MODELLING CHLOROPHYLL-A CONCENTRATIONS IN DANISH COASTAL WATERS USING A BAYESIAN MODELLING APPROACH

Documentation report

Scientific Report from DCE – Danish Centre for Environment and Energy No. 469 2021

AARHUS
UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

DHI

MODELLING CHLOROPHYLL-A CONCENTRATIONS IN DANISH COASTAL WATERS USING A BAYESIAN MODELLING APPROACH

Documentation report

Scientific Report from DCE – Danish Centre for Environment and Energy No. 469 2021

Nisha Shetty[1]
Jesper Philip Aagaard Christensen[1]
Christian Damgaard[1]
Karen Timmermann[2]

[1]Aarhus University, Department of Ecoscience
[2]DTU Aqua

AARHUS
UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

Data sheet

Series title and no.:	Scientific Report from DCE – Danish Centre for Environment and Energy No. 469
Category:	Scientific advisory report
Title:	Modelling chlorophyll-a concentrations in Danish coastal waters using a Bayesian modelling approach
Subtitle:	Documentation report
Authors:	Nisha Shetty[1], Jesper Philip Aagaard Christensen[1], Christian Damgaard[1] & Karen Timmermann[2]
Institutions:	[1]Aarhus University, Department of Ecoscience & [2]DTU Aqua
Publisher:	Aarhus University, DCE – Danish Centre for Environment and Energy ©
URL:	http://dce.au.dk/en
Year of publication:	November 2021
Editing completed:	November 2021
Referee:	Jacob Carstensen
Quality assurance, DCE:	Anja Skjoldborg Hansen
Linguistic QA:	Anne van Acker
Financial support:	The Danish Environmental Protection Agency
Please cite as:	Shetty N, Christensen JPA, Damgaard C & Timmermann K. 2021. Modelling chlorophyll-a concentrations in Danish coastal waters using a Bayesian modelling approach. Documentation report. Aarhus University, DCE – Danish Centre for Environment and Energy, 62 pp. Scientific Report No. 469.

Abstract: In this study a Bayesian group model and 46 single station models were developed for estimation of chlorophyll a concentration, nutrient input relationship, in Danish waterbodies. The models relationships were derived to support the Danish EPA, setting input targets to fulfil the EU water framework directive. As expected nutrient input was the most common predictor of chlorophyll a concentration in single model stations (31 of 46) and temperature the second most abundant predictor. In the group model nutrient input, water temperature, salinity and water column stability, were the best predictors of chlorophyll a concentration in the 42 coastal monitoring stations, in the model.

Keywords:	Chlorophyll a, Danish water-bodies, Water Framework Directive, Bayesian modelling.
Layout:	Anne van Acker
Front page photo:	Peter Bondo Christensen
ISBN:	978-87-7156-639-0
ISSN (electronic):	2244-9981
Number of pages:	62
Internet version:	The report is available in electronic format (pdf) at http://dce2.au.dk/pub/SR469.pdf

Contents

Preface

This report was commissioned and funded by the Danish Environmental Protection Agency (EPA) as part of the project "Application of the Danish EPA's Marine Model Complex and Development of a Method Applicable for the River Basin Management Plans (RBMP) 2021-2027". The work reported was managed and performed by AU/DCE. DTU and DHI. During the project, a steering committee followed the development and was involved through dialogue and follow-up on progress, etc. The steering committee consisted of members from the Danish Ministry of Environment and Food (MFVM), the Danish EPA (MST), DHI and AU. In addition, a follow-up group consisting of members from The Danish Agriculture & Food Council. SEGES, National Association of Sustainable Agriculture, the Danish Society for Nature Conservation, the Danish Sports Fishing Association, Danish Fishermen PO (DFPO), the Danish Ports and KL/municipalities was affiliated with the project. The follow-up group has been continuously informed about the progress of the project at meetings convened by the MFVM. The group has had the chance to comment on the report. No changes has been made based upon comments from the group. Choice of methods, data processing, description and presentation of results have been solely AU/DCE's, DTU's and DHI's decision and responsibility.

Summary

Bayesian hierarchical models were developed to predict chlorophyll-a concentration in Danish water-bodies, as a function of land-based nitrogen and phosphorus loadings, physicochemical and climatic predictors. The objective of the model development was to support the Danish implementation of the Water Framework Directive (WFD) by providing tools applicable for estimating chlorophyll-a reference conditions and maximum allowable nutrient input (MAI) to the Danish coastal waters covered by the water framework directive.

We developed single station Bayesian models for 46 Danish water quality monitoring stations representing 43 water-bodies, as well as an overall group model that included data from 42 monitoring stations representing 39 water bodies. In the group model, nutrient loadings, water temperature, salinity, and water column stability were the best predictors of chlorophyll-a concentration in the 39 water bodies represented by the model. For the single station model, we found that nutrient loading was the most abundant predictor for chlorophyll-a concentration, followed by temperature and salinity. As expected, we found a positive slope coefficient between nutrient loadings and chlorophyll-a concentration in all but two stations.

Model evaluation plots and performance statistics revealed that most of the models could capture the levels and year-to-year variation in chlorophyll-a concentrations reasonable well indicating that the models can be used to produce reliable predictions of chlorophyll-a concentrations in Danish coastal waters. Not surprisingly, model performance for single station models was, in general, better than for the group model, but the risk of “over parameterization” is also higher for single station models.

Dansk resume

Vi har udviklet bayesianske hierarkiske modeller, der beskriver klorofyl-a koncentrationen i de danske vandområder, som en funktion af landbaseret kvælstof- og fosforbelastning, fysisk-kemiske og klimatiske prediktorer. Formålet med modeludviklingen var at tilvejebringe værktøjer, der kan anvendes til at estimere reference tilstand for klorofyl samt bestemme den maksimale næringsstoftilførsel, som vil sikre målopfyldelse (målbelastninger) til de danske farvande, der er omfattet af vandrammedirektivet.

Vi har udviklet enkelt-stations modeller for 46 danske vandkvalitetsstationer i det nationale overvågningsprogram, som tilsammen repræsenterede 43 vandområder, og udviklet en samlet gruppemodel, indeholdende data fra 42 stationer. I gruppemodellen var næringsstofbelastninger, vandtemperatur, saltholdighed og vandsøjlestabilitet de bedste uafhængige variable til at beskrive klorofyl-a-koncentrationen i de 39 vandområder repræsenteret af modellen. For enkelt-station-modellerne fandt vi, at næringsstofbelastning var den bedste variable til estimering af klorofylkoncentration, efterfulgt af vandtemperaturen og dernæst salinitet. Som forventet fandt vi en positiv hældningskoefficient mellem næringsstofbelastning og klorofyl-a- koncentrationen i alle, undtagen to, stationer

Modelevalueringsplot og model-performance-statistik viste, at de fleste modeller kunne fange niveauer og år-til-år variationen i klorofyl-a koncentrationen i de enkelte vandområder og at der var en rimelig god overensstemmelse mellem modelestimater og målinger, hvilket indikerer, at modellerne kan anvendes til modelscenarier for klorofyl koncentrationer i danske kystvande.

1 Introduction

In this study, a Bayesian approach was used to develop chlorophyll-a (chl-a) models to support the establishment of the Danish River Basin Management Plans 2021-2027 (RBMP 2021-27) as part of the implementation of the Water Framework Directive (WFD). Bayesian statistics has gained popularity in recent years, as modern computing power has facilitated the development of new algorithms and model types. Further, it was recommended by the panel of international experts evaluating models and methods used for the Danish RBMP 2015-2021 (Herman et al. 2017).

In Bayesian statistics, parameter estimates are not considered point estimates, but rather expressed as probability distributions, involving more uncertainty information. Parameter uncertainty is quantified using the prior knowledge from earlier studies or literature, along with the sample data, which is not the case in frequentist or classical statistics where parameter estimates are inferred from sample data only. For ecological models, it is essential to have information on the uncertainty in the parameter estimates and the resulting model predictions (Beck 1987; Ellison 1996; Omlin & Reichert 1999). Thus, the Bayesian perspective is more comprehensive, and incorporation of prior knowledge makes it more consistent in terms of the scientific process of progressive learning (Germano 1999) as well as in the policy practice of adaptive management (Walters 1986). Bayesian statistics has been used previously in water quality modeling (Malve & Qian 2006; Gronewold et al. 2010; Cha et al. 2016).

The chl-a concentration is a crucial indicator for assessing ecological status as part of the WFD. In Denmark, the chl-a indicator is defined as the average concentration from May to September and this indicator has been intercalibrated with Sweden, Germany, and Norway as part of the intercalibration process. The main purpose of this study was to develop reliable Bayesian models for the chlorophyll-a indicator for as many coastal water bodies as possible, allowing for quantification of the relationship between the response variable chl-a and the predictor variables, especially the nutrient loadings, which can be managed.

2 Materials and methods

2.1 Data collection and preparation of predictor and response variables

The data used for model development consist of chl-a concentration (chl-a, µg L^{-1}) as response variable and data describing the conditions assumingly affecting the coastal environmental status, i.e., nutrient loadings and climate variables (predictor variables). Data used to construct time series of chl-a data were obtained from the Danish National Aquatic Monitoring and Assessment Programme database (DNAMAP). Data for the predictor variables (nutrient inputs, salinity, sea surface temperature, and buoyancy frequency) were also obtained from DNAMAP, whereas irradiance and wind data were obtained from two weather stations in Copenhagen and one in Sprogø.

Chl-a concentration is an indicator of phytoplankton biomass, which is one of the biological quality elements required to assess ecological status according to the Water Framework Directive (WFD). The chl-a indicator is defined as the average chl-a concentration from May to September and was computed based on average chl-a concentration measurements from the surface down to 10 m depth.

Nutrient loadings consisted of land-based nitrogen loadings (N, tonnes y^{-1}) and phosphorous loadings (P, tonnes y^{-1}), and physical variables included were salinity, sea surface temperature, buoyancy frequency, irradiance, and wind energy. In general, data were sampled with weekly to biweekly intervals.

For all the semi-enclosed water bodies, nutrient inputs were provided as monthly values, obtained from integrating continuous flow measurements and nutrient spot samples from gauged catchments combined with model predictions from ungauged catchments. In the relatively open coastal areas, a larger but still local catchment was used.

Salinity and water temperature predictor variables were calculated as monthly means of the mean salinity or temperature of the surface layer (0-10 m).

Buoyancy frequency was calculated as the Brunt-Väisälä buoyancy frequency (N) based on the difference between surface (0-1 m) and bottom density (1 m above bottom):

$$N = \sqrt{-\frac{g}{p0}\frac{dp}{dz}}$$

where g is the regional gravitational constant (9.82 ms^{-2}), $p0$ is the potential density (surface density, kg m^{-3}), dp is the difference between bottom and surface density (kg m^{-3}), dz is the depth difference between bottom and surface (m) and N is the buoyancy frequency (s^{-1}).

The mechanical force of the wind on the water surface is proportional to the cubed wind speed (Alexander et al. 2000), and we, therefore, cubed the wind speed to obtain a relative measure on the wind energy delivered to the sea surface.

Irradiance data were obtained as half-hourly values of global irradiance from 1990 to 2017, and these were then converted to PAR values based on an algorithm from Copenhagen University. Data gaps were filled with data from Sprogø (SPØ) after adjustment of the level based on the maximum level of irradiance (0.96 of the level measured in Copenhagen). Data from the two sites in Copenhagen (HCØ and HBG) have the same level and slope, and the final unit is µmol photons m^{-2} s^{-1} calculated from global irradiance (W m^{-2}).

Data from 1990 to 1993 were hourly data, which were interpolated linearly to obtain 30 min intervals using the "Proc Expand" procedure in SAS. For some years, values were adjusted (dark values subtracted) due to a significant sensor offset. In addition, all values below 2 µmol photons m^{-2} s^{-1} were set to zero due to low sensor sensitivity within that range and problems in some years with a dark offset. This is significant in some winter months, where a dark offset may constitute a substantial part of the daily sum. Finally, the data were translated into monthly mean values. The remaining gaps were filled with average values for the same day and time from other years.

Model development was restricted to data from monitoring stations within the WFD zone of water bodies with at least 15 years' data series between 1990 and 2017, ensuring that both year-to-year variations as well as potential long-term trends could be resolved. Only time series with a minimum of one bi-monthly observation was included to have enough data points to get robust monthly interpolated values.

The data were filtered and interpolated linearly between observations using the expand procedure to obtain daily values using statistical software SAS® (https://support.sas.com/documentation/onlinedoc/ets/132/expand.pdf). Before performing interpolation, chl-a values were log-transformed, assuming that these observations were lognormal distributed. From daily values, monthly averages were calculated before the values were back transformed to the original scale. The monthly means on the original scale were averaged to gain annual summer means (May-September). These summer means were combined with annual nutrient inputs from January to September, and different physical variables aggregated over the same seasonal window as chl-a (May-September).

The two predictor variables, nitrogen inputs, and phosphorous inputs, were highly correlated. To avoid potential collinearity between these predictor variables, the variables were considered as on latent variable referred to as "load". This was done using principal component analysis (Wold et. al.1987) performed on N and P inputs to obtain on common vector, the first principal component (PC), which explained most of the total variation (an average of 91 %) in the two datasets. The first PC was used instead of the original N and P inputs. Predictor variables were measured in different units and scales. Therefore they were standardized to a mean of zero and a standard deviation of one, making the variances of predictor variables comparable.

Scatter plots of chl-a vs. load (first principal component of nutrient loadings) for different stations were used to identify deviating observations from the typical "behavior," and such outlier observations were discarded before the chl-a modelling.

2.2 Bayesian approach

To quantify the relationship between the chl-a and the predictor variables, i.e. nutrient loadings and climate variables, we used a Bayesian modeling framework.

Bayesian inference is a way of combining information from data along with the prior knowledge from expertise, earlier studies, or literature. Bayesian statistics is based on principles of conditional probability. It interprets probability as a measure of believability or how confident we are in an event occurring.

Bayes' original theorem applied to point probabilities given as follows:

$$p(B/A) = \frac{p(A/B)p(B)}{p(A)}$$

The theorem illustrates that a conditional probability for event B given event A is equal to the conditional probability of event A given event B, multiplied by the marginal probability for event B and divided by the marginal probability for event A.

In other ways, Bayes' rule states how the prior information $p(B)$ and the likelihood $p(A/B)$ are combined to arrive at the posterior distribution $p(B/A)$. *p(A)* is often ignored since it is, in many cases difficult to calculate and can often be assumed constant.

Thus, we can write Bayes' rule as:

$$posterior \propto likelihood * prior$$

where likelihood is likelihood function, which reflects information about the parameters contained in the data, and the prior is prior probability distribution, which quantifies prior belief about the parameters before observing data. Posterior is posterior probability distribution; prior distribution and likelihood are combined to form the posterior distribution, which describes total knowledge about the parameters after the data have been observed (Gelman et al. 2003; Glickman & van Dyk 2007).

In the Bayesian method, model parameters intercept and slopes are considered as random variables, and prior belief about these parameters so-called prior probability distribution is assigned. Thus, the Bayesian approach incorporates prior understandings, and evidence, to produce new posterior beliefs. Additionally, Bayesian inference quantifies the uncertainty explicitly, which is appealing in environmental decision-making (Gelman et al. 2003; Ellison 2004; Clark 2005).

In the current study, chl-a prediction models for 42 coastal monitoring stations were fitted simultaneously with the random intercept random slope using the Bayesian hierarchical modeling approach. Where chl-a at all Danish water bodies are assumed to be described by the overall model, but with a random slope and intercept on individual station level. Combining data from different sources, e.g., different sampling locations within similar river types using hierarchical modeling method, often results in improved model accuracy and reduced model uncertainty (Qian et al. 2004, 2005). In a hierarchical structure, each subsystem is believed to be unique; however, some

commonality exists between other subsystems, which is captured in the population structure (Borsuk et al. 2001). The combining of data using a hierarchical method may be useful in a variety of model applications within aquatic and ecological science where we need to estimate empirical parameter estimation. However, available local data are very limited (Borsuk et al. 2001).

Four additional coastal monitoring stations had broader chl-a range and had high random effects, compared to other monitoring stations when combined in a Bayesian hierarchical modeling approach due to very high specific nutrient loading and low retention time for one of the stations (Randers inderfjord, Hjarbæk fjord, Mariager fjord, and Nissum fjord). Thus, for these four stations, non-hierarchical station-specific Bayesian models were developed. Additionally, to compare with the 42 grouped station model results, non-hierarchical station-specific models were developed for all of these stations. Developing separate models for each monitoring station, i.e., "no pooling", is essential as many factors can affect the relationship between the model response variable chl-a and the predictor variable to differ among monitoring stations. However, focusing on site-specific features increases the risk of overfitting. A comparison of results (in terms of parameter posterior distribution) from the hierarchical model and the station-specific models will qualify the balance between robustness (hierarchical model) and precision (site-specific model). In the ideal case, the differences between the hierarchical and station-specific models are only minor, and the model will be both precise and robust.

The Bayesian linear model used in this study summarized as shown:

$$Chla_{observed_{ij}} \sim Gamma\left(chl_{true_{ij}}, \nu\right)$$

The chl-a data are right-skewed and always positive; thus a reparameterized gamma distribution is used to model the chl-a concentration. It will also ensure that the model will put more emphasis on the lower values with less variation.

Compare to the traditional gamma parameterization:

$$shape = \frac{chl_{true_{ij}}}{scale} \text{ and } scale = \nu$$

Where $chl_{true_{ij}}$ is the mean and ν is the scale parameter.

The mean model for the hierarchical (grouped station) model:

$$Chla_{true_{ij}} = (\mu_\alpha + \alpha_j) + X_{1,ij}\left(\mu_{\beta^{x1}} + \beta_j^{x1}\right) + X_{2,ij}\left(\mu_{\beta^{x2}} + \beta_j^{x2}\right)$$

where

μ_α is the group mean intercept parameter.

α_j is the deviation from mean intercept for grouped station model.

$X_{1,ij}$, $X_{2,ij}$, … are predictors of ith sample in jth station.

$\mu_{\beta^{x1}}$, $\mu_{\beta^{x2}}$,… are group mean slope parameters for respective predictor in a grouped station model.

β_j^{x1} , β_j^{x2} … are deviation from mean slope for grouped station model for respective predictor variable.

The mean model for station-specific model:

$$Chla_{true_{ij}} = \alpha_i + X_{1,ij}\,\delta_j^{x1} + X_{2,ij}\delta_j^{x2}$$

where $Chla_{ij}$ is i[th] sample of chl-a concentration, which occurs in the j[th] station.

α_i is the intercept term for station-specific model.

δ_j^{x1} , δ_j^{x2} … are station-specific slope parameters.

The non-informative prior distributions were used for model intercept and overall slope parameters:

α_j , $\mu_{\beta^{x1}}$, $\mu_{\beta^{x2}}$ $\sim N(0{,}100)$

α_i , δ_j^{x1}, δ_j^{x2} $\sim N(0{,}1000)$.

In the grouped station model, station-specific slope parameters (i.e., mean slope deviation) had narrower priors, which intended to regularize (McElreath 2016): β_j^{x1} , β_j^{x2} ... $\sim N(0, \sigma)$. Such regularization in the hierarchical approach attempts to improve the generalizability of the model but with station-specific slope and intercept as random effects. This will help reduce uncertainty and potential bias, which may occur while fitting station-specific models. Thus, a hierarchical model can be generalized to all representative monitoring stations.

Scale parameter ν had non-informative prior:
ν $\sim Uniform(0{,}100)$

Initial values were defined for parameters as shown:

α_i = mean of $Chla_i$

$\mu_{\beta^{x1}}$, $\mu_{\beta^{x2}}$, δ_j^{x1}, δ_j^{x2}… = 0 and

ν = Standard deviation of $Chla_i$

Simulation of parameter posterior distributions was performed using Hamiltonian Monte Carlo (HMC) - a Markov chain Monte Carlo (MCMC) method. The MCMC method allows the user to sample all unknown parameters using joint posterior distributions that otherwise cannot be directly calculated (Gilks et al. 1996; Gamerman & Lopes 2006). In this study, sampling consisted of 20,000 iterations and 4,000 warmup iterations. Initially, two independent chains were used to sample from, but in the final models, one chain was used, which was sufficient. The mean value of the posterior samples was considered as the estimate of each parameter.

Convergence diagnostics such as effective number of samples, trace and posterior density plots were evaluated to ensure that sufficient number of chains was used, trajectory of the chain was stationary around the similar values, mixing was good and posterior had appropriate target distribution (McElreath 2016).

Statistical software SAS® (SAS 9.4, SAS Institute Inc, Cary, North Carolina, USA) was used for relevant data extraction from database and data management. Bayesian analysis was performed using R software (R Core Team 2018) using Rethinking (McElreath 2016) and RStan (Stan Development Team 2019) packages.

2.3 Choosing informative predictor variables

In this study, six predictor variables were available for chl-a prediction. However, not all predictor variables are of importance for chl-a modeling at different monitoring stations. We used the method proposed by Lindeman, Merenda, and Gold (LMG) (Lindeman et al. 1980) to exclude predictor variables with no explanatory power for chl-a. The method uses sequential sums of squares from the multiple linear regression model, and overall assessment is obtained through the coefficient of determination (R^2) partitioned by averaging over orderings among predictors. R package '*relaimpo*' (Grömping 2006) was used to perform this analysis. For a few selected stations, partial least squares regression (Wold et al. 2001) was also used to check for variable importance for prediction, and the variable importance was then compared with the LMG method.

2.4 Model comparison and final model selection

Once the informative predictors were selected based on their relative importance in predicting chl-a, different combinations of importance predictors were used in the Bayesian approach. The Watanabe-Akaike or widely applicable information criterion (WAIC) (Watanabe, 2010) was used to compare the model with different predictors. Lower WAIC values correspond to better model performance, and therefore models with the lowest WAIC were considered as the best model. However, for some monitoring stations, the final selected model was chosen based on the best balance between both WAIC value and R^2 value.

