*Rutilus rutilus*) as an important predator on blue mussel (*Mytilus edulis*) populations in a brackish water environment, the northern Baltic Sea. *Marine Biology*, **147**, 323-330.
- Bekkby, T. et al., 2009. Spatial predictive distribution modelling of the kelp species *Laminaria hyperborea*. *ICES Journal of Marine Science: Journal du Conseil*, 66(10), pp.2106 -2115.
- Blomqvist, M. 2007. Transektinventering av marina bottenar. Manual för inmatningsapplikationen, Hafok AB, på uppdrag av Naturvårdsverket, 2007-05-29.
- Burnham, K.P. & Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* 2nd ed., New York: Springer.
- Eriksson, B.K. & Bergström, L., 2005. Local distribution patterns of macroalgae in relation to environmental variables in the northern Baltic Proper. *Estuarine, Coastal and Shelf Science*, 62(1-2), pp.109-117.
- Eriksson, B.K. & Johansson, G., 2005. Effects of sedimentation on macroalgae: species-specific responses are related to reproductive traits. *Oecologia*, 143(3), pp.438-448.
- Eriksson, B.K. & Johansson, G., 2003. Sedimentation reduces recruitment success of *Fucus vesiculosus* (Phaeophyceae) in the Baltic Sea. *European Journal of Phycology*, 38(3), pp.217-222.
- Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.
- Florin, A.-B., Sundblad, G. & Bergström, U. (2009) Characterisation of juvenile flatfish habitats in the Baltic Sea. *Estuarine, Coastal and Shelf Science*, **82**, 294-300.
- Hansen, J.P., Sagerman, J & Wikström, S.A. (2010) Effects of plant morphology on small-scale distribution of invertebrates. *Marine Biology*, **157**, 2143-2155.
- Isæus, M., 2004. Factors structuring *Fucus* communities at open and complex coastlines in the Baltic Sea. Dissertation. Stockholm: Stockholm University.

Kautsky, H., 1992. Methods for monitoring of phytobenthic plant and animal communities in the Baltic Sea. In M. Plinski, ed. The ecology of Baltic terrestrial, coastal and offshore areas - protection and management. *Gdansk: Sopot*, pp. 21-59.

Kautsky, H., 1995. Quantitative distribution of sublittoral plant and animal communities along the Baltic Sea gradient. In A. Eleftheriou, A. D. Ansell, & C. J. Smith, eds. *Biology and Ecology of Shallow Coastal Water, Proceedings of the 28th European Marine Biology Symposium, Crete. Fredensborg, Denmark: Olsen & Olsen*, pp. 23-30.

Kautsky, H. & van der Maarel, E., 1990. Multivariate approaches to the variation in phytobenthic communities and environmental vectors in the Baltic Sea. *Marine Ecology Progress Series*, 60, pp.169-184.

Kautsky, N. et al., 1986. Decreased depth penetration of *Fucus vesiculosus* (L.) since the 1940's indicates eutrophication of the Baltic Sea. *Marine Ecology Progress Series*, 28(1-2), pp.1-8.

Kiirikki, M., 1996. Mechanisms affecting macroalgal zonation in the northern Baltic Sea. *European Journal of Phycology*, 31(3), pp.225-232.

Kirk, J. T. O. 1994. Light and photosynthesis in aquatic ecosystems. 2nd edition. Cambridge University Press, Melbourne.

Krause-Jensen, D. et al., 2007. Spatial patterns of macroalgal abundance in relation to eutrophication. *Marine Biology*, 152(1), pp.25-36.

Lee, Z., K. L. Carder, C. D. Mobley, R. G. Steward, and J. S. Patch. 1998. Hyperspectral remote sensing for shallow waters: 1. A semianalytical model. *Applied Optics* 37:6329-6338.

Lindfors, A., K. Rasmus, and N. Strömbeck. 2004. Point or pointless - quality of ground data. *International Journal of Remote Sensing* 26:415-423.

Länsstyrelsen i Södermanland (2009)
<http://www.lansstyrelsen.se/sodermanland/SiteCollectionDocuments/sv/miljo-och-klimat/verksamheter-med-miljopaverkan/forenade-omraden/tillsyn-och-tillsynsvagledning/Slutrapportinklabilagortillsynsprojekt20062009.pdf>.

Länsstyrelsen i Södermanland (2010)

http://www.lansstyrelsen.se/sodermanland/SiteCollectionDocuments/sv/miljo-och-klimat/verksamheter-med-miljopaverkan/forenade-omraden/tillsyn-och-tillsynsvagledning/Slutrapporttillsynsprojekt2009_2010.pdf.

Petzold, T. J. 1972. Volume scattering functions for selected ocean waters. SIO Ref. 72-78, Scripps Institute for Oceanography, La Jolla.

Remane, A. & Schlieper, C., 1971. *Biology of Brackish Water*, New York: John Wiley & Sons.