WAIC is regarded as an improvement on the deviance information criterion (DIC) for Bayesian models and is defined as (McElreath 2016):

WAIC = -2(lppd - pWAIC)

where lppd is log-pointwise-predictive density, averaged over the posterior distribution, given by:

$$lppd = \sum_{i=1}^{N} \log \Pr(yi)$$

$\Pr(yi)$ is the average likelihood of the i^{th} observation of training sample.

pWAIC is the effective number of parameters, given by:

$$\text{pWAIC} = \sum_{i=1}^{N} \text{V}(yi)$$

$\text{V}(yi)$ is the variance in log-likelihood of the i^{th} observation of training sample.

2.5 Dunn-Smyth residuals

The current study used gamma distribution to model the chl-a concentration, and therefore the obtained residuals were far from normally distributed residuals. Thus, randomized quantile residuals or Dunn-Smyth residuals were calculated and used as a diagnostic tool. Dunn-Smyth residuals are based on the principle of inverting the fitted distribution function at each response value to obtain the equivalent standard normal residuals (Dunn & Smyth 1996). Dunn-Smyth residuals are obtained by using two transformations; first, transforming the response values into approximately independent uniformly distributed random variables using the fitted cumulative distribution function and then the inverse of the standard normal cumulative distribution function is used to obtain variables that are approximately independent with standard normal distribution.

2.6 Posterior predictive evaluation

- Model performance for both the grouped station model and single-station models was assessed and quantified by comparing model predictions and observations of Chl-a concentrations using a suite of quantitative and visual measures listed below.

- Time series plots of modeled and observed Chl-a values (appendix A+B) together with the 95 % highest posterior density interval (HPDI) for each predicted observation obtained using samples from a posterior density.

- Plots and correlation analysis of observed vs. modeled Chl-a values (appendix A+B) were used to evaluate the ability of the model to capture the variation in observed Chl-a values. The coefficient of determination R^2 (Table 2.a and 2.b) was used to quantify the variability captured by the model and an F test (table 2a and 2b) was used to determine if the correlation was significant.

- Time series plots of residuals (appendix A+B) were used to identify any non-random pattern and the ability to capture year-to-year variation. An F test (Table 2a and 2b) was used to test if the bias was time-dependent (i.e. if the regression between the residuals and time was significant).

- Plots of residuals vs observed chl-a values (appendix A+B) were used to detect if residuals were distributed as expected, i.e. “U-shaped” with smallest residuals close to the mean and larger residuals for small or large chl-a values. The Shapiro-Wilk test for residuals was used to test if the residuals were normally distributed.

- Mean Dunn-Smyth residuals (bias) were used to quantify any systematic deviation between model results and observations

- Root mean square error (RMSE) was used to assess the applicability of the model to capture high values or produced high values not reflected by the observations (Table 2a and 2b).

As there do not exists any objective or formal criteria for determining when a model “is good enough” the following guidelines were used to identify any potential problematic areas that might influence overall model performance and the applicability of the model to perform model scenarios.

Table 1. Overview of statistical model assessment methods used for evaluation of model performance.

Statistical method	**Description**
Coefficient of determination (R^2) for the correlation between model results and corresponding observations	Quantify the variation captured by the model and should be as close to 1 as possible. R^2 and the assessment criteria is only meaningful if there is "sufficient" variation in the data
Significance test (F-test) for the correlation between model results and observations	Model results and observations should be significantly correlated
Average bias of model results and observation	Average bias identify any overall systematic deviation between model and observations and should be as close to 0 as possible
Significant test (F test) for the correlation between time and residuals	Correlation between time and residuals should be not significant as significant time trends in residuals could pose a problem for model scenarios
Root mean square error (RMSE)	RMSE should be as small as possible and not exceed the standard deviation of the observations.
Shapiro-Wilk test for residuals	Test if residuals are significantly different from a normal distribution. NS implies that residuals are normally distributed as expected

3 Results and discussion

3.1 Bayesian Chl-a models

In total, 46 monitoring stations fulfilled the requirement of more than 15 years of regular monitoring, and data from these stations were included in the model development.

The grouped (hierarchical) Bayesian model included 42 monitoring stations covering 39 water bodies (*figure 1*). Based on the WAIC analysis four (out of six) predictor variables: load, Temperature, Salinity, and Buoyancy frequency were selected as being important for predicting chl-a concentration at these 42 sites. Although better model performance could be obtained by including additional predictor variables, these were excluded by the WAIC in order to reduce the risk of "overfitting". The MCMC simulation with given prior distributions converged quickly. MCMC runs were assessed using different convergence statistics. The effective number of samples was sufficiently large, and trace plots indicated that the center of the chain appeared to be stationary around similar values (not shown). The posterior density plot (*Appendix A, Figure A*) of each posterior parameter distribution showed it had appropriate target distribution, i.e. bell-shaped curve for the posterior distributions of parameters. Posterior density is the true probability density and summarizes the relative belief weight of each possible value of the parameter.

Four stations (*figure 1*, marked in green) showed high random effects compared to other monitoring stations when combined together in hierarchical way. Thus, these four stations were excluded from the grouped station model and station-specific Bayesian models were developed.

Single station models were developed for 46 monitoring stations covering 43 water bodies (*figure 1*). Despite the lower number of observations in the single-station models compared to the grouped station model, the MCMC runs converged quickly with trace plot indicating that the centre of the chain appeared to be stationary around the similar values and with bell-shaped curves for the posterior distributions of parameters (not shown).

Load was selected as predictor variables in 31 of the station-specific models making it the most common predictor variable followed by temperature, which was a predictor in nearly half of the station-specific models.

The parameter selection for single station models was based on LMG and WAIC. Both LMG method using multiple linear regression and partial least squares method (in combination with WAIC) and WAIC alone yielded the same list of important predictor variables for chl-a prediction when compared for selected monitoring stations. This indicates that parameter selection was relatively insensitive to the choice of method.

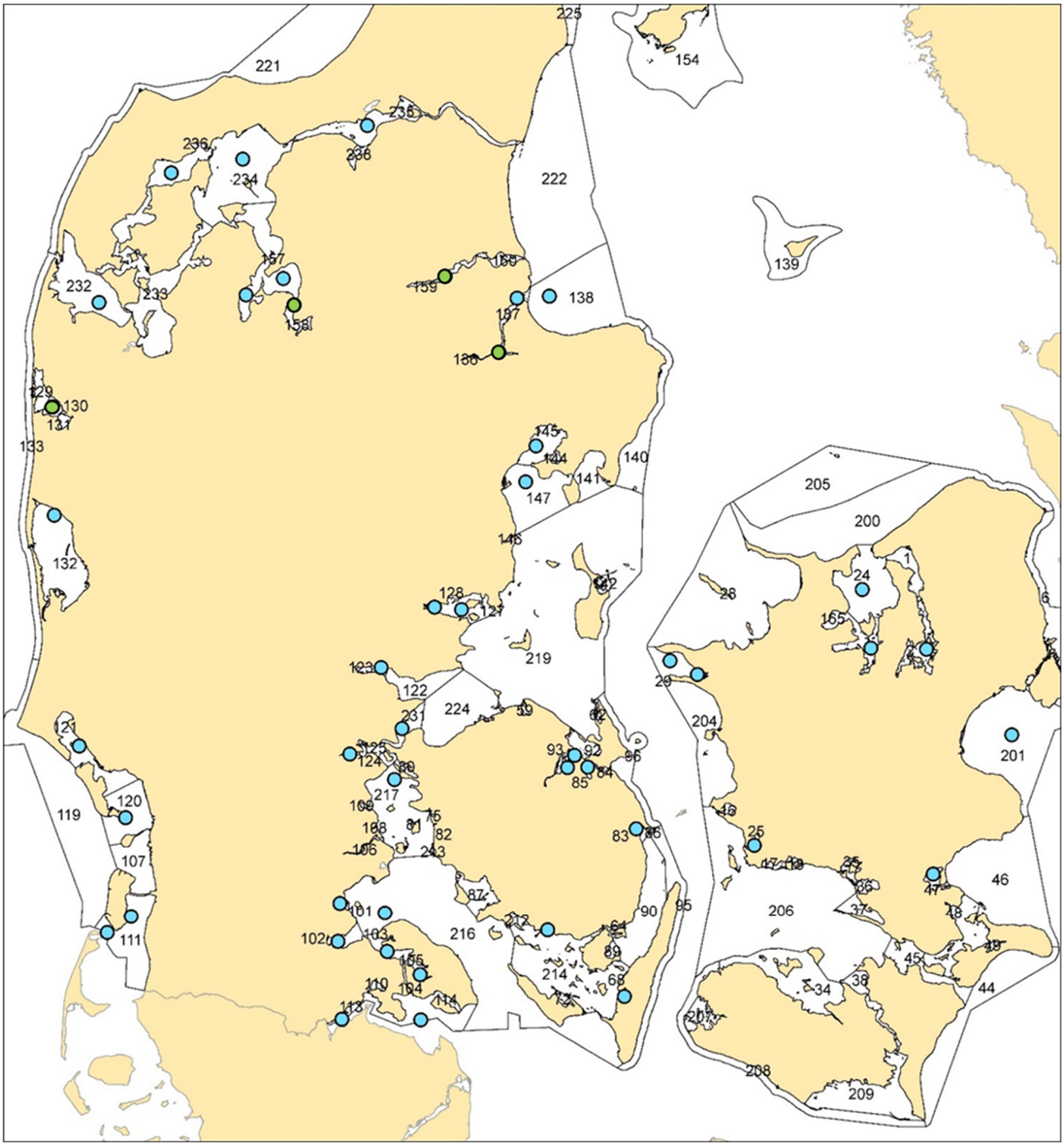

Figure 1. Map showing location of the 42 monitoring stations included in the grouped (hierarchical) Bayesian chlorophyll-a model (blue circle) as well as the 46 monitoring stations (blue and green circles) for which single station chlorophyll-a models were developed.

3.2 Model fit and evaluation

The grouped station model included 42 monitoring stations. *Table 2a* gives an overview of model goodness-of-fit using different model evaluation parameters. The grouped station model explained overall 72 % of the chl-a variation with Dunn-Smyth bias of -0.021 and RMSE of 0.94. Also, at individual station level, the grouped station model seemed to perform acceptable for most of the stations, with significant correlations between model results and observations and no obvious systematic deviations as indicated by fairly low bias and RMSE as well as normally distributed residuals without time trends. Average

bias was generally low but negative indicating an overall tendency for the model to underestimate the chl-a concentration. For 24 stations R^2 values for the correlation between model and observations was low (< 0,3) indicating that the model did not capture the variation in observed Chl-a concentrations very well.

Visual inspection of the model performance plots (Appendix A) confirmed the overall acceptable performance of the grouped station models also on the level of individual monitoring stations. Time series plots of chl-a predictive distribution means and its 95 % credible intervals plots showed that the model most often could capture the overall trend as well as the year-to-year variation in observations and only in few cases did the residual plots reveal potential undesirable patterns in residual distribution. There was, however, a tendency for a mismatch between the model-observation regression line and the 1:1 correspondence line coinciding with a low R^2 value (Appendix A).

The performance statistics for the 46 single station models (*table 2b*) indicated a better model performance of single station models compared to the grouped station models. Significant ($p < 0,05$) correlations between model and observations were obtained for 41 out of 46 of the models, and R^2 values were in general acceptable ($R^2 > 0,3$). Mean residual (Bias), as well as RMSE, were low, but there was a tendency for the models to underestimate chl-a concentrations (negative bias). The residuals were normally distributed except for three monitoring stations, and generally, no time trends in residuals were detected.

Visual inspection of the model performance plots (Appendix B) confirmed the overall good performance of the single station models. Both time series plots as well as regression plots of observed vs. modeled chl-a concentrations, showed that the models could capture the variation in the observation. In contrast to the grouped station models, regression lines of observed vs. model results were close to the 1:1 correspondence line.

Although model performance statistics are better for single station models compared to the grouped model, the generality is lower for the single station models, and the risk of over parametrisation / ”false” predictor variables is somewhat higher.

Table 2a. Summarized model performance statistics for the grouped station model. R^2 is the coefficient of determination for the relationship between observed and modelled chl-a concentrations; RMSE is the root mean squared error; Shapiro-Wilk test is the normality test results for the residuals and F test for Dunn-Smyth residual vs. time indicate if residuals are time dependent or not. Stars (*) indicate significance level, NS means "not significant"..

Station name	Station ID	R^2	Mean Dunn-Smyth residual (Bias)	RMSE	Shapiro-Wilk test for Dunn-Smyth residual	F test for Dunn-Smyth residual vs. time
Overall	Overall	0.72***	-0.021	0.94	NS	-
Kalø Vig	ARH170002	0.53***	-0.116	0,53	NS	NS
Århus Bugt	ARH170006	0.22*	-0.105	0,57	NS	NS
Hevring Bugt	ARH190004	0.45**	-0.119	0,56	NS	NS
Randers ydre	ARH230905	0.13ns	0.004	1,59	NS	*
Nakkebølle Fjord	FYN0018361	0.19*	-0.034	1,29	NS	NS
Lindelse Nor	FYN0018571	0.25*	-0.187	0,37	NS	NS
Holcken Havn	FYN0018752	0.31*	0.201	4,46	NS	NS
Kerteminde Fjord/Kertinge Nor	FYN0018843	0.31*	-0.041	1,17	NS	NS
Bredningen, Lillebælt	FYN6100021	0.15.	-0.070	0,88	NS	NS
Lillebælt nord - ved Fredericia	FYN6100051	0.12 ns	-0.062	0,85	NS	NS
Nord for Als (Lillebælt vest)	FYN6200901	0.21.	-0.088	0,75	NS	NS
Odense ydre	FYN6900017	0.10 ns	-0.042	0,96	NS	*
Odense indre	FYN6910008	0.01 ns	-0.018	1,24	NS	NS
Vadehavet Grådyb - Ho Bugt v Langli	RIB1610002	0.29**	0.053	2,74	NS	NS
Vadehavet Knudedyb	RIB1620014	0.14.	0.062	2,93	NS	NS
Ringkøbing Fjord nord	RKB1	0.08 ns	0.081	3,21	NS	NS
Køge Bugt	ROS1727	0.45***	-0.118	0,59	NS	NS
Roskilde indre	ROS60	0.37**	-0.061	0,73	NS	NS
Lister Dyb	SJY1	0.3**	-0.008	1,57	NS	*
Augustenborg Fjord	SJY12	0.09 ns	-0.044	1,29	NS	NS
Als Fjord	SJY13B	0.36**	-0.015	1,59	NS	NS
Aabenraa Fjord	SJY15	0.31**	-0.046	0,97	NS	NS
Genner Bugt	SJY19	0.02 ns	-0.051	1,07	NS	NS
Lister Dyb ved Rømø Havneby	SJY3	0.21*	0.034	2,43	NS	**
Flensborg Fjord inder	SJYKFF2	0.23*	0.025	2,45	NS	NS
Flensborg Fjord yder	SJYKFF5	0.46***	-0.077	0,77	NS	**
Præstø Fjord	STO0802008	0.33**	0.005	2,08	NS	NS
Kolding Fjord	VEJ0003350	0.37**	0.107	2,85	NS	NS
Vejle Fjord	VEJ0004273	0.59***	-0.002	1,42	NS	*
Horsens inder	VEJ0005790	0.35**	-0.007	1,42	NS	NS
Horsens yder	VEJ0006489	0.07 ns	-0.071	0,73	NS	NS
Nissum Bredning	VIB3702-00001	0.37***	-0.045	1,03	NS	**
Løgstør Bredning	VIB3708-00001	0.26**	-0.011	1,76	NS	NS
Nibe Bredning	VIB3711-00001	0.37**	-0.012	1,76	NS	*
Thisted Bredning	VIB3723-00001	0.2.	0.086	3,24	NS	NS
Skive Fjord	VIB3727-00001	0.22*	0.079	3,08	NS	NS
Lovns Bredning	VIB3728-00001	0.03 ns	0.240	4,86	NS	*
Isefjord dybt bassin	VSJ10003	0.4***	-0.057	1,21	NS	NS
Isefjord inderbredning	VSJ10006	0.26*	-0.036	1,34	NS	NS
Kalundborg Fjord yder	VSJ41007	0.24*	-0.122	0,55	NS	NS
Kalundborg Fjord inder	VSJ41008	0.17.	-0.126	0,54	NS	NS
Skælskør Fjord	VSJ51013	0.000002 ns	-0.038	1,41	*	NS

***p ≤ 0.001; **p ≤ 0.01; * p ≤ 0.05; .p ≤ 0.1; ns p > 0.1 (not significant).

Table 2b. Summarized model performance statistics for the station specific models. R^2 is the coefficient of determination for the relationship between observed and modelled chl-a concentrations; RMSE is the root mean squared error; Shapiro-Wilk test is the normality test results for the residuals and F test for Dunn-Smyth residual vs. time indicate if residuals are time dependent or not. Stars (*) indicate significance level, NS means "not significant"

Station name	Station ID	R^2	Mean Dunn-Smyth residual (Bias)	RMSE	Shapiro-Wilk test for Dunn-Smyth residual	F test for Dunn-Smyth residual vs. time	Predictors
Kalø vig	ARH170002	0.52***	-0,020	0,46	NS	#N/A	Load
Århus bugt	ARH170006	0.22*	-0,017	0,51	NS	#N/A	Load
Hevring bugt	ARH190004	0.52***	-0,041	0,46	NS	NS	Load, Sali
Randers ml	ARH230902	0.51***	-0,036	4,16	NS	*	Load
Randers ydre	ARH230905	0.56***	-0,053	1,14	NS	NS	Temp, Load, Irr, Wind
Nakkebølle fjord	FYN0018361	0.25*	-0,047	1,24	Yes	#N/A	Load, Sali, Wind
Lindelse nor	FYN0018571	0.47**	-0,038	0,18	NS	#N/A	Wind, Sali, Temp
Holcken havn	FYN0018752	0.35**	-0,075	4,01	NS	#N/A	Temp, Sali
Kerteminde fjord/kertinge nor	FYN0018843	0.51***	-0,056	0,98	NS	NS	Load, Temp, Sali
Bredningen, Lillebælt	FYN6100021	0.22*	-0,038	0,83	NS	#N/A	Load, BV, Wind
Lillebælt nord - ved Frederecia	FYN6100051	0.50**	-0,046	0,63	NS	NS	Load, Wind, Temp, Irr
Nord for Als (Lillebælt vest)	FYN6200901	0.36*	-0,056	0,65	NS	NS	Load, Wind, Irr
Odense ydre	FYN6900017	0.18*	-0,026	0,88	NS	*	Load, BV, Wind
Odense indre	FYN6910008	0.19*	-0,041	1,09	NS	#N/A	Temp, Wind, Irr
Mariager fjord	NOR5503	0.71***	-0,056	3,23	Yes	NS	Load, Irr, Sali
Vadehavet Grådyb - Ho bugt v Langli	RIB1610002	0.34**	-0,028	2,59	NS	NS	Sali
Vadehavet Knudedyb	RIB1620014	0.42***	-0,056	2,42	NS	NS	Load, BV, Temp, Irr
Ringkøbing fjord nord	RKB1	0.24*	-0,064	2,95	NS	*	Irr, BV, Temp
Nissum fjord	RKB22	0.61***	-0,085	10,80	NS	NS	Load, BV, Sali
Køge bugt	ROS1727	0.46***	-0,037	0,53	NS	#N/A	Load, Temp
Roskilde indre	ROS60	0.43**	-0,025	0,66	NS	#N/A	Load, BV
Lister dyb	SJY1	0.51***	-0,032	1,30	NS	*	Sali, Irr
Augustenborg fjord	SJY12	0,11	-0,031	1,29	NS	#N/A	Load
Als fjord	SJY13B	0.36**	-0,055	1,54	NS	NS	Load
Aabenraa fjord	SJY15	0.39***	-0,031	0,89	NS	#N/A	Load, Temp
Genner bugt	SJY19	0.48**	-0,063	0,77	NS	NS	Load, Temp, Irr
Lister dyb v Rømø havneby	SJY3	0.54***	-0,038	1,83	Yes	*	BV, Irr, Temp
Flensborg fjord inder	SJYKFF2	0.34**	-0,060	2,22	NS	NS	Load, BV, Temp
Flensborg fjord yder	SJYKFF5	0.44***	-0,025	0,74	NS	#N/A	Load
Præstø fjord	STO0802008	0.31**	-0,062	2,03	NS	#N/A	Temp, Sali
Kolding fjord	VEJ0003350	0.46***	-0,033	2,49	NS	NS	Load, Wind
Vejle fjord	VEJ0004273	0.59***	-0,024	1,29	NS	NS	Load
Horsens inder	VEJ0005790	0.46***	-0,031	1,27	NS	NS	Load, Wind, Irr, Temp

Horsens yder	VEJ0006489	0.17.	-0,026	0,62	NS	#N/A	BV, Temp
Hjarbæk fjord	VIB3729-00001	0.74***	-0,106	13,61	NS	NS	Load, BV, Sali
Nissum bredning	VIB3702-00001	0.40***	-0,028	0,98	NS	#N/A	Load, Wind
Løgstør bredning	VIB3708-00001	0.24**	-0,035	1,76	NS	#N/A	Temp, Sali
Nibe bredning	VIB3711-00001	0.45***	-0,042	1,58	NS	#N/A	BV, Wind
Thisted bredning	VIB3723-00001	0,16	-0,075	3,30	NS	#N/A	Sali
Skive fjord	VIB3727-00001	0.25**	-0,031	2,99	NS	#N/A	BV, Sali
Lovns bredning	VIB3728-00001	0,14	-0,060	4,51	NS	#N/A	Irr
Isefjord dybt bassin	VSJ10003	0.44***	-0,052	1,11	NS	NS	Load, BV, Temp
Isefjord inderbredning	VSJ10006	0.28*	-0,054	1,30	NS	NS	Load, Temp
Kalundborg fjord yder	VSJ41007	0.37**	-0,060	0,45	NS	NS	Load, BV, Sali, Temp
Kalundborg fjord inder	VSJ41008	0.32*	-0,046	0,42	NS	*	Sali, Wind, BV
Skælskør fjord	VSJ51013	0,14	-0,077	1,32	NS	NS	Load, Wind

***p ≤ 0.001; **p ≤ 0.01; * p ≤ 0.05; . p ≤ 0.1; ns p > 0.1 (not significant).

[1] Shapiro Wilk Normality test indicated residuals are significantly different from normally distributed residuals, however, Q-Q plot (not shown) indicated normality of residuals.

Load is the first principal component of nitrogen and phosphorous loadings (nutrient loadings); Wind is cubed wind speed; Temp is sea surface temperature; BV is Brunt-Väisälä buoyancy frequency for the whole water column; Irr is incoming PAR radiation; Sali is salinity in the water surface (upper 10 m).

3.3 Model coefficients

Load coefficients estimated from either the hierarchical grouped station model (table 3a) or single station models (table 3b) showed an overall positive correlation between nutrient loadings and chl-a concentrations, as expected. Only at two stations (Randers ydre and Thisted), the Load coefficient appeared to be negative in both the grouped station model and single station model.

Overall, there was a good agreement between Load coefficients estimated with the grouped station model and single station models indicating that the quantification is robust even at the single station models with fewer observations.

Table 3a. Estimated slope parameter for the grouped station model for predictor variables Load, Temperature, Brunt-Väisälä buoyancy frequency and Salinity along with the standard deviation.

Station name	Station ID	Slope for Load ± SD	Slope for Temp ± SD	Slope for Sali ± SD	Slope for BV ± SD
Overall	Overall	0.37 ± 0.08	0.17 ± 0.07	-0.15 ± 0.08	0.13 ± 0.07
Kalø Vig	ARH170002	0.43 ± 0.19	0.02 ± 0.15	0.01 ± 0.18	0.02 ± 0.15
Århus Bugt	ARH170006	0.32 ± 0.19	0.004 ± 0.15	-0.01 ± 0.19	0.02 ± 0.15
Hevring Bugt	ARH190004	0.30 ± 0.22	0.04 ± 0.17	-0.10 ± 0.22	0.02 ± 0.16
Randers ydre	ARH230905	-0.04 ± 0.31	0.35 ± 0.22	-0.30 ± 0.27	0.22 ± 0.2
Nakkebølle Fjord	FYN0018361	0.38 ± 0.25	0.15 ± 0.19	-0.21 ± 0.22	0.02 ± 0.15
Lindelse Nor	FYN0018571	0.15 ± 0.19	0.04 ± 0.15	-0.03 ± 0.16	0.01 ± 0.13
Holcken Havn	FYN0018752	0.14 ± 0.37	0.83 ± 0.39	-0.91 ± 0.4	0.09 ± 0.24
Kerteminde Fjord/Kertinge Nor	FYN0018843	0.55 ± 0.29	0.14 ± 0.2	0.07 ± 0.25	0.13 ± 0.16
Bredningen, Lillebælt	FYN6100021	0.37 ± 0.21	0.04 ± 0.17	0.04 ± 0.19	0.11 ± 0.15
Lillebælt nord - ved Fredericia	FYN6100051	0.40 ± 0.24	-0.0001 ± 0.19	0.07 ± 0.22	0.09 ± 0.16
Nord for Als (Lillebælt vest)	FYN6200901	0.40 ± 0.25	0.04 ± 0.18	0.04 ± 0.22	0.04 ± 0.16
Odense ydre	FYN6900017	0.42 ± 0.25	0.13 ± 0.18	-0.04 ± 0.23	0.07 ± 0.15
Odense indre	FYN6910008	0.12 ± 0.3	0.25 ± 0.2	-0.20 ± 0.25	0.19 ± 0.2
Vadehavet Grådyb - Ho Bugt v Langli	RIB1610002	0.02 ± 0.3	0.50 ± 0.25	-0.74 ± 0.32	0.22 ± 0.21
Vadehavet Knudedyb	RIB1620014	0.67 ± 0.31	-0.06 ± 0.24	-0.13 ± 0.27	0.15 ± 0.2
Ringkøbing Fjord nord	RKB1	0.56 ± 0.29	0.21 ± 0.24	-0.09 ± 0.26	0.32 ± 0.22
Køge Bugt	ROS1727	0.36 ± 0.17	0.11 ± 0.14	0.03 ± 0.15	-0.0005 ± 0.12
Roskilde indre	ROS60	0.47 ± 0.26	0.12 ± 0.18	-0.04 ± 0.23	0.03 ± 0.16
Lister Dyb	SJY1	0.17 ± 0.3	0.22 ± 0.21	-0.49 ± 0.3	0.12 ± 0.18
Augustenborg Fjord	SJY12	0.35 ± 0.26	0.06 ± 0.19	-0.05 ± 0.23	0.06 ± 0.16
Als Fjord	SJY13B	0.70 ± 0.25	-0.05 ± 0.2	0.09 ± 0.22	0.03 ± 0.17
Aabenraa Fjord	SJY15	0.55 ± 0.24	-0.11 ± 0.2	0.07 ± 0.21	0.05 ± 0.15
Genner Bugt	SJY19	0.37 ± 0.28	-0.01 ± 0.21	0.01 ± 0.25	0.06 ± 0.16
Lister Dyb v Rømø Havneby	SJY3	0.13 ± 0.32	0.30 ± 0.23	-0.22 ± 0.29	0.42 ± 0.27
Flensborg Fjord inder	SJYKFF2	0.60 ± 0.28	0.04 ± 0.22	-0.18 ± 0.25	0.26 ± 0.21
Flensborg Fjord yder	SJYKFF5	0.50 ± 0.21	0.08 ± 0.17	0.02 ± 0.19	-0.01 ± 0.15
Præstø fjord	STO0802008	0.18 ± 0.22	0.44 ± 0.21	-0.56 ± 0.22	0.02 ± 0.17
Kolding fjord	VEJ0003350	1.09 ± 0.36	0.02 ± 0.27	0.01 ± 0.29	0.35 ± 0.26
Vejle fjord	VEJ0004273	0.96 ± 0.28	-0.01 ± 0.19	0.11 ± 0.22	0.06 ± 0.17
Horsens inder	VEJ0005790	0.70 ± 0.27	0.13 ± 0.2	-0.04 ± 0.22	0.20 ± 0.17
Horsens yder	VEJ0006489	0.22 ± 0.25	0.15 ± 0.19	-0.08 ± 0.23	0.10 ± 0.17
Nissum Bredning	VIB3702-00001	0.61 ± 0.23	0.05 ± 0.17	0.05 ± 0.2	0.07 ± 0.15
Løgstør Bredning	VIB3708-00001	0.23 ± 0.23	0.38 ± 0.21	-0.33 ± 0.21	0.11 ± 0.16
Nibe Bredning	VIB3711-00001	0.41 ± 0.25	0.25 ± 0.2	-0.17 ± 0.23	0.27 ± 0.21
Thisted Bredning	VIB3723-00001	-0.002 ± 0.32	0.49 ± 0.27	-0.61 ± 0.32	0.18 ± 0.21
Skive Fjord	VIB3727-00001	0.42 ± 0.3	0.41 ± 0.26	-0.50 ± 0.29	0.40 ± 0.26
Lovns Bredning	VIB3728-00001	0.32 ± 0.38	0.48 ± 0.32	-0.56 ± 0.36	0.44 ± 0.3
Isefjord dybt bassin	VSJ10003	0.46 ± 0.22	0.14 ± 0.16	0.09 ± 0.19	0.19 ± 0.15
Isefjord inderbredning	VSJ10006	0.35 ± 0.27	0.28 ± 0.21	-0.12 ± 0.24	0.13 ± 0.17
Kalundborg Fjord yder	VSJ41007	0.19 ± 0.18	0.12 ± 0.15	-0.10 ± 0.18	-0.04 ± 0.15
Kalundborg Fjord inder	VSJ41008	0.12 ± 0.21	0.12 ± 0.17	-0.13 ± 0.18	0.07 ± 0.15
Skælskør Fjord	VSJ51013	0.08 ± 0.27	0.13 ± 0.2	-0.11 ± 0.23	0.16 ± 0.16

Table 3b. Estimated slope parameter for station-specific models along with the standard deviation.

Station name	Coef. for Load	SD for Load	Coef. for Temp.	SD for Temp	Coef. for Sali	SD for Sali	Coef. for BV	SD for BV	Coef. for Irr	SD for Irr	Coef. for Wind	SD for Wind
ARH170002	0,52	0,12	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
ARH170006	0,31	0,11	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
ARH190004	0,23	0,16	NI	NI	-0,32	0,16	NI	NI	NI	NI	NI	NI
ARH230902	3,96	0,98	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
ARH230905	-0,65	0,36	1,51	0,54	NI	NI	NI	NI	-1,23	0,56	-0,68	0,41
FYN0018361	0,50	0,36	NI	NI	-0,26	0,35	NI	NI	NI	NI	0,33	0,32
FYN0018571	NI	NI	0,09	0,06	-0,10	0,06	NI	NI	NI	NI	0,13	0,06
FYN0018752	NI	NI	2,13	1,03	-1,70	1,22	NI	NI	NI	NI	NI	NI
FYN0018843	1,09	0,38	0,47	0,30	0,84	0,42	NI	NI	NI	NI	NI	NI
FYN6100021	0,29	0,19	NI	NI	NI	NI	0,21	0,20	NI	NI	0,17	0,20
FYN6100051	0,20	0,22	-0,41	0,23	NI	NI	NI	NI	0,49	0,27	0,47	0,25
FYN6200901	0,35	0,22	NI	NI	NI	NI	NI	NI	0,20	0,25	0,27	0,24
FYN6900017	0,39	0,22	NI	NI	NI	NI	-0,23	0,22	NI	NI	-0,18	0,22
FYN6910008	NI	NI	0,68	0,40	NI	NI	NI	NI	-0,47	0,47	0,31	0,37
NOR5503	3,88	1,66	NI	NI	2,38	1,51	NI	NI	3,24	1,49	NI	NI
RIB1610002	NI	NI	NI	NI	-1,72	0,51	NI	NI	NI	NI	NI	NI
RIB1620014	1,14	0,63	-1,47	0,58	NI	NI	-1,08	0,73	0,88	0,54	NI	NI
RKB1	NI	NI	0,98	1,03	NI	NI	0,70	0,79	-1,52	0,99	NI	NI
RKB22	9,81	2,88	NI	NI	8,71	2,43	-4,95	2,74	NI	NI	NI	NI
ROS1727	0,40	0,12	0,20	0,11	NI	NI	NI	NI	NI	NI	NI	NI
ROS60	0,48	0,20	NI	NI	NI	NI	-0,25	0,19	NI	NI	NI	NI
SJY1	NI	NI	NI	NI	-1,21	0,39	NI	NI	0,48	0,37	NI	NI
SJY12	0,31	0,29	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
SJY13B	1,00	0,43	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
SJY15	0,53	0,21	-0,45	0,21	NI	NI	NI	NI	NI	NI	NI	NI
SJY19	0,39	0,26	-0,94	0,34	NI	NI	NI	NI	0,92	0,31	NI	NI
SJY3	NI	NI	-0,81	0,59	NI	NI	2,11	0,53	1,20	0,55	NI	NI
SJYKFF2	1,04	0,57	-0,78	0,60	NI	NI	1,12	0,60	NI	NI	NI	NI
SJYKFF5	0,64	0,16	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
STO0802008	NI	NI	0,70	0,35	-0,83	0,35	NI	NI	NI	NI	NI	NI
VEJ0003350	2,26	0,70	NI	NI	NI	NI	NI	NI	NI	NI	-0,97	0,60
VEJ0004273	1,58	0,30	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
VEJ0005790	1,00	0,33	-0,37	0,42	NI	NI	NI	NI	0,68	0,46	-0,15	0,36
VEJ0006489	NI	NI	0,17	0,24	NI	NI	NI	NI	NI	NI	NI	NI
VIB370200001	0,77	0,23	NI	NI	NI	NI	NI	NI	NI	NI	-0,24	0,22
VIB370800001	NI	NI	0,73	0,41	-0,41	0,42	NI	NI	NI	NI	NI	NI
VIB371100001	NI	NI	NI	NI	NI	NI	0,99	0,38	NI	NI	-0,70	0,37
VIB372300001	NI	NI	NI	NI	-1,81	1,07	NI	NI	NI	NI	NI	NI
VIB372700001	NI	NI	NI	NI	-0,75	0,72	1,20	0,69	NI	NI	NI	NI
VIB372800001	NI	NI	NI	NI	NI	NI	NI	NI	2,07	1,40	NI	NI
VIB372900001	20,40	8,38	NI	NI	18,09	8,41	-8,99	6,75	NI	NI	NI	NI
VSJ10003	0,46	0,24	0,38	0,22	NI	NI	0,60	0,23	NI	NI	NI	NI
VSJ10006	0,47	0,45	0,69	0,40	NI	NI	NI	NI	NI	NI	NI	NI
VSJ41007	0,13	0,14	0,14	0,14	-0,21	0,15	-0,23	0,15	NI	NI	NI	NI
VSJ41008	NI	NI	NI	NI	-0,17	0,15	0,18	0,14	NI	NI	-0,13	0,16
VSJ51013	-0,34	0,43	NI	NI	NI	NI	NI	NI	NI	NI	-0,30	0,42

Station name is the name of the monitoring station from which data were used to develop the models; Load is the first principal component of nitrogen and phosphorous loadings (nutrient loadings); Temp is sea surface temperature; Sali is salinity in the water surface (upper 10 m); BV is Brunt-Väisälä buoyancy frequency for the whole water column; Irr is incoming PAR radiation; Wind is cubed wind speed; SD is standard deviation; NI: not included (respective predictor variable is not included for that station).

3.4 Model applicability

The developed Bayesian model framework provides chl-a models for 46 monitoring stations distributed in 43 coastal water bodies. The model performance statistics and evaluation plots indicate that most of the models can be used for scenario run, at least when the scenarios are within or not too far from the model calibration area. As for all types of models, the uncertainty will increase when moving away from the calibration area.

Although the performance statistics for some of the models indicate potential problematic model performance, these models could still produce reasonable scenario results provided that the slope coefficients used in the scenarios are robust as determined by standard deviation. However, as a low model performance imply that important processes or mechanisms are not included in the model, this could potentially influence the reliability of model scenario results.

The models have been developed with the aim of producing nutrient load scenarios making the estimated Load slope coefficients and associated uncertainty particularly important. For the grouped station model Load was included as one of four predictor variables, and hence nutrient load scenarios can be performed and the uncertainty evaluated. Single station models could only be applied for nutrient load scenarios if Load was included as predictor variable in the model. For models that do not contain Load as predictor, the relation between nutrient loadings and chl-a concentrations have not been quantified. This does not necessarily imply that nutrient loadings and chl-a concentrations are not linked in that particular water body, but only that other factors are more important and that the available data do not support at quantification of the Load coefficient. Although the slope coefficient for the Load predictor was not significantly different from 0 at all stations, Load was selected as predictor variables and hence the information adds explanatory power to the model and the estimated slope coefficients should be used for running nutrient load scenarios.

4 References

Alexander M, Scott J & Deser C. 2000. Processes that influence sea surface temperature and ocean mixed layer depth variability in a coupled model. Journal of Geophysical Research: Oceans 105: 16823-16842 http://onlinelibrary.wiley.com/doi/10.1029/2000JC900074/full.

Beck MB. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resources Research 23: 1393-1442.

Borsuk ME, Higdon D, Stow CA & Reckhow KH. 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecological Modelling 143: 165-181.

Cha YK, Park SS, Lee WH & Stow CA. 2016. A Bayesian hierarchical approach to model seasonal algal variability along an upstream to downstream river gradient. Water Resources Research 52(1): 348-357.
https://doi.org/10.1002/2015WR017327

Clark JS. 2005. Why environmental scientists are becoming Bayesians. Ecology Letters 8: 2-14.

Dunn PK. & Smyth GK. 1996. Randomized Quantile Residuals. Journal of Computational and Graphical Statistics 5: 236-244.

Ellison AM. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6: 1036–1046.

Ellison AM. 2004. Bayesian inference in ecology. Ecology Letters 7: 509-520.

Gamerman D & Lopes HF. 2006. Markov chain Monte Carlo: Stochastic simulation for Bayesian inference. Boca Raton: Chapman & Hall/CRC. 342 pp.

Gelman A, Carli, JB, Stern HS & Rubin DB. 2003. Bayesian Data Analysis (2nd edition). Chapman & Hall/ CRC. 668 pp.

Germano JD. 1999. Ecology, statistics, and the art of misdiagnosis: The need for a paradigm shift. Environmental Reviews 7: 167-190.

Glickman ME & van Dyk DA. 2007. Basic Bayesian Methods. In: Ambrosius WT (Ed) Topics in Biostatistics. Methods in Molecular Biology™, vol 404. Humana Press, pp 319-338.

Grömping U. 2006. Relative Importance for Linear Regression in R: The Package relaimpo. Journal of Statistical Software 17(1): 1-27.

Gronewold A & Borsuk ME. 2010. Improving Water Quality Assessments through a Hierarchical Bayesian Analysis of Variability. Environmental Science and Technology. John Wiley & Sons, Ltd., Indianapolis, IN, 44(20):7858-7864.

Herman P, Newton A, Schernewski G, Gustafsson B & Malve O. 2017. International evaluation of the Danish marine models. https://mfvm.dk/fileadmin/user_upload/MFVM/Nyheder/Bilag_1_Evalueringsrapport_om_de_danske_kvaelstofmodeller__10._oktober_2017-2.pdf

Gilks WR, Richardson S, & Spiegelhalter DJ (Eds). 1996. Markov chain Monte Carlo in practice. Chapman and Hall/CRC, Boca Raton.

Lindeman RH, Merenda PF & Gold, RZ. 1980. Introduction to Bivariate and Multivariate Analysis. Glenview, IL: Scott, Foresman and Company.

Malve O & Qian SS. 2006. Estimating nutrients and chlorophyll a relationships in Finnish lakes. Environmental Science & Technology 40: 7848-7853.

McElreath R. 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press.

Omlin M & Reichert P 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecological Modelling 115: 45-59.

Qian SS, Donnelly M, Schmelling DC, Messner M, Linden KG & Cotton C. 2004. Ultraviolet light inactivation of protozoa in drinking water: a Bayesian meta-analysis. Water Research 38: 317-326.

Qian SS, Linden KG, & Donnelly M. 2005. A Bayesian analysis of mouse infectivity data to evaluate the effectiveness of using ultraviolet light as a drinking water disinfectant. Water Research 39: 4229-4239.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
https://www.R-project.org/

Stan Development Team. 2019. RStan: the R interface to Stan. R package version 2.19.2. http://mc-stan.org/.

Walters CJ. 1986. Adaptive Management of Renewable Resources. McMillan, New York.

Watanabe S. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research 11: 3571-3594.

Wold S, Esbensen K & Geladi P. 1987. Principal component analysis. Chemometrics and Intelligent Laboratory Systems 2(1-3): 37-52.

Wold S, Sjöström M & Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58(2): 109-130.

Appendix A: Bayesian model evaluation for the grouped station model

Figure A: Bayesian model evaluation plots. In posterior density plot, “a” is intercept, “bP” is Load, “bB” is Brunt-Väisälä buoyancy frequency, “bS” is salinity and “sigma” is residual error (v).

Figures A1-A40 and B1 to B26: Bayesian model evaluation plots. In the time series plots, grey shaded areas represent 95 % highest posterior density interval (HPDI) for each predicted observation.

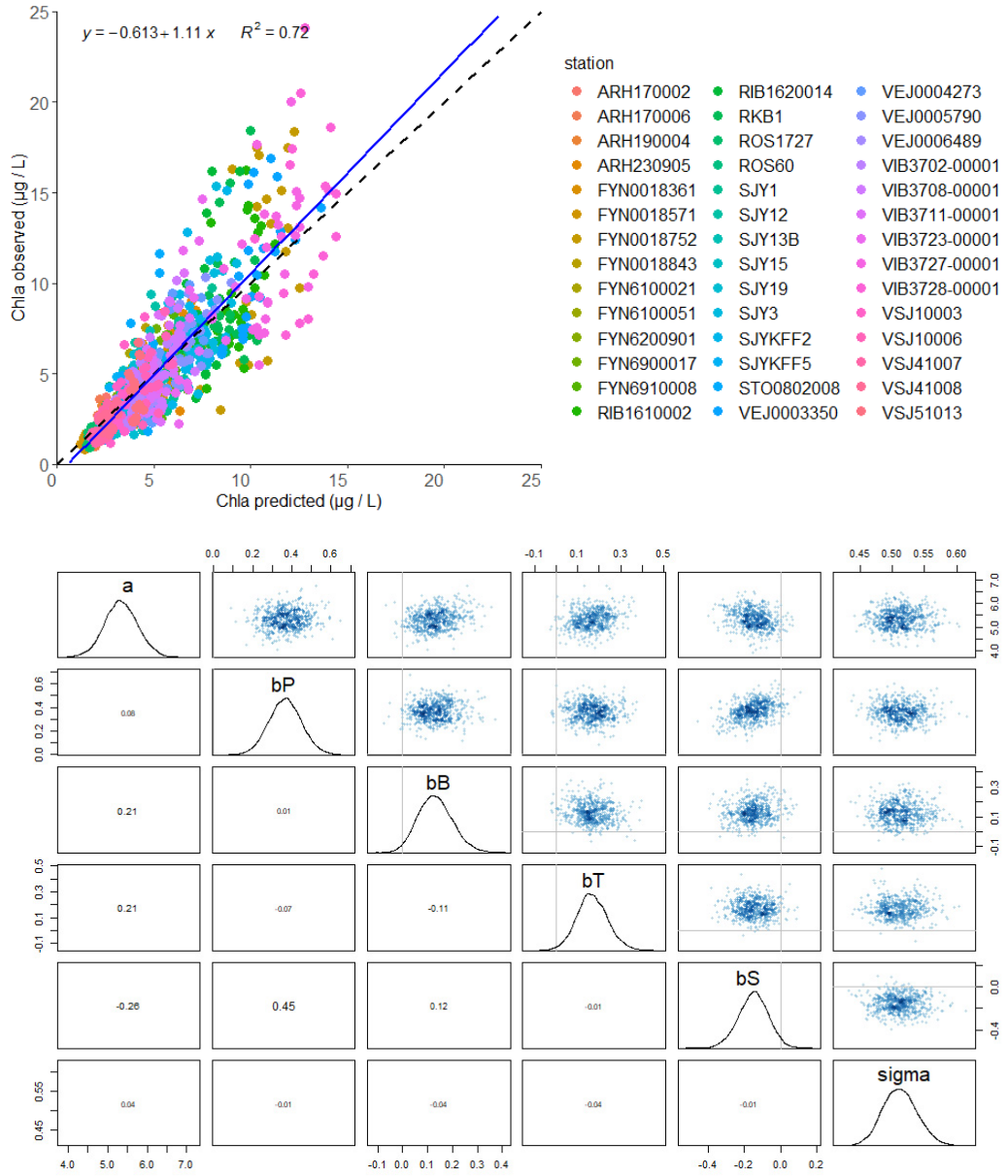

Figure A.