Sandström, A., B. K. Eriksson, P. Karås, M. Isæus and H. Schreiber (2005). "Boating activities influences the recruitment of near-shore fishes in a Baltic Sea archipelago area." *Ambio* 34(2): 125-130.

Snickars, M., Sundblad, G., Sandström, A., Ljunggren, L., Bergström, U., Johansson, G. & Mattila, J. (2010) Habitat selectivity of substrate-spawning fish: modelling requirements for the Eurasian perch *Perca fluviatilis*. *Marine Ecology Progress Series*, **398**, 235-243.

Sundblad, G., Bergström, U. & Sandström, A. (2011) Ecological coherence of marine protected area networks: a spatial assessment using species distribution models. *Journal of Applied Ecology*, **48**, 112-120.

Tedengren, M. & Kautsky, N., 1986. Comparative study of the physiology and its probable effect on size in blue mussels (*Mytilus edulis* L.) from the North Sea and the northern Baltic proper. *Ophelia*, 25(3), pp.147–155.

Wallentinus, I. (1979) Environmental influences on benthic macrovegetation in the Trosa-Askö area, northern Baltic proper. II. *The ecology of macroalgae and submersed phanerogams*. – Contrib. Askö Lab. Univ. Stockholm 25: 1-210

Wood, S.N. & Augustin, N.H. (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling*, **157**, 157-177.

12 APPENDIX 1

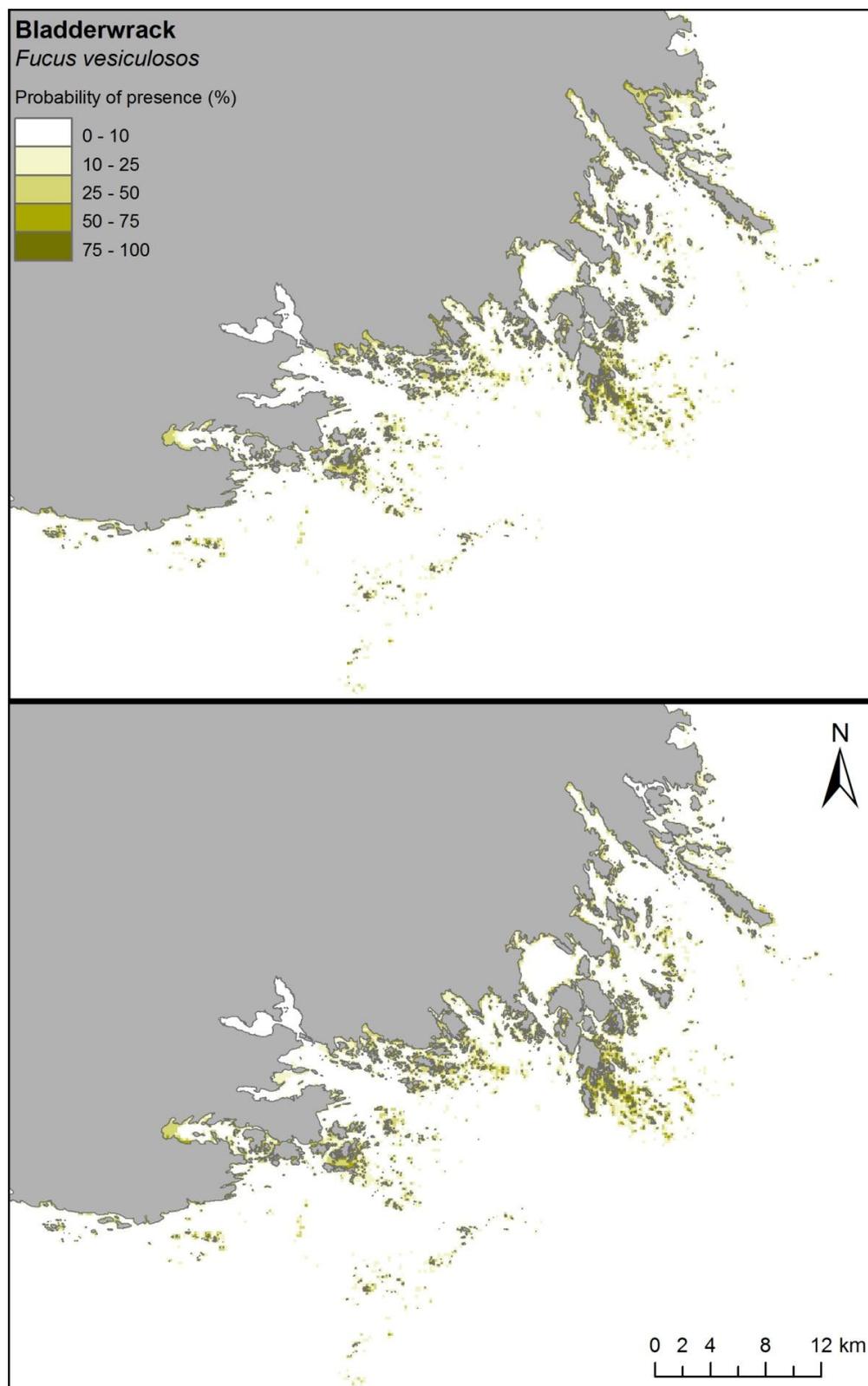


Figure 37. Modeled probability of presence for Bladderwrack in Södermanland marine area. Top: model without Secchi depth. Bottom: model with Secchi depth.