Kalø Vig, St. ARH170002

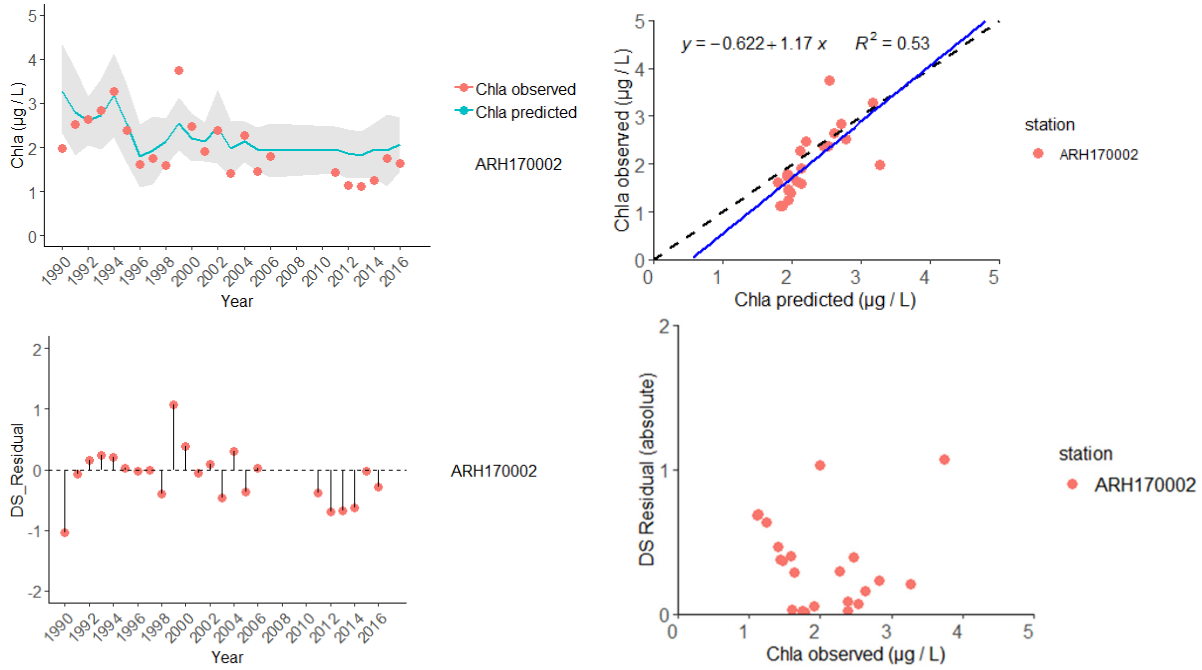

Figure A1.

Århus Bugt, St. ARH170006

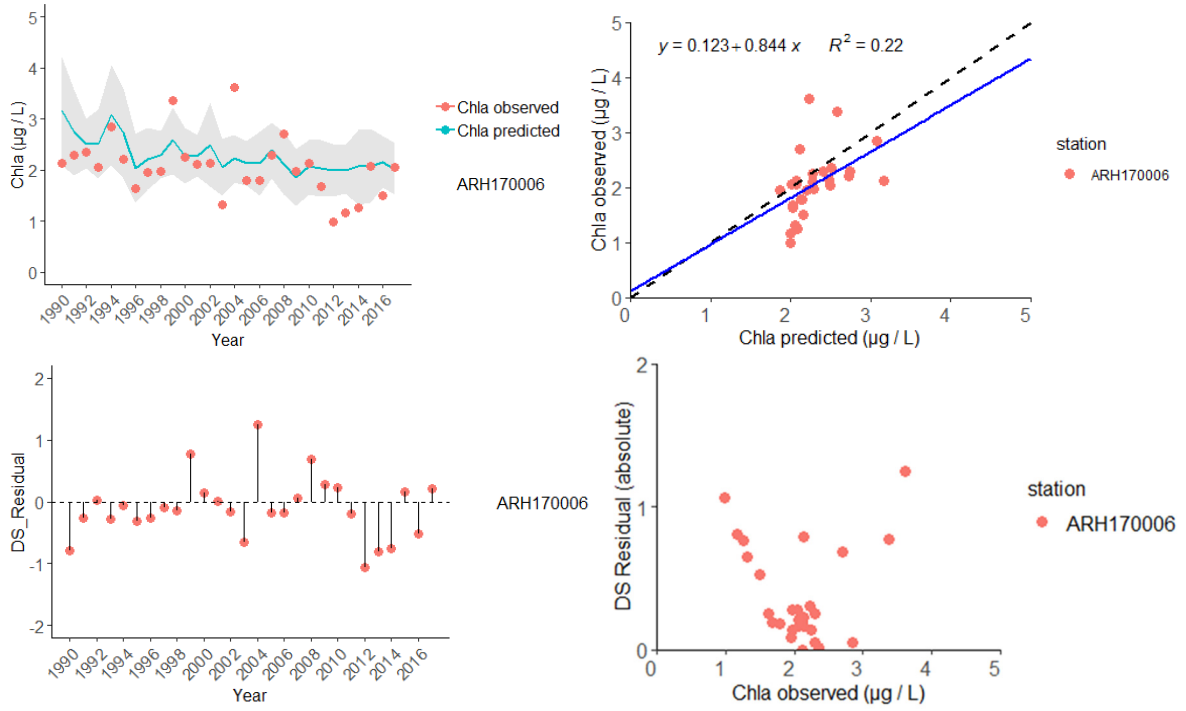

Figure A2.

Hevring Bugt, St. ARH190004

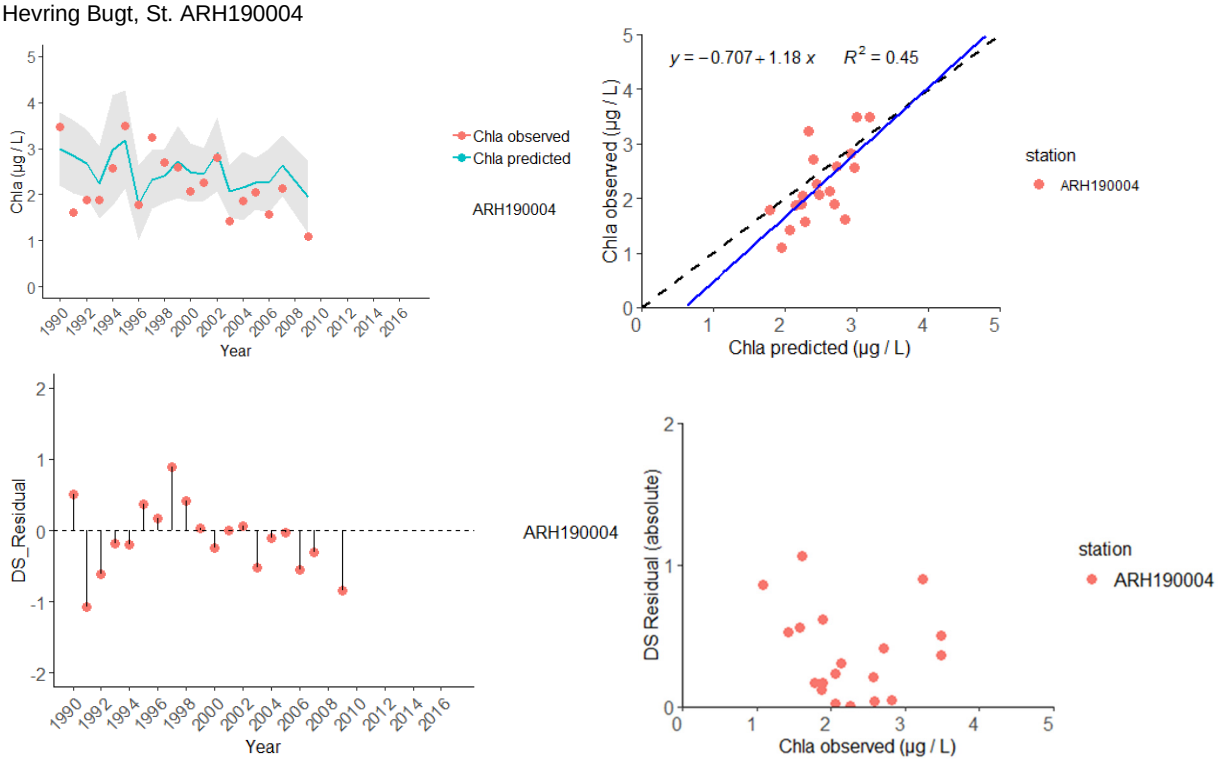

Figure A3.

Randers ydre, St. ARH230905

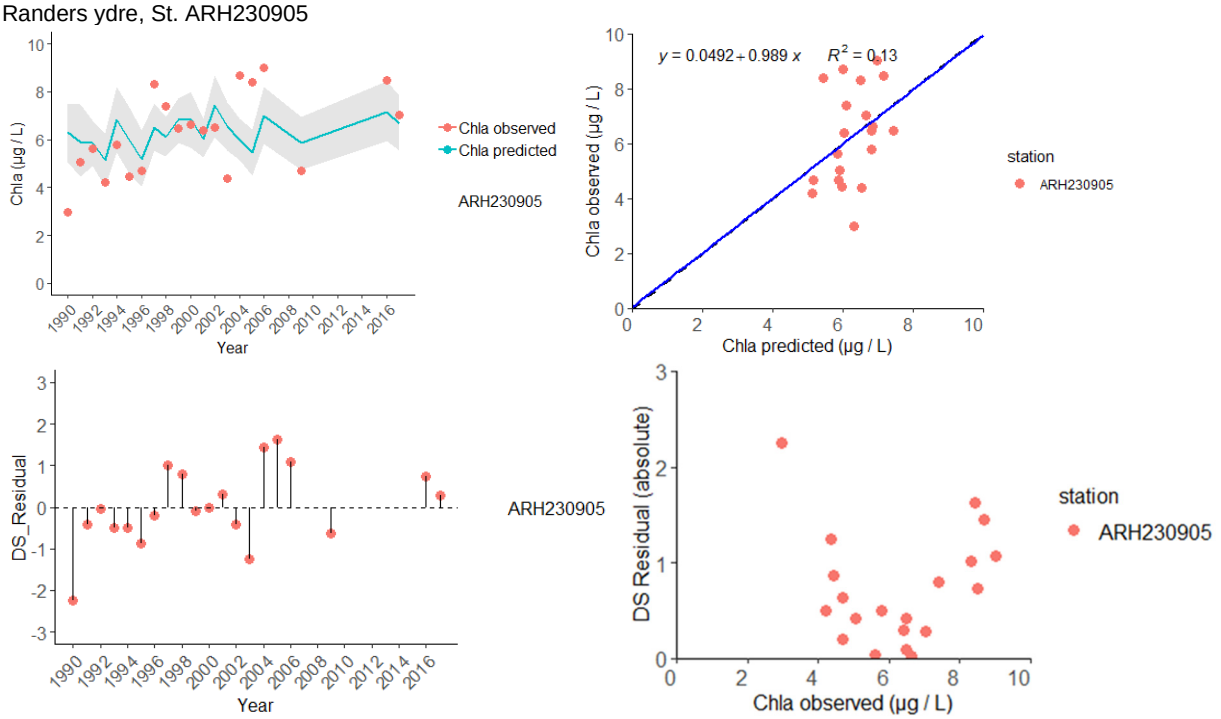

Figure A4.

Nakkebølle Fjord, St. FYN0018361

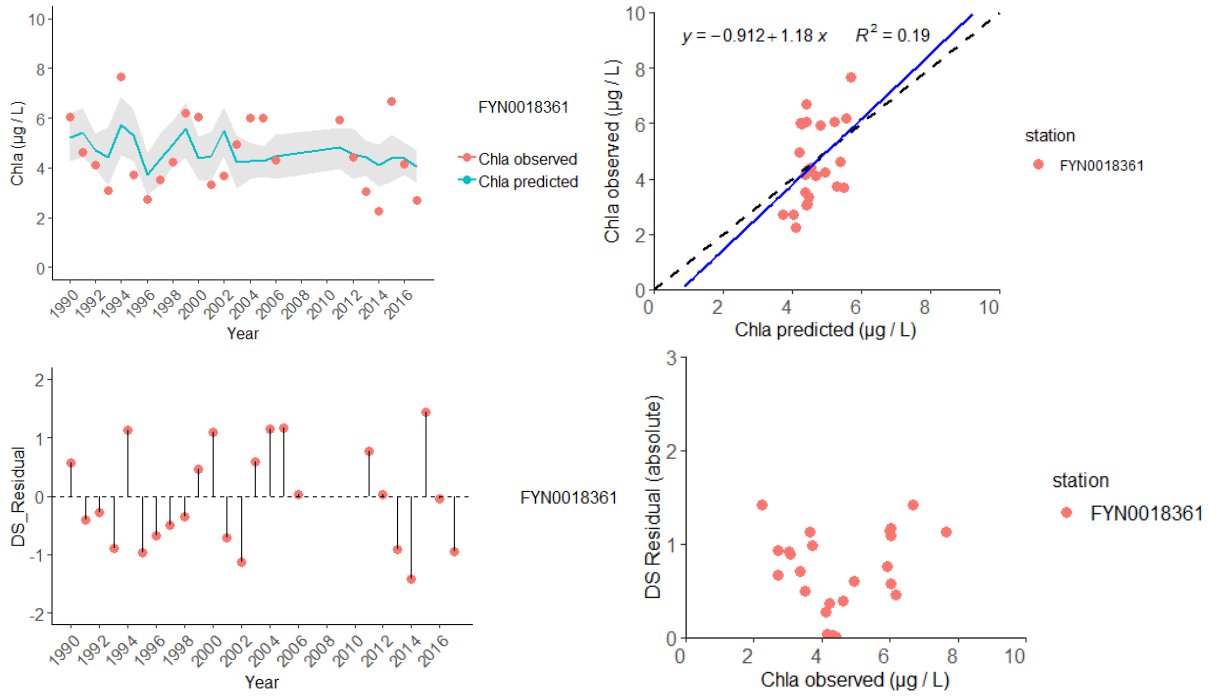

Figure A5.

Lindelse Nor, St. FYN0018571

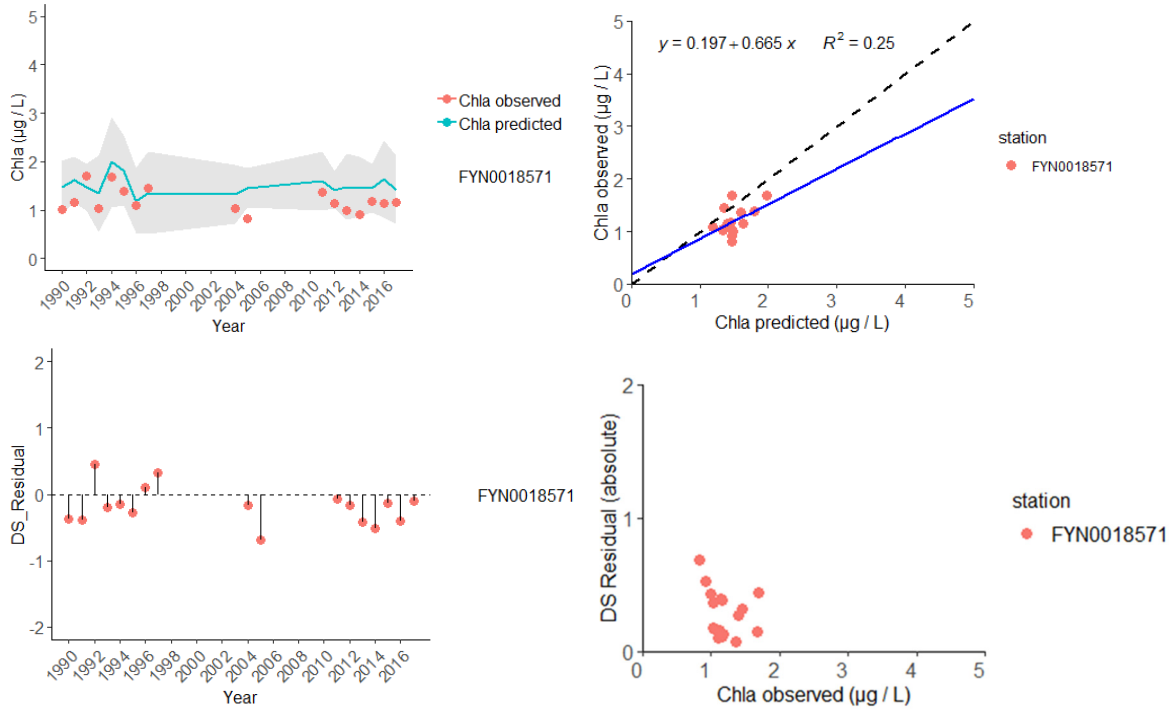

Figure A6.

Holcken Havn, St. FYN0018752

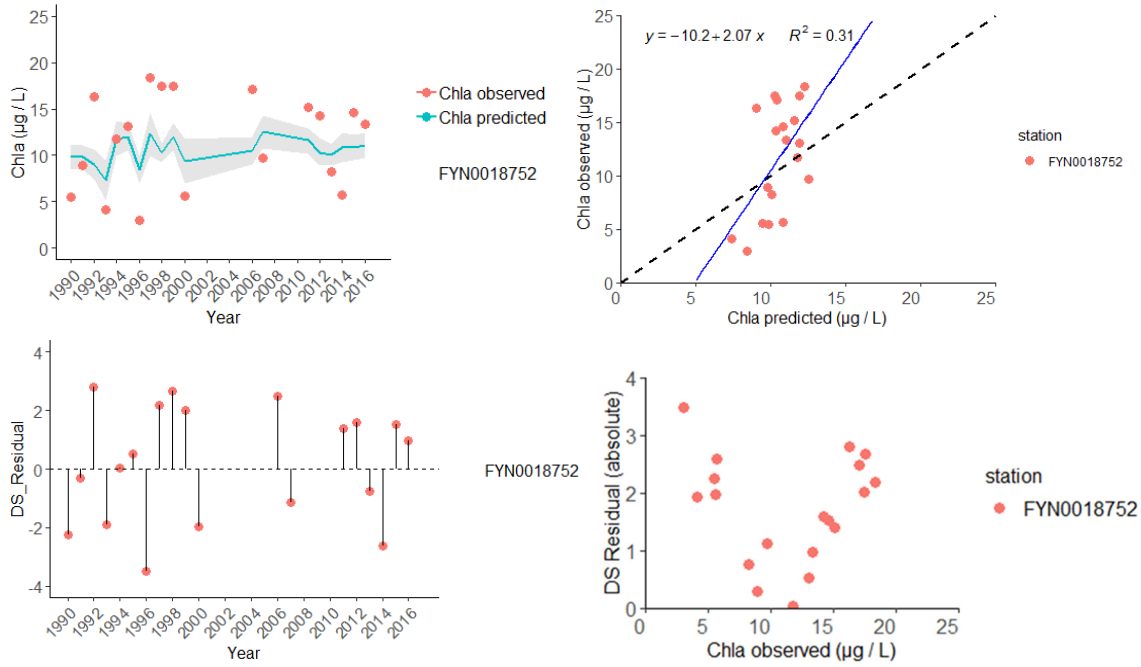

Figure A7.

Kerteminde Fjord/Kertinge Nor, St. FYN0018843

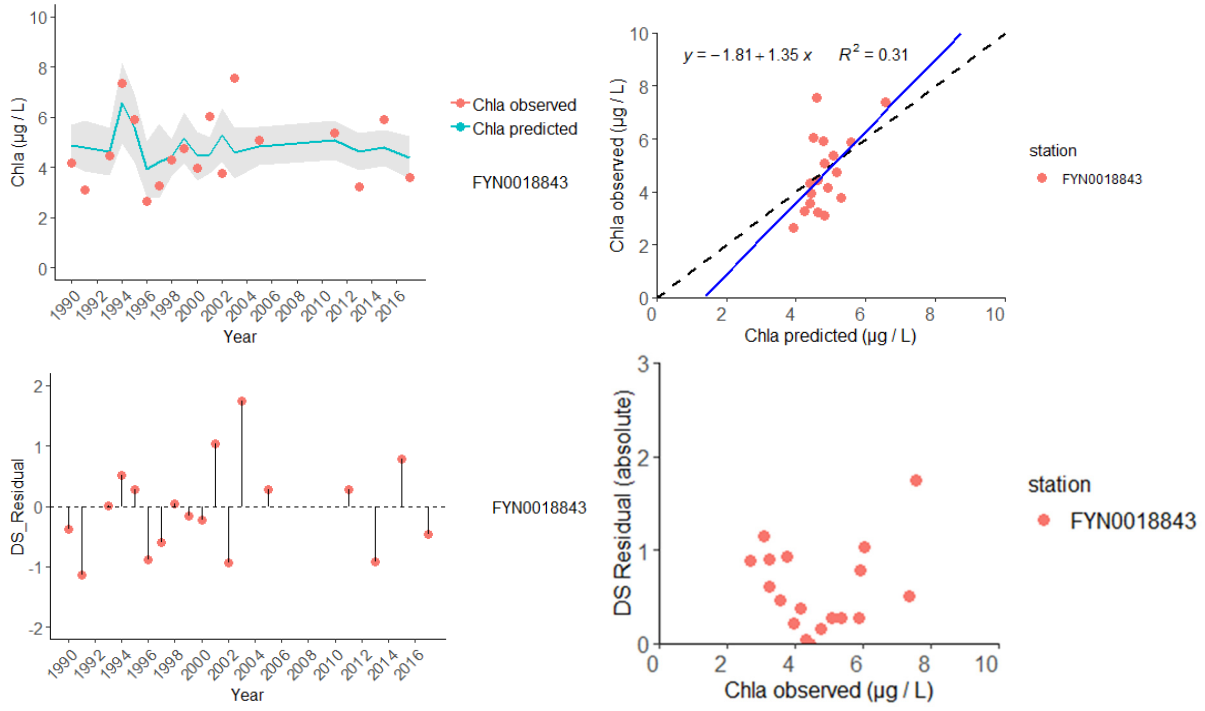

Figure A8.

Bredningen, Lillebælt, St. FYN6100021

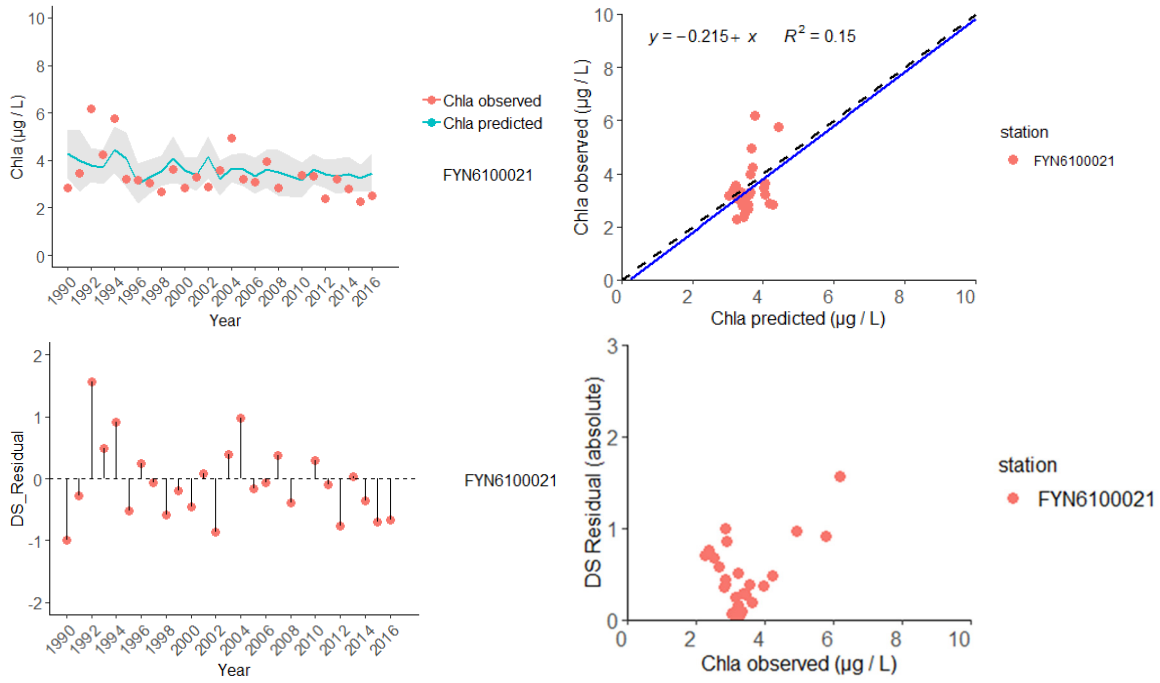

Figure A9.

Lillebælt nord - ved Fredericia, St. FYN6100051

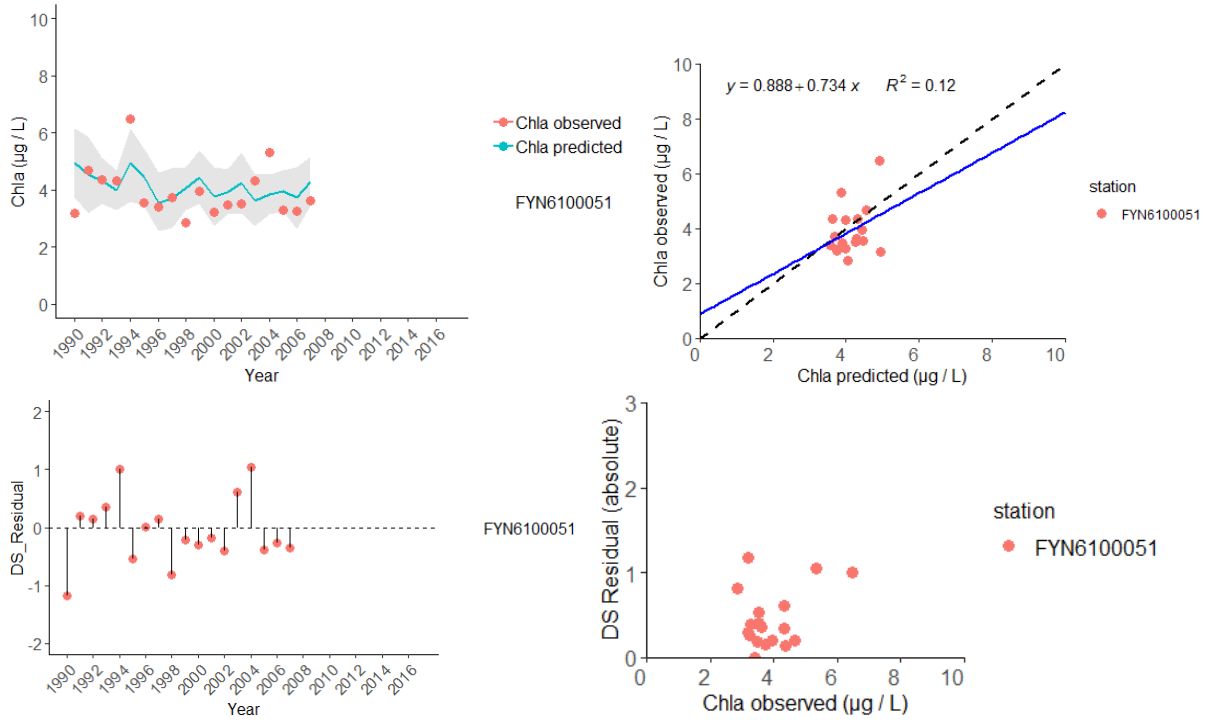

Figure A10.

Nord for Als (Lillebælt vest), St. FYN6200901

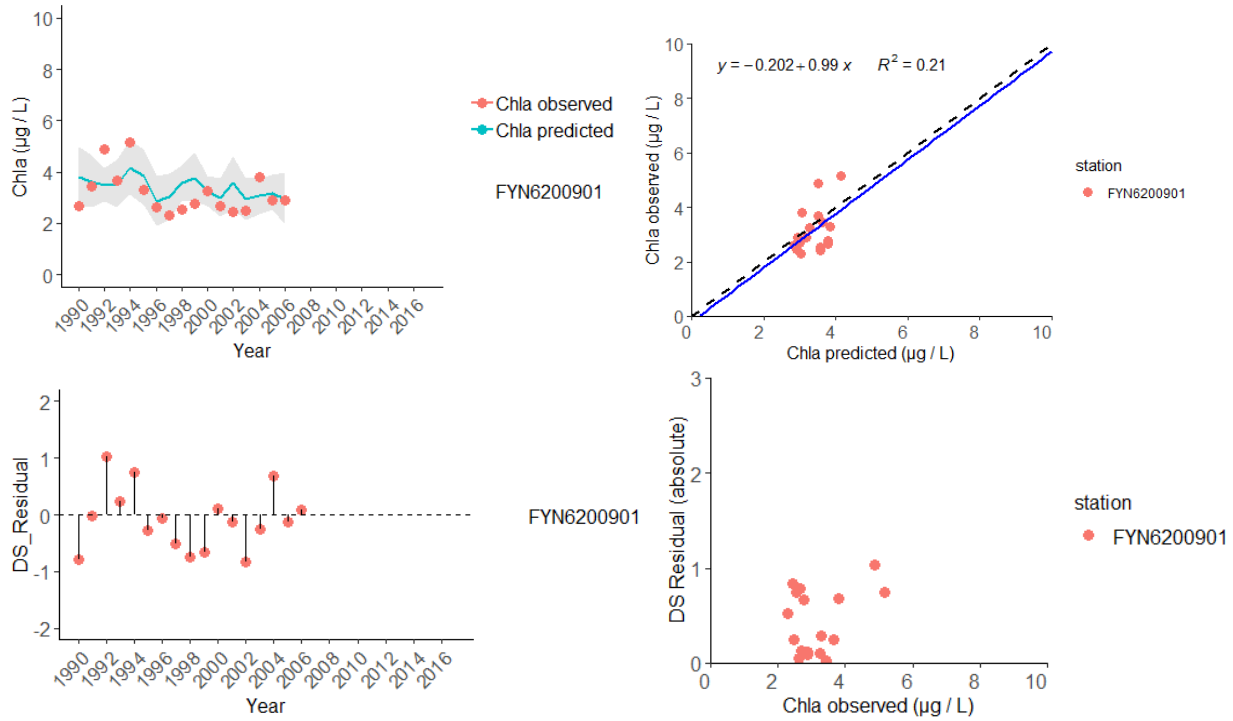

Figure A11.

Odense ydre, St. FYN6900017

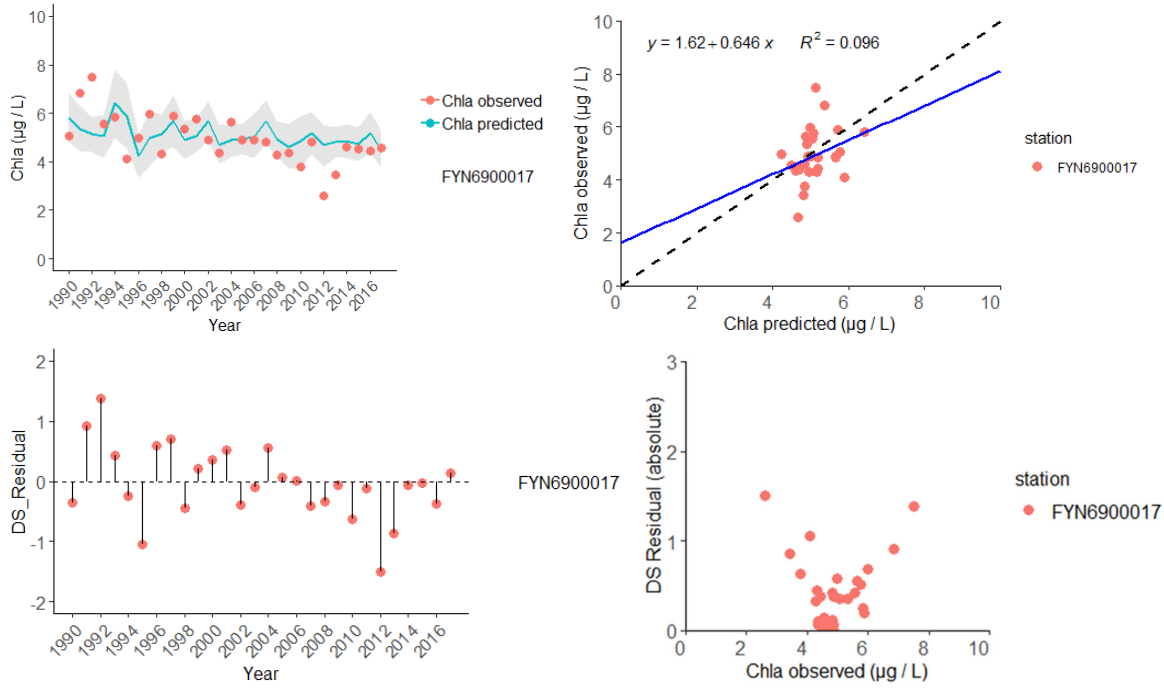

Figure A12.

Odense indre, St. FYN6910008

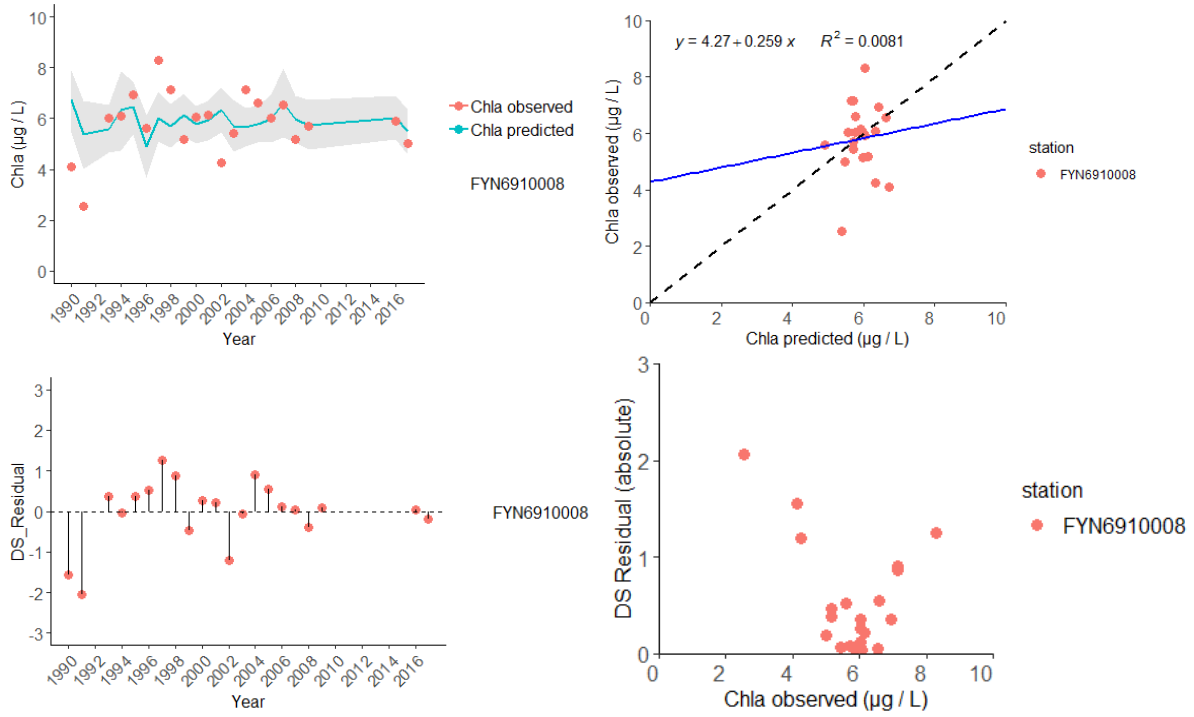

Figure A13.

Vadehavet Grådyb - Ho Bugt v Langli, St. RIB1610002

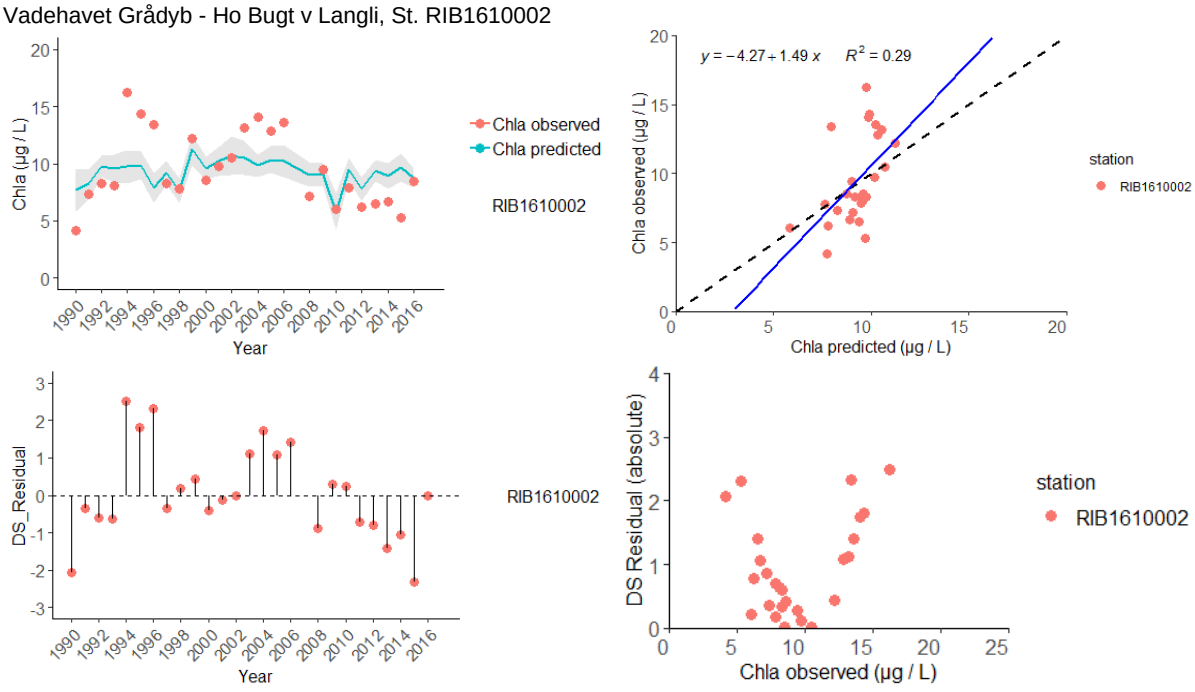

Figure A14.

Vadehavet Knudedyb, St. RIB1620014

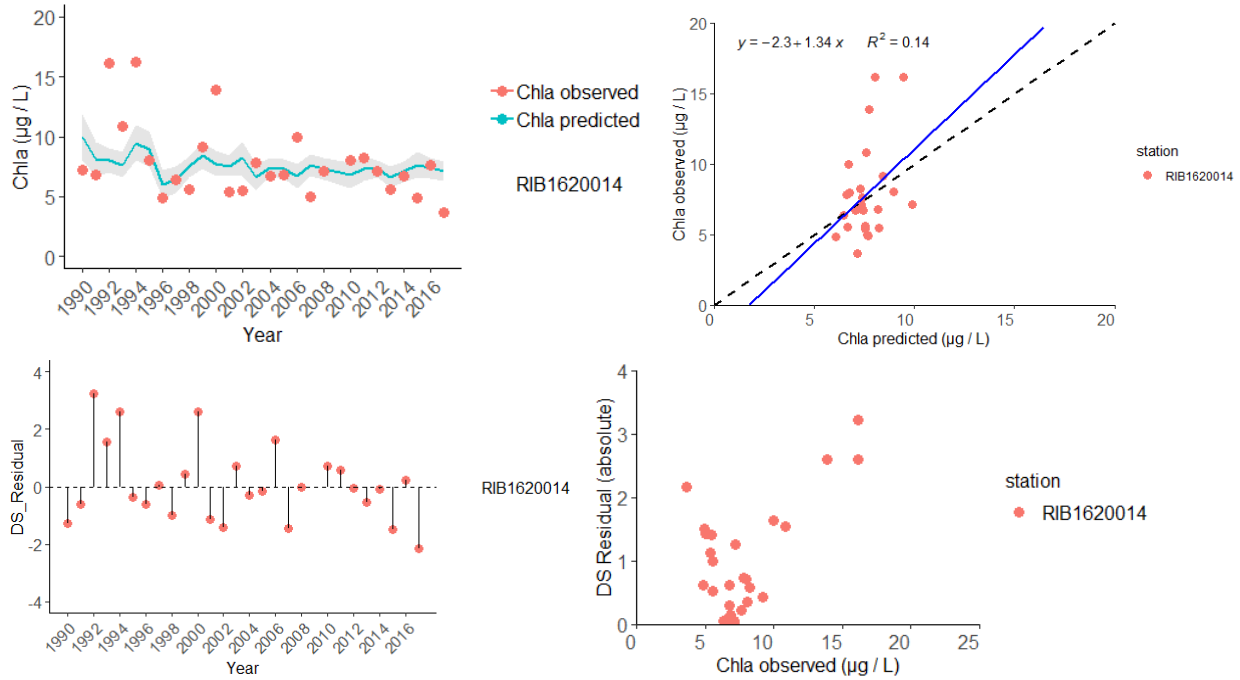

Figure A15.

Ringkøbing Fjord nord, St. RKB1

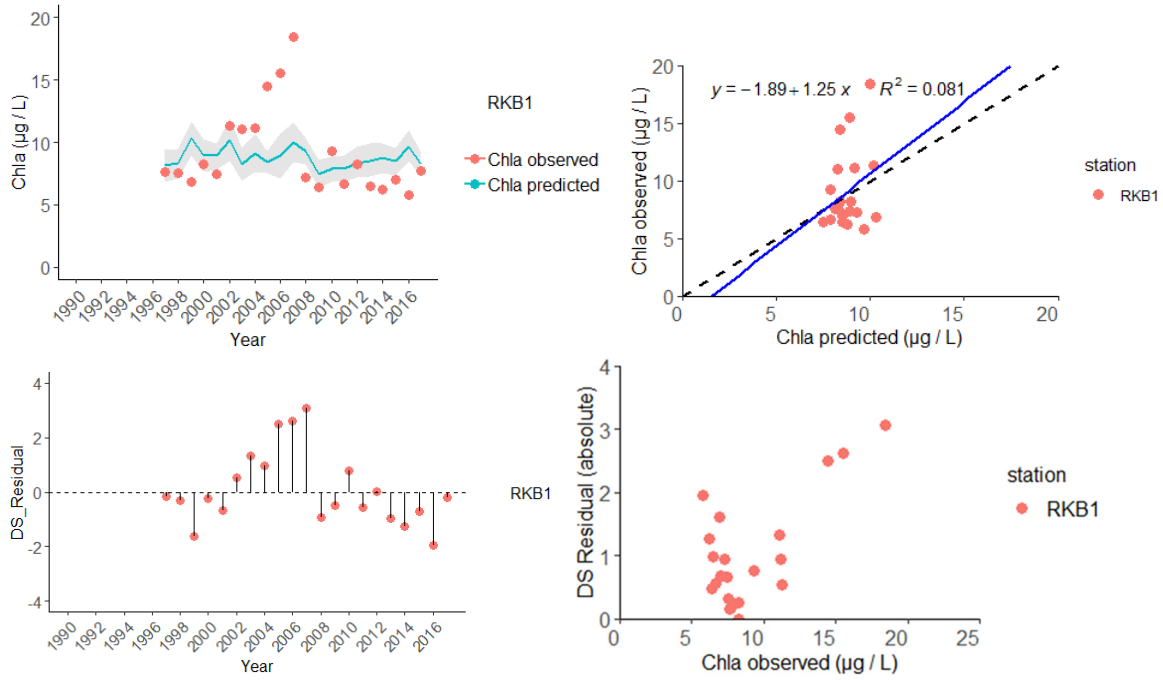

Figure A16.