Figure 38. Modeled probability of presence for Fennel pondweed in Södermanland marine area. Top: model without Secchi depth. Bottom: model with Secchi depth.

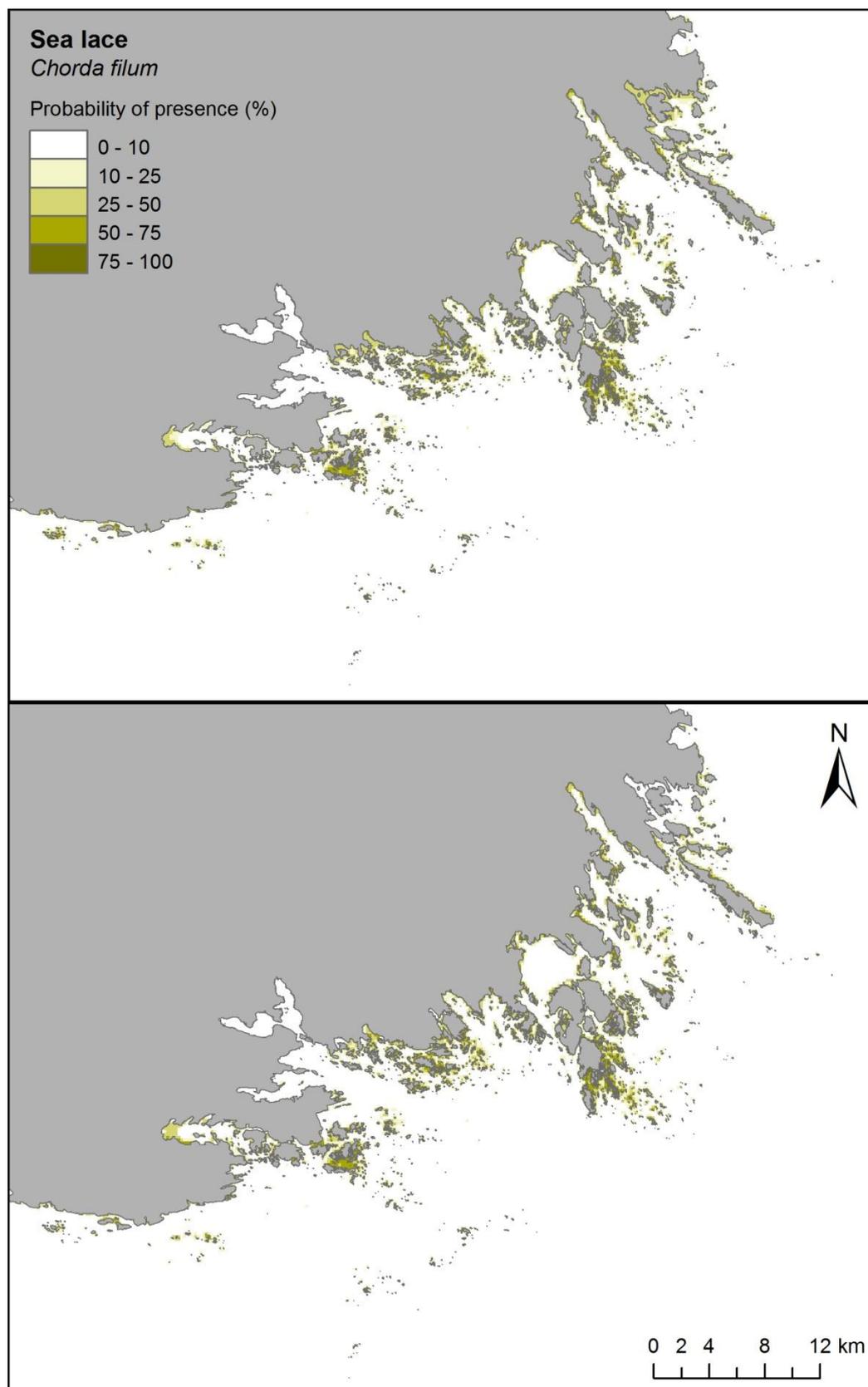


Figure 39. Modeled probability of presence for Sea lace in Södermanland marine area. Top: model without Secchi depth. Bottom: model with Secchi depth.

www.aquabiota.se