Køge Bugt, St. ROS1727

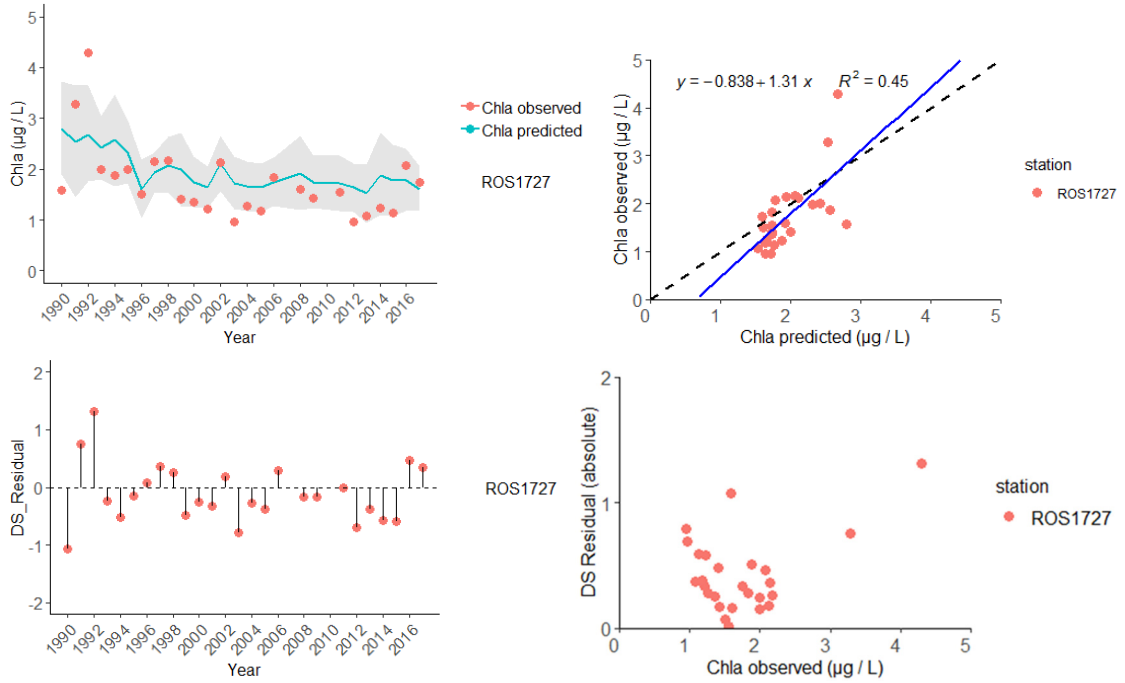

Figure A17.

Roskilde indre, St. ROS60

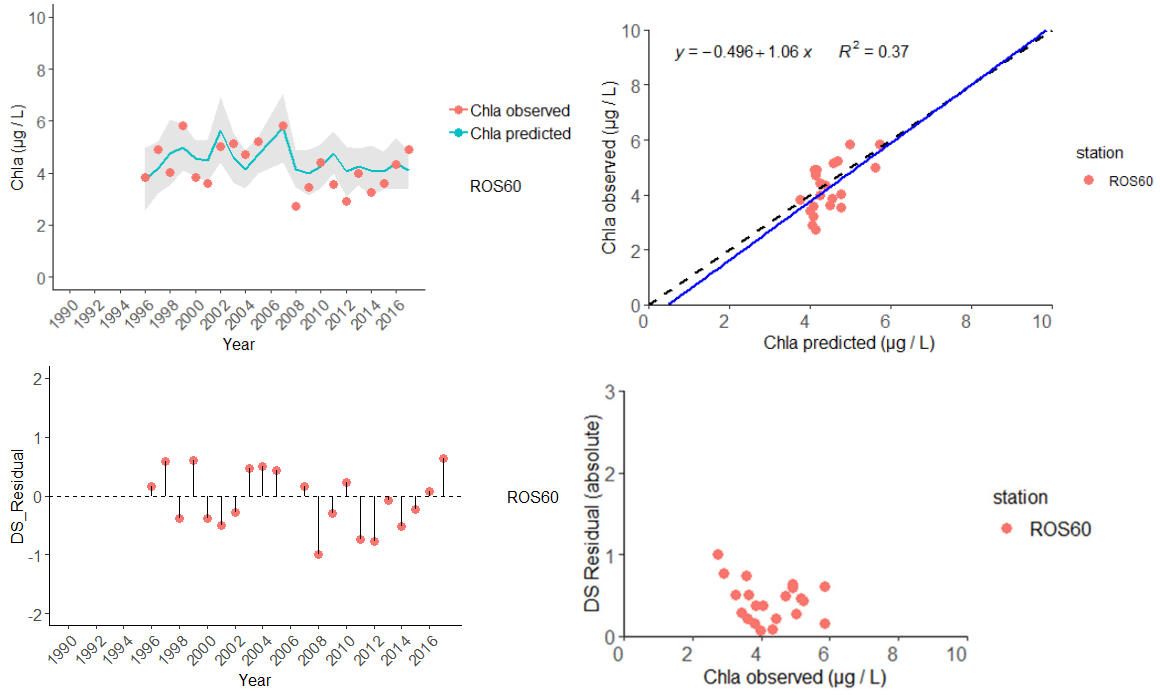

Figure A18.

Lister Dyb St. SJY1

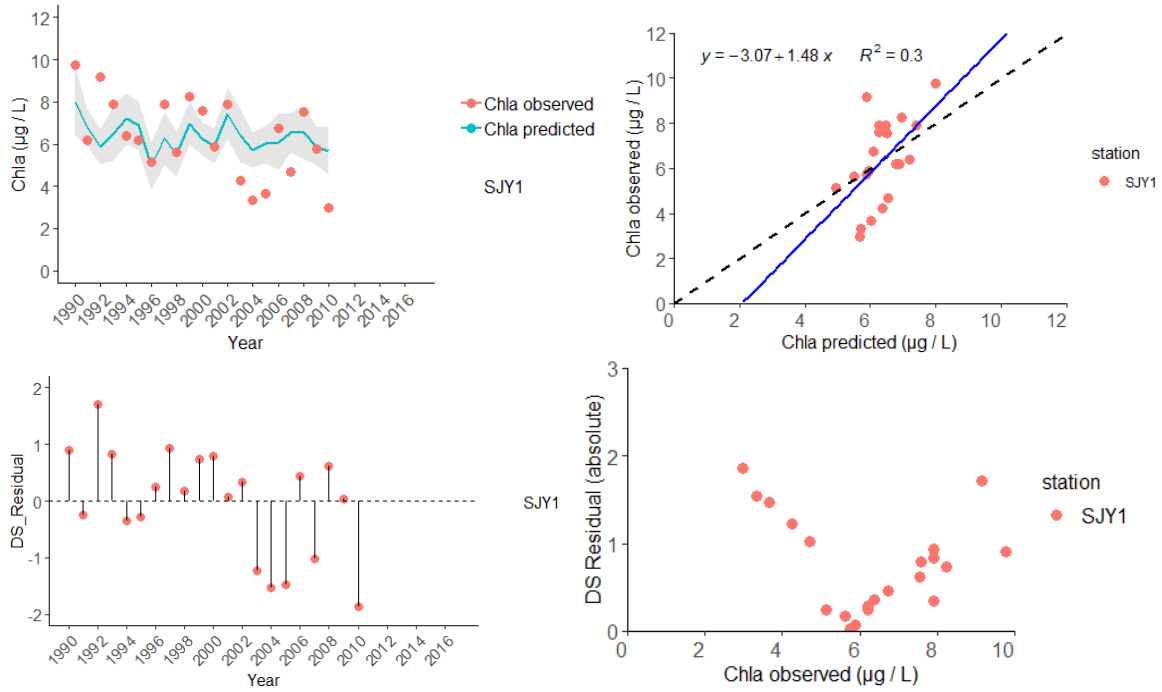

Figure A19.

Augustenborg Fjord, St. SJY12

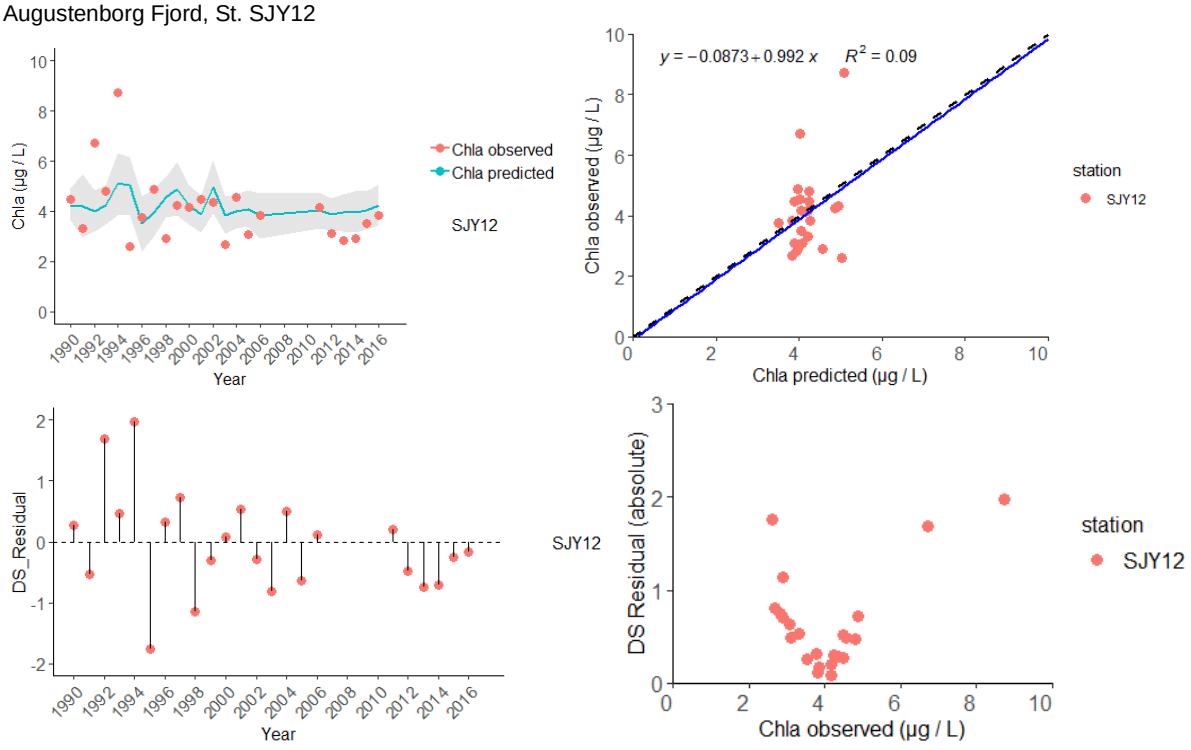

Figure A20.

Als Fjord, St. SJY13B

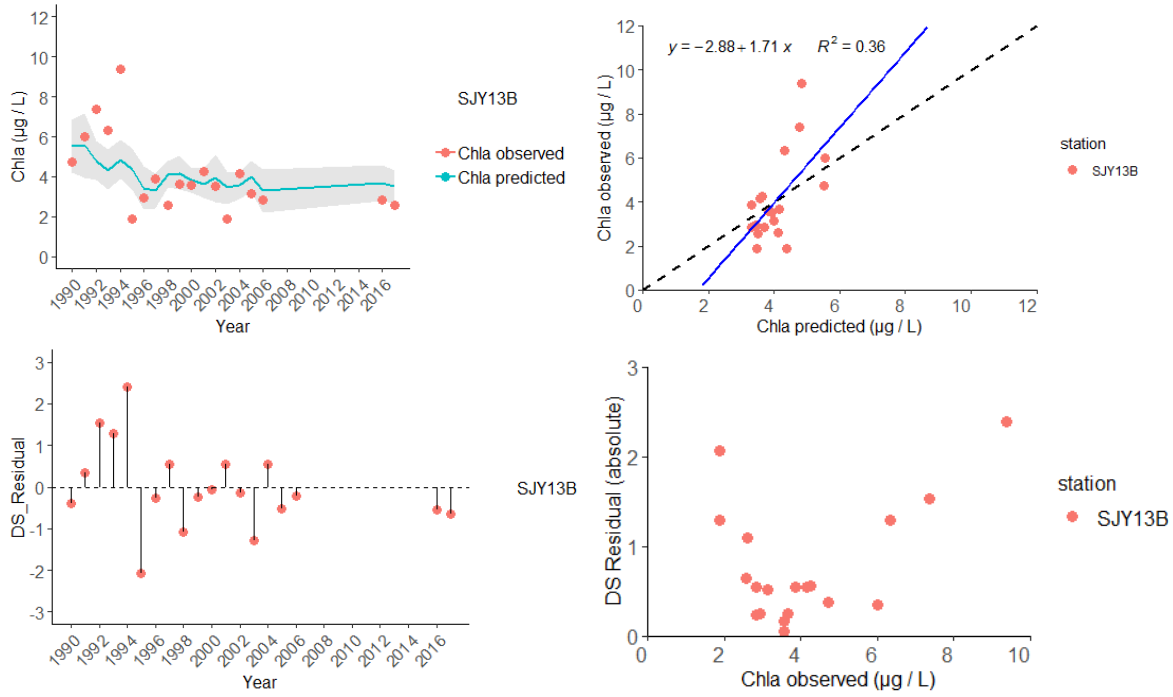

Figure A21.

Aabenraa Fjord, St. SJY15

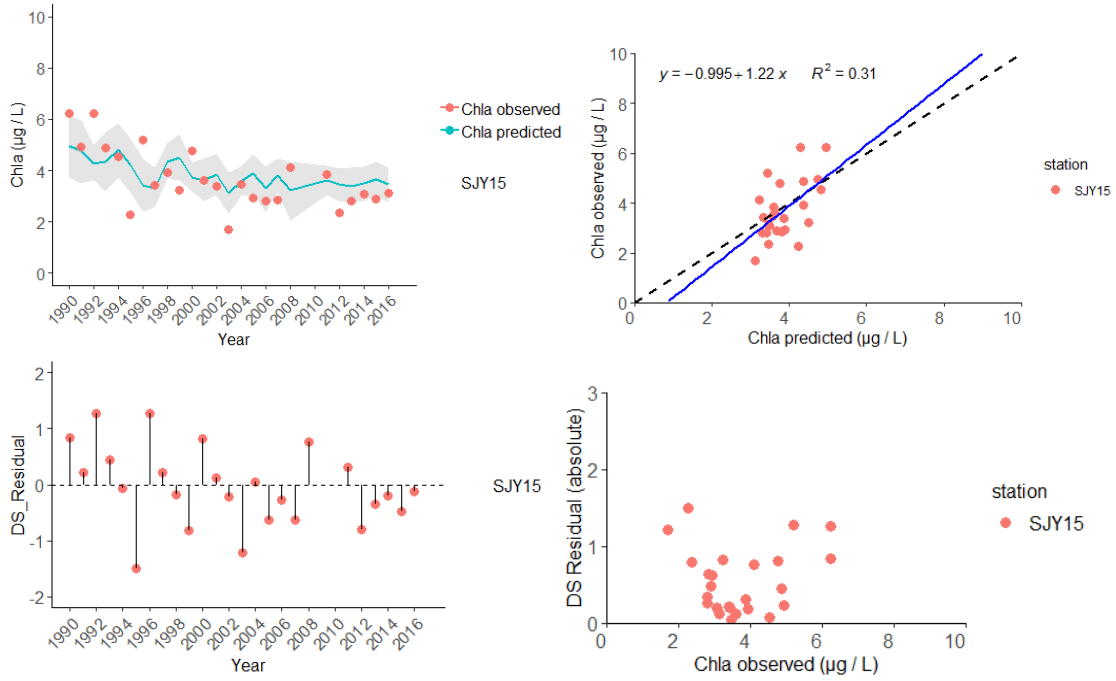

Figure A22.

Genner Bugt, St. SJY19

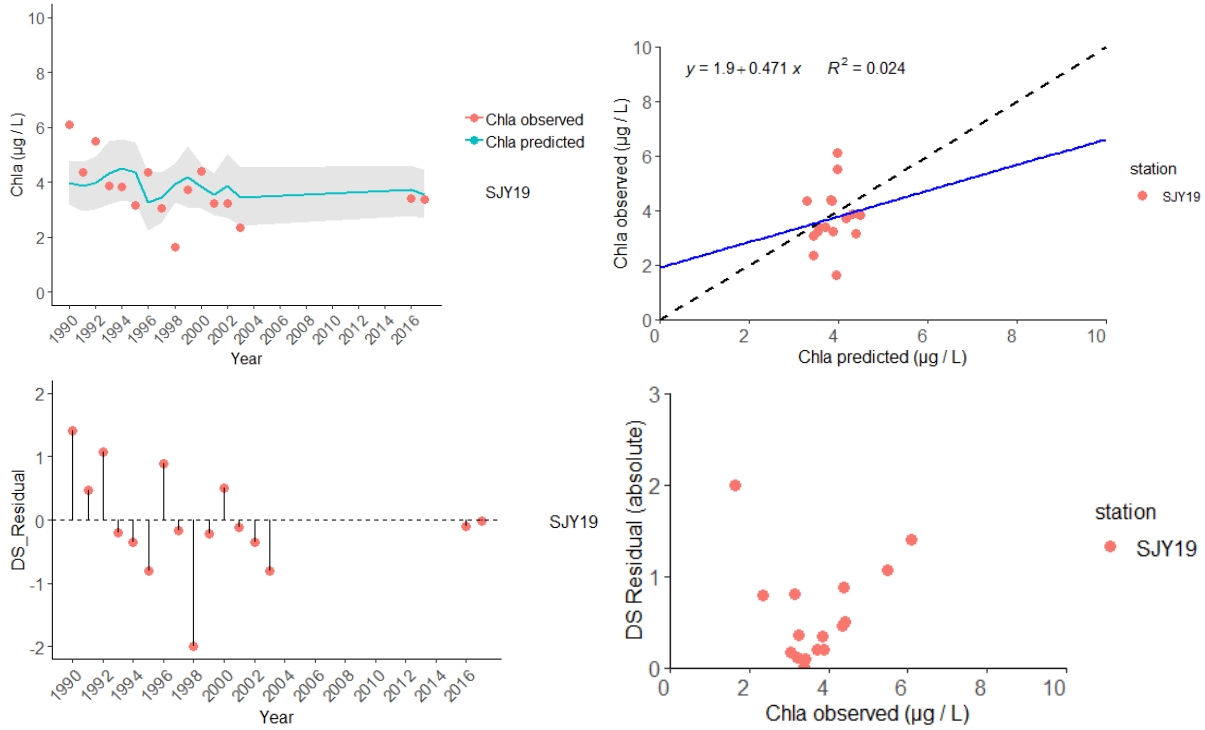

Figure A23.

Lister Dyb v Rømø Havneby, St. SJY3

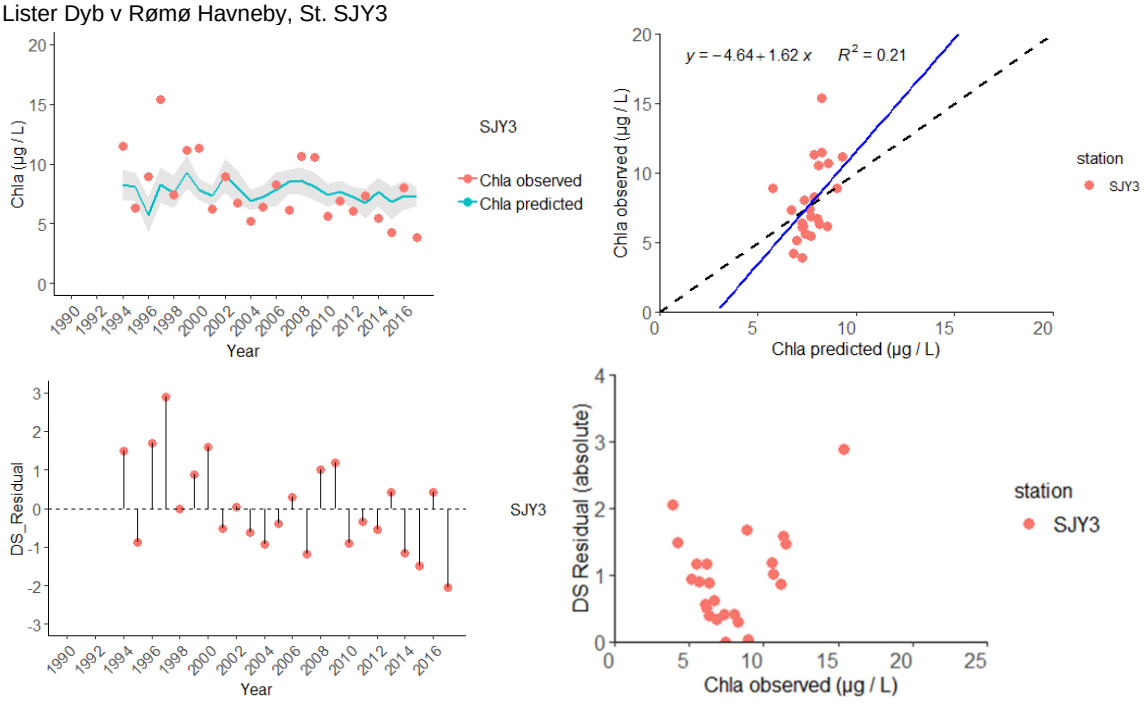

Figure A24.

Flensborg Fjord inder, St. SJYKFF2

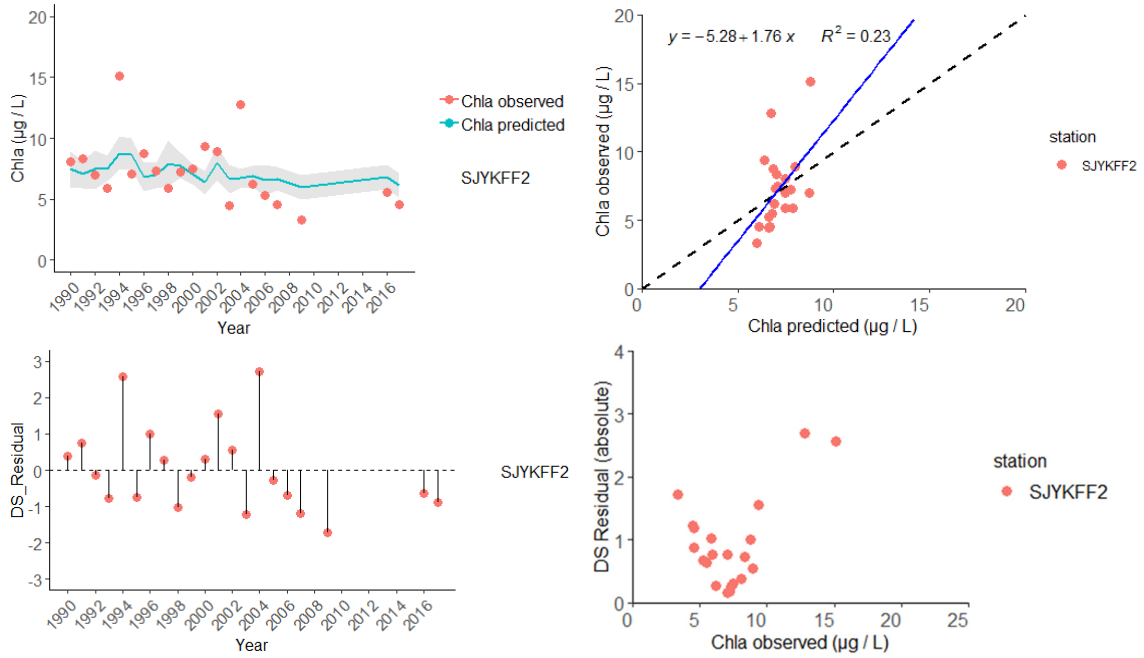

Figure A25.

Flensborg Fjord yder, St. SJYKFF5

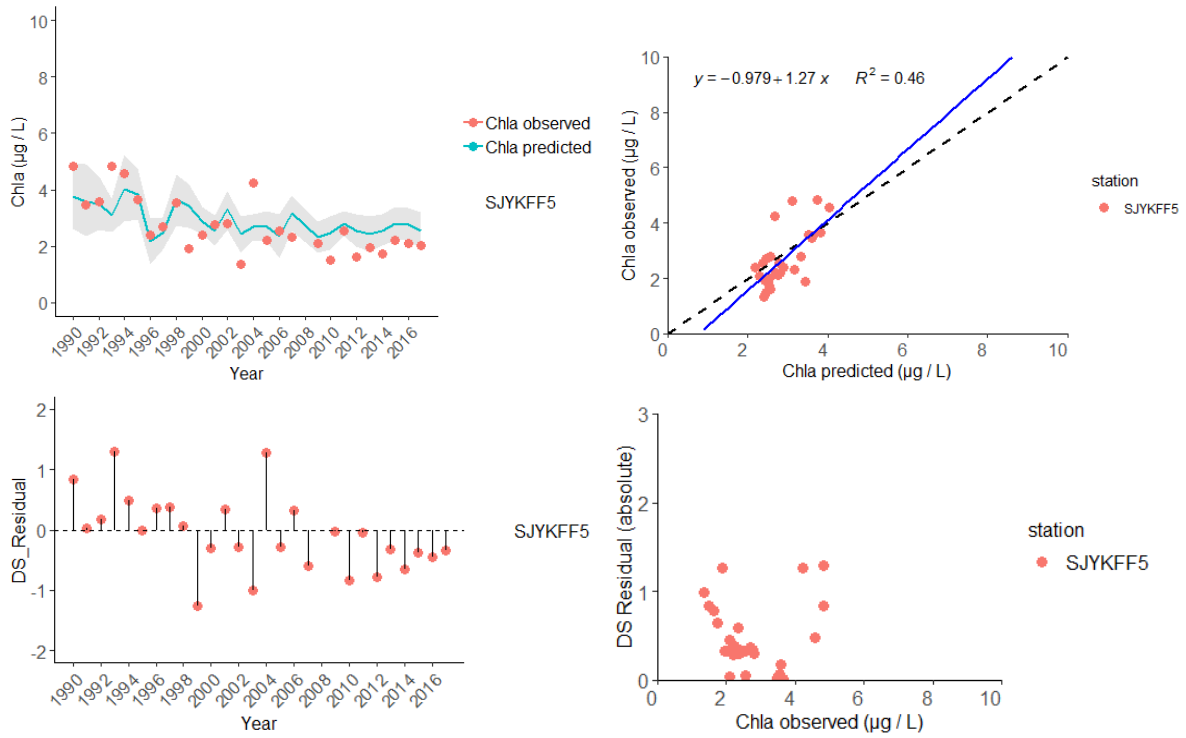

Figure A26.

Præstø Fjord, St. STO0802008

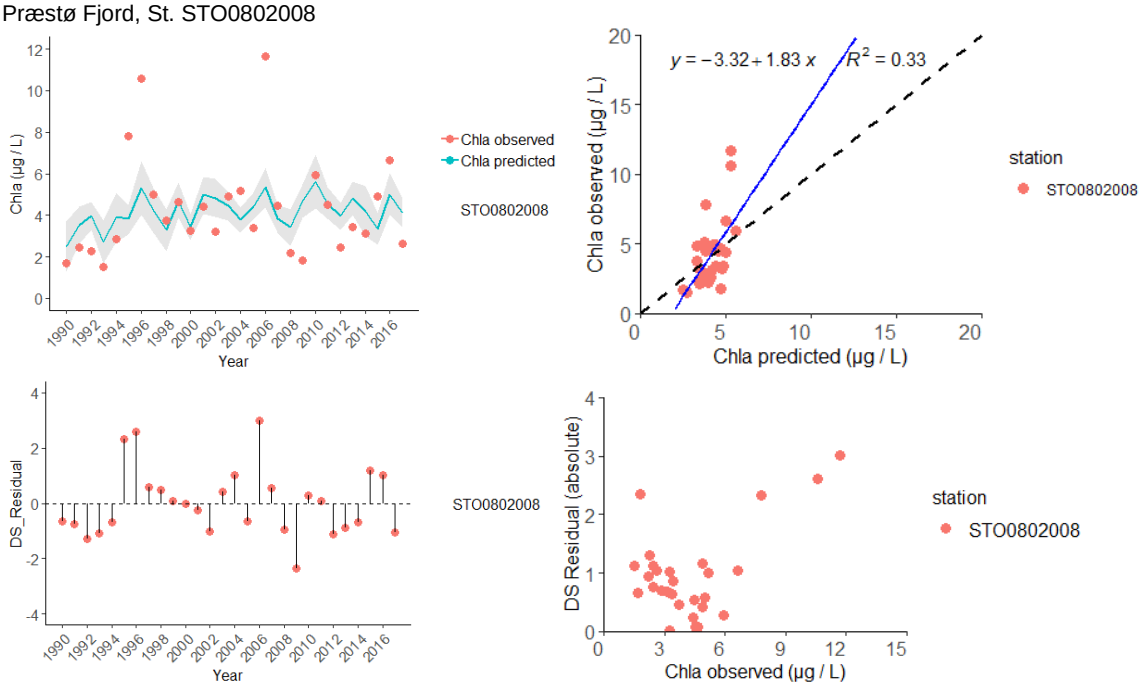

Figure A27.

Kolding Fjord, St. VEJ0003350

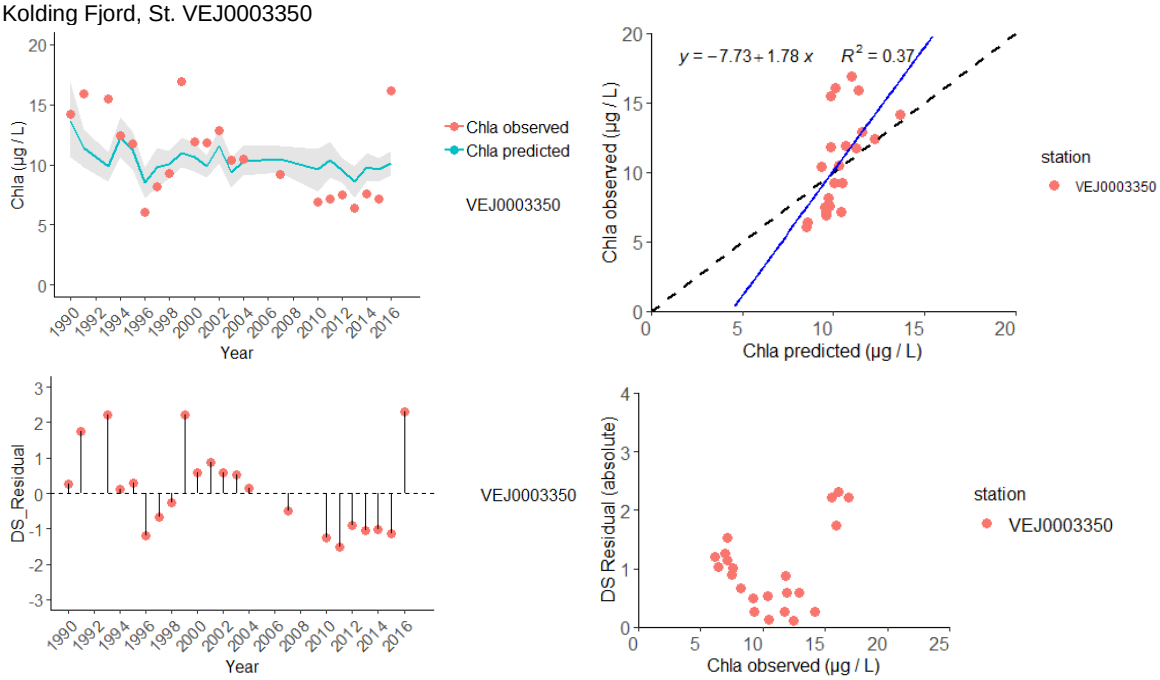

Figure A28.

Vejle Fjord, St. VEJ0004273

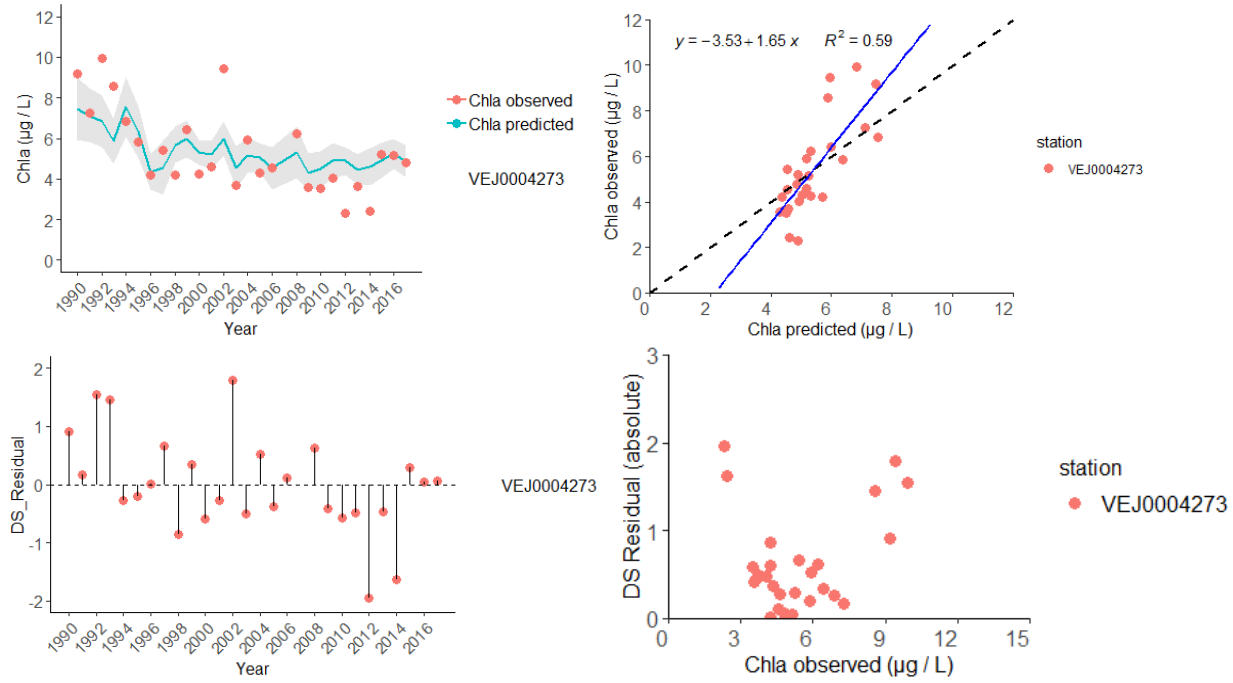

Figure A29.

Horsens inder, St. VEJ0005790

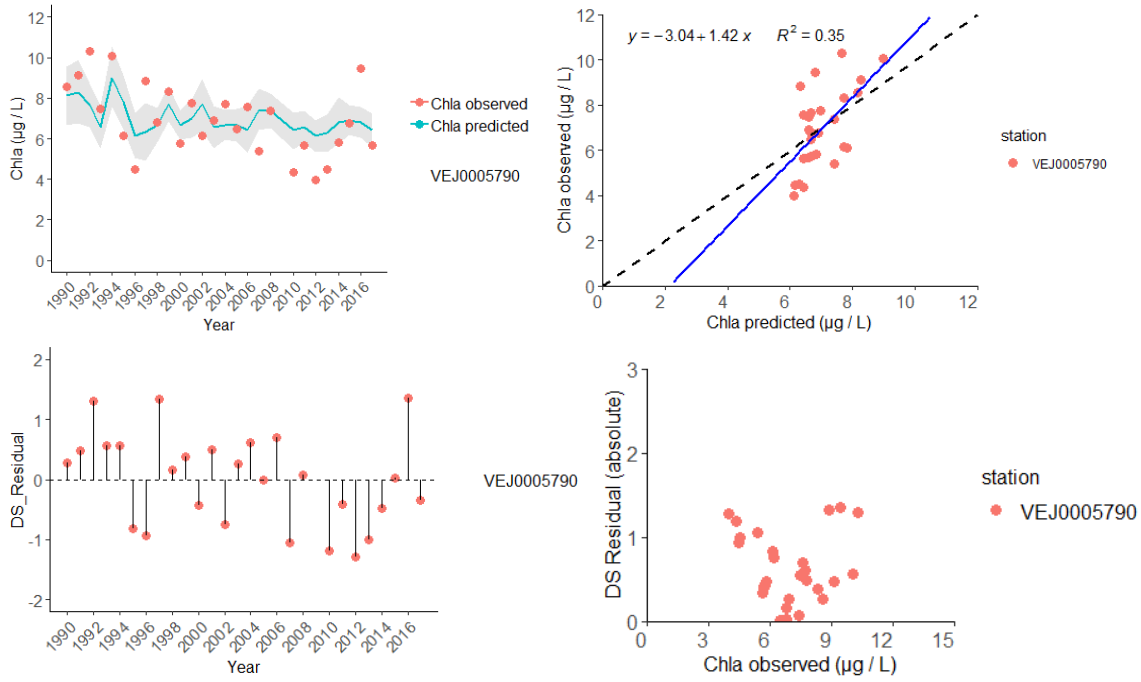

Figure A30.

Horsens yder, St. VEJ0006489

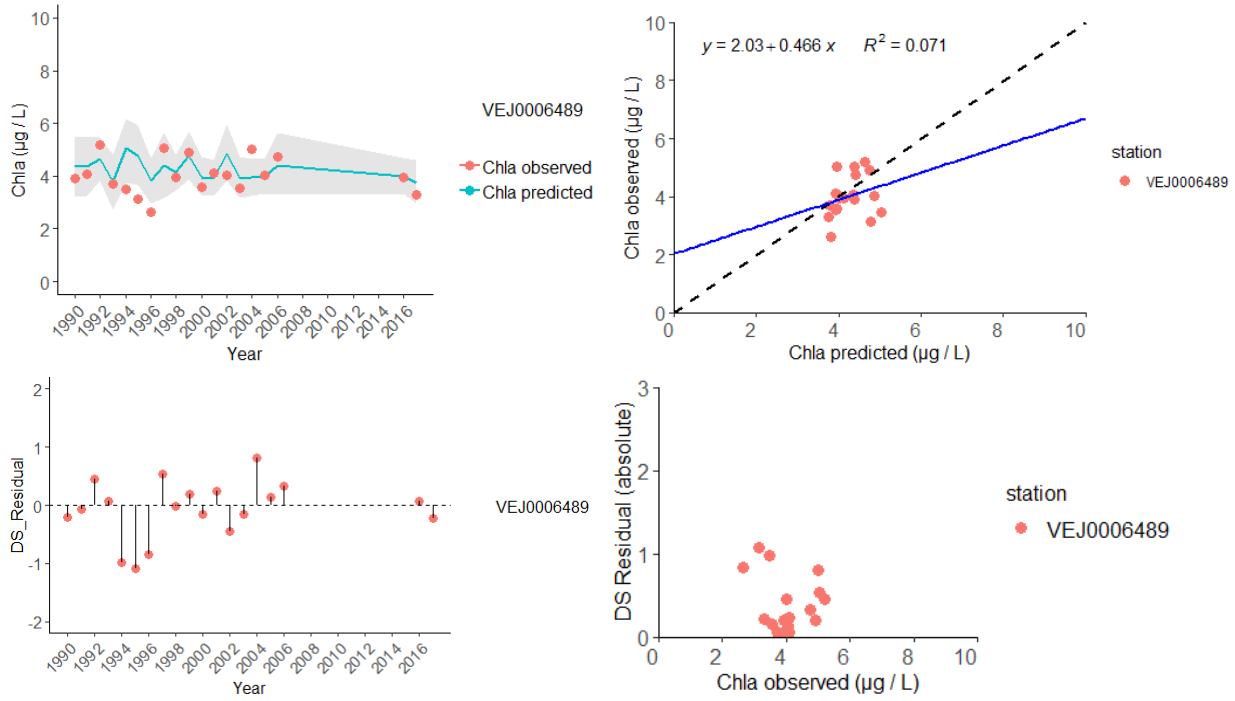

Figure A31.

Nissum Bredning, St. VIB3702-00001

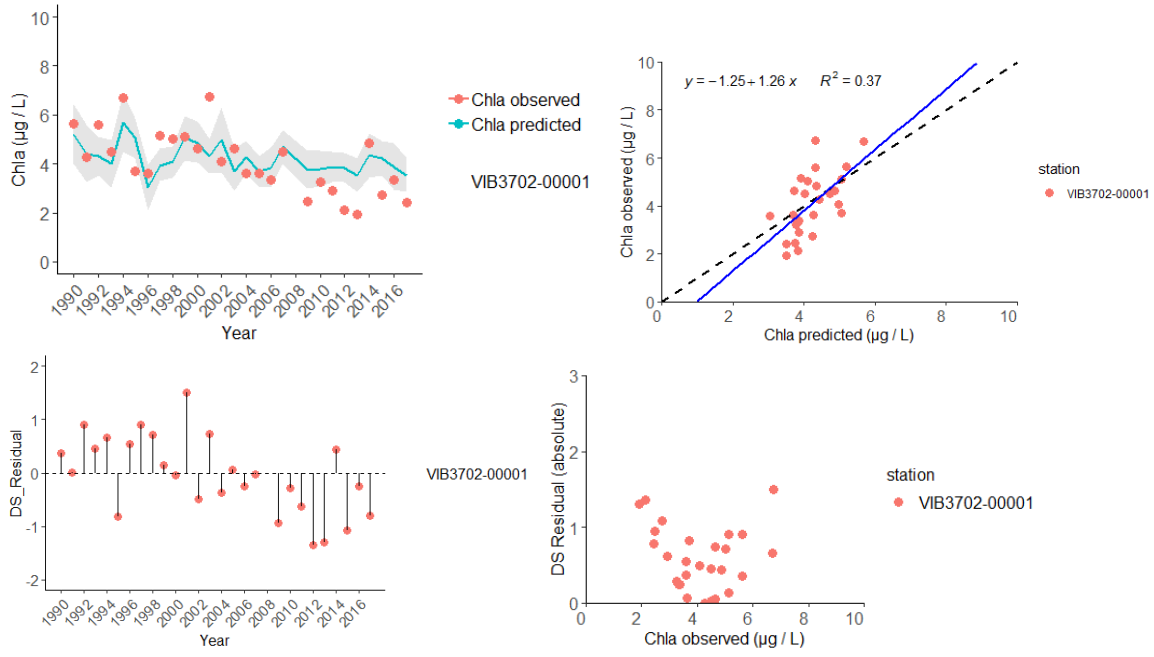

Figure A32.

Løgstør bredning, St. VIB3708-00001

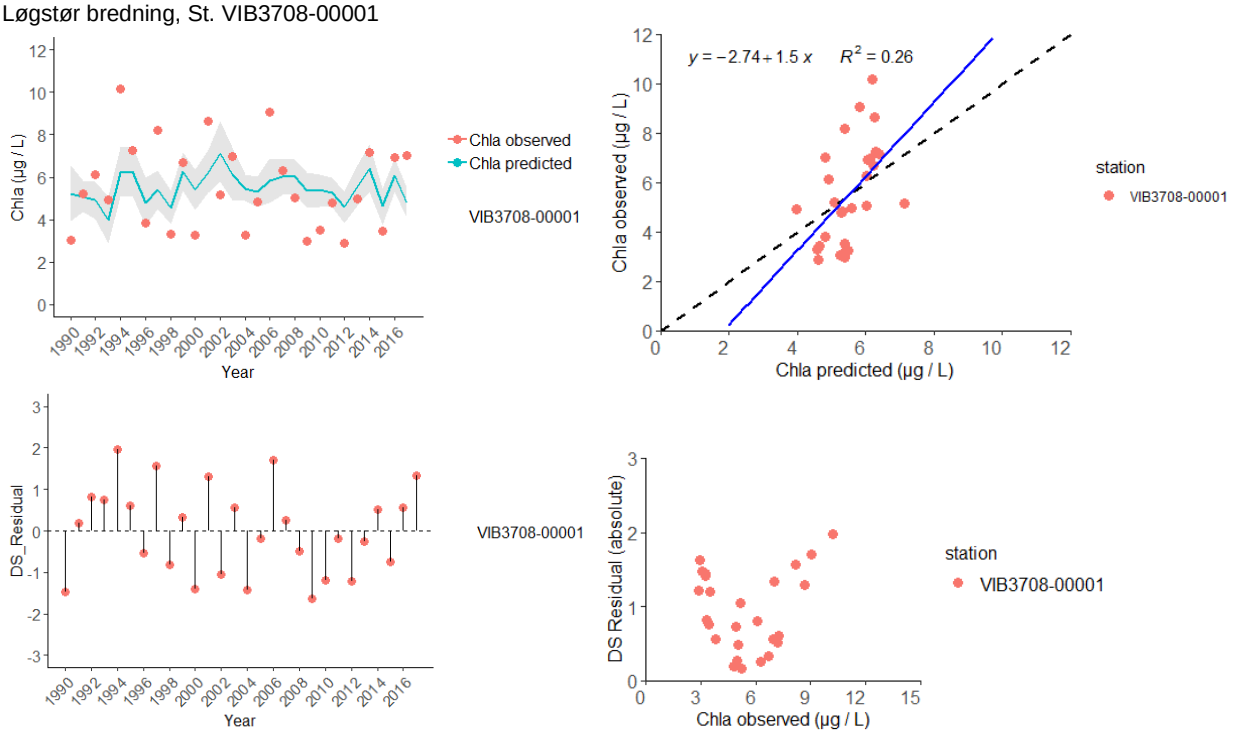

Figure A33.

Nibe Bredning, St. VIB3711-00001

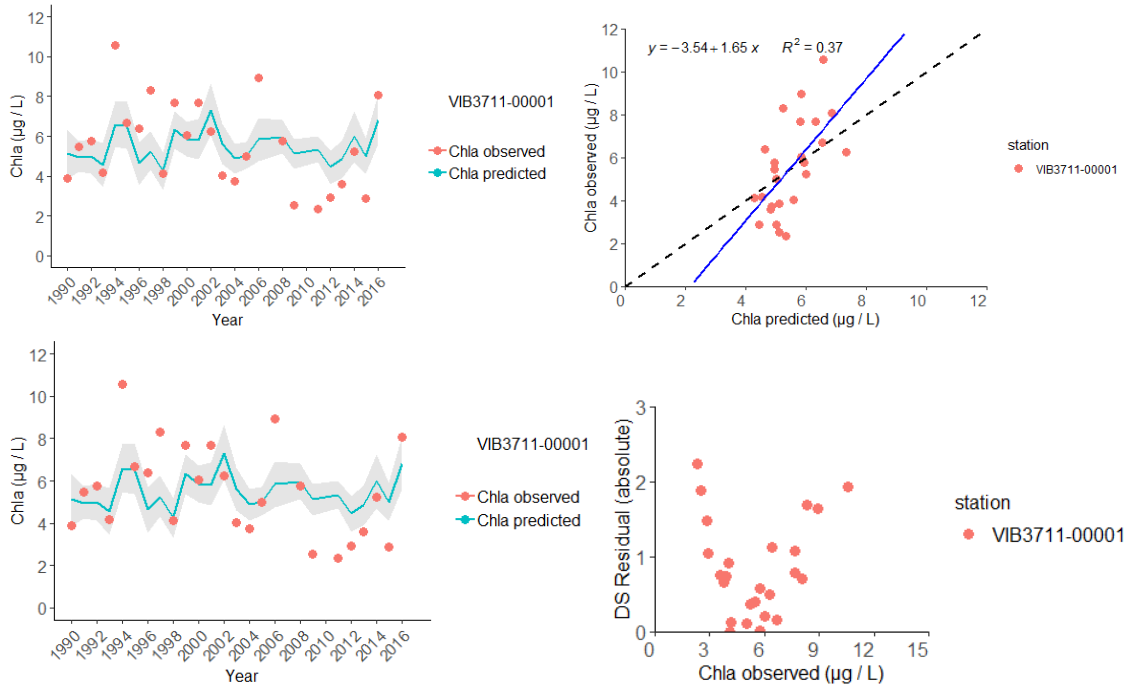

Figure A34.

Thisted Bredning, St. VIB3723-00001

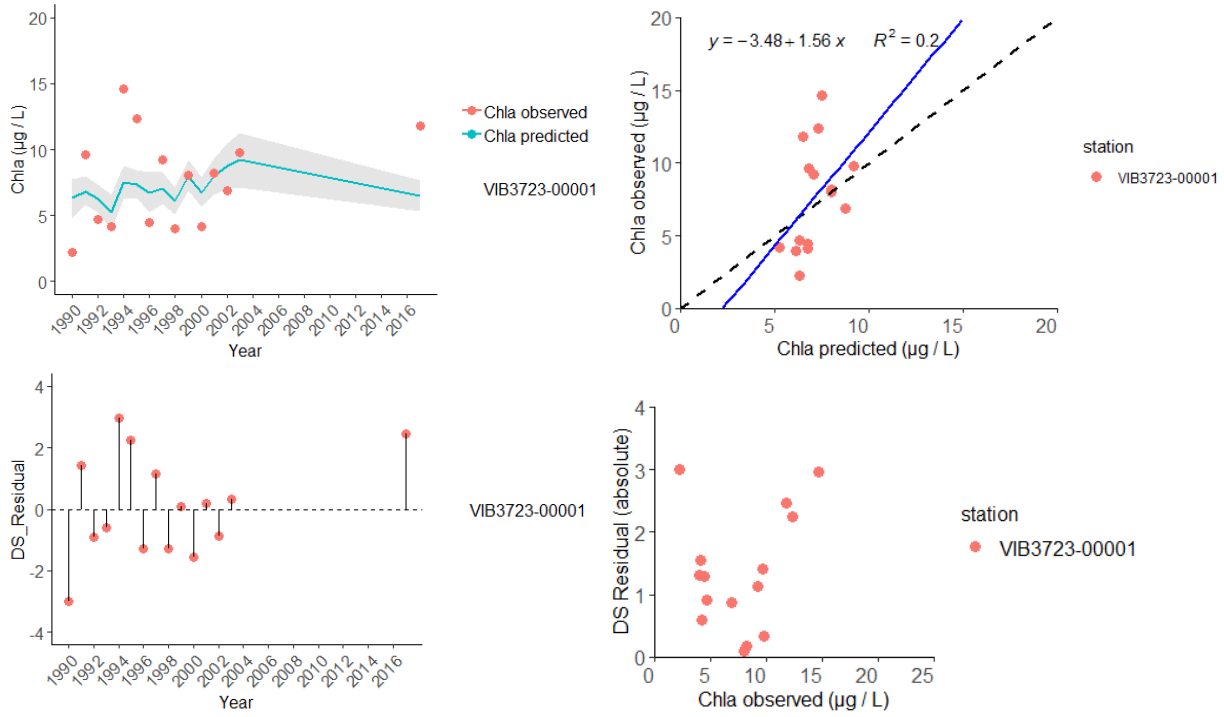

Figure A35.

Skive Fjord, St. VIB3727-00001

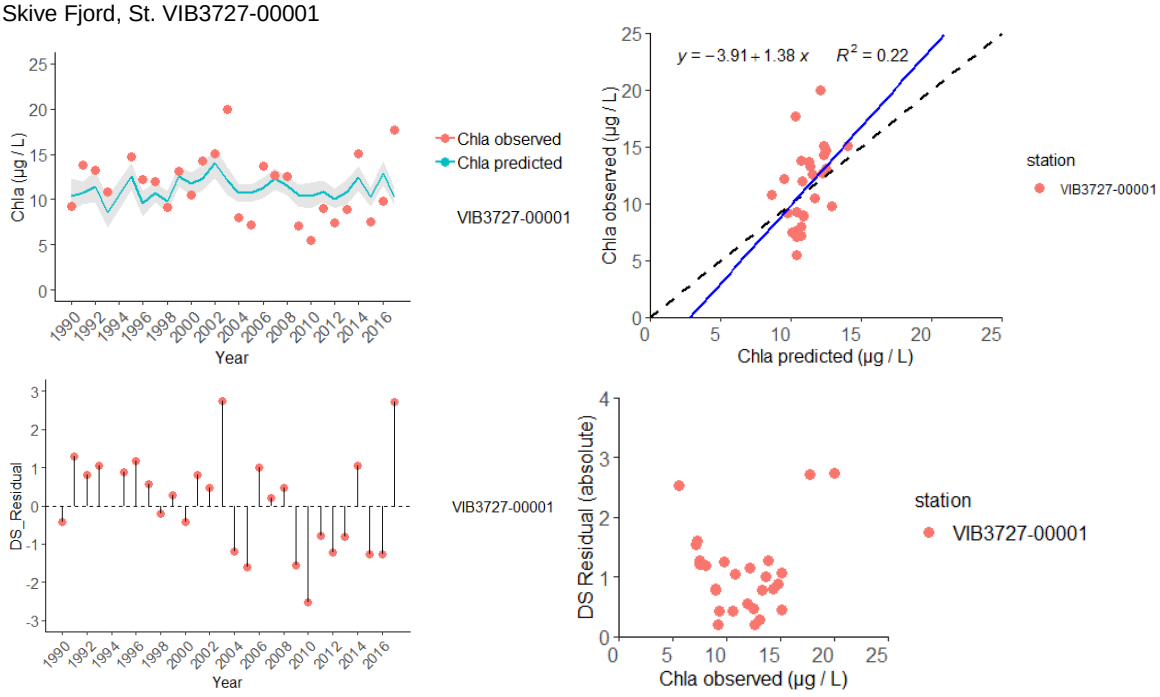

Figure A36.

Lovns Bredning, St. VIB3728-00001

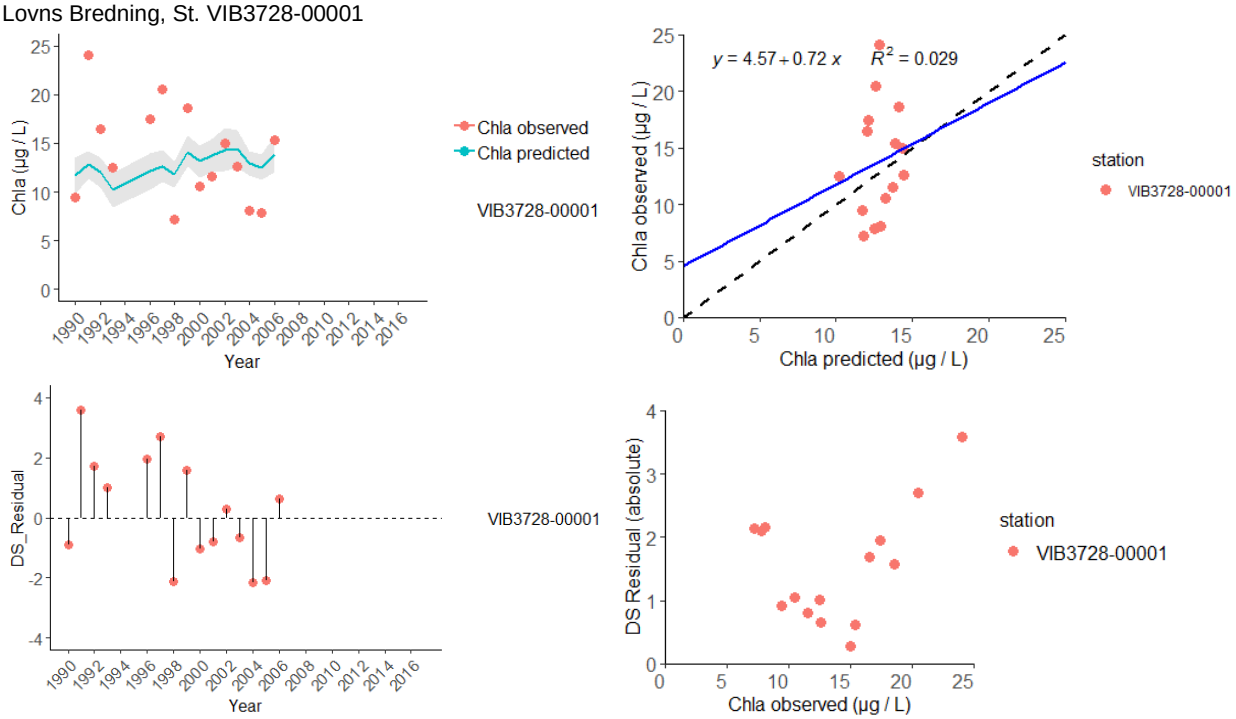

Figure A37.

Isefjord dybt basin, St. VSJ10003

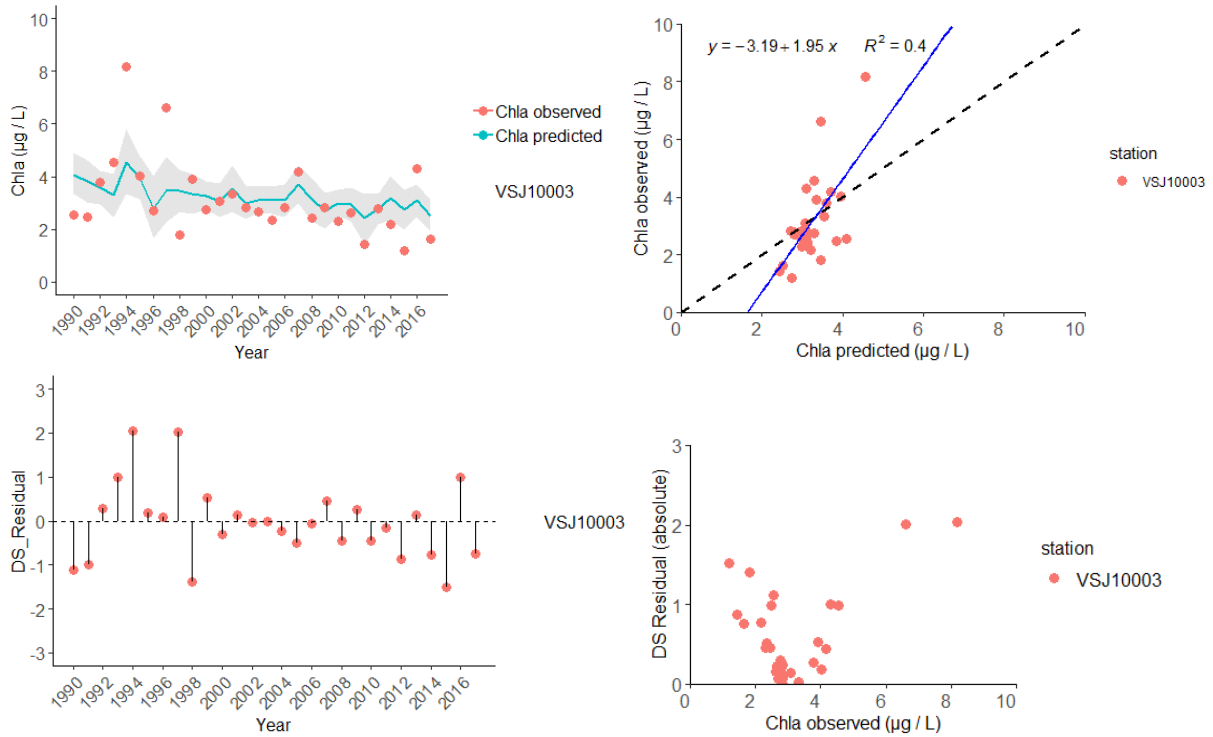

Figure A38.

Isefjord inderbredning, St. VSJ10006

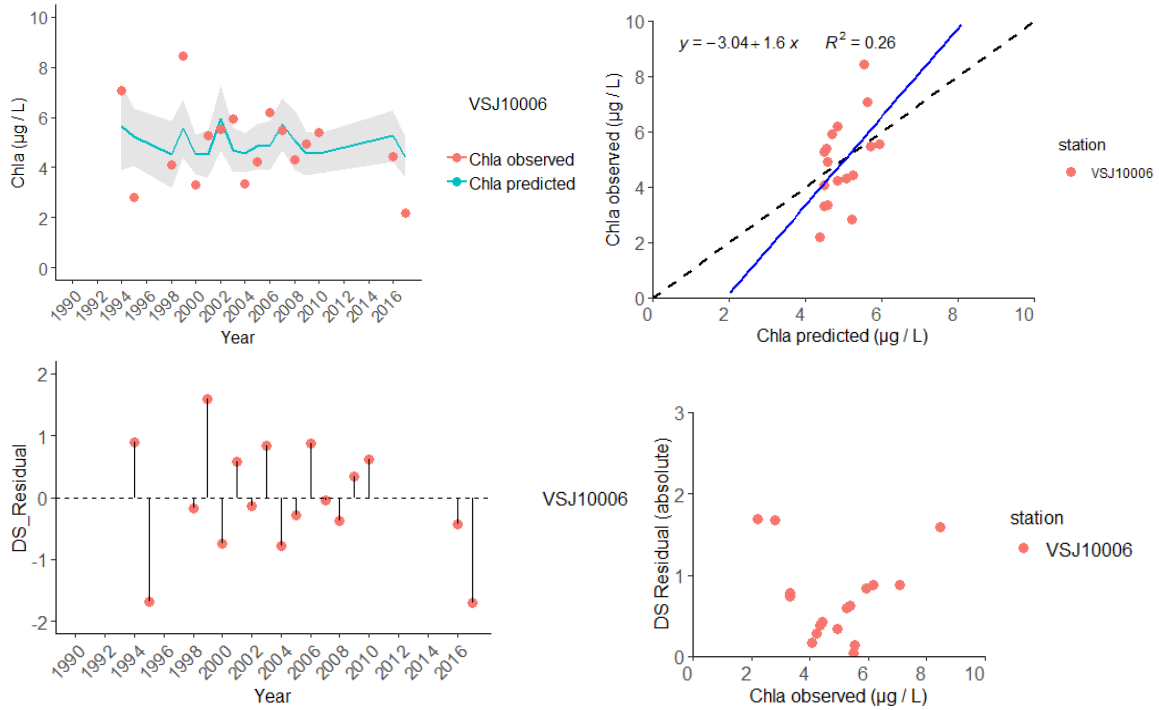

Figure A39.

Kalundborg Fjord yder, St. VSJ41007

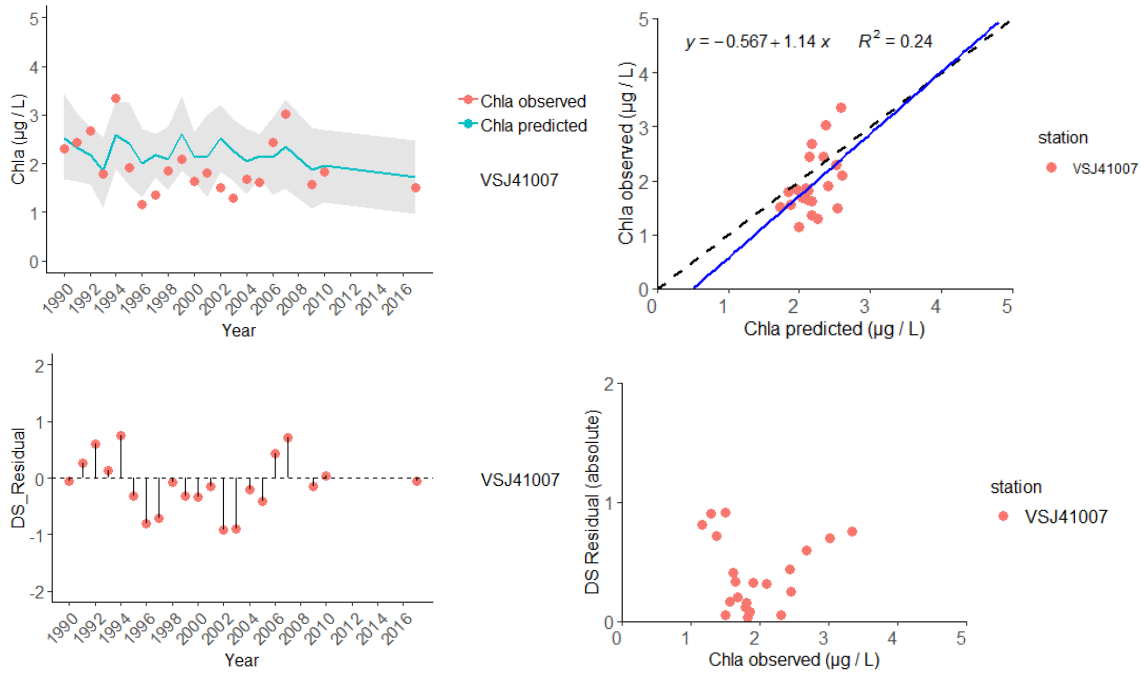

Figure A40.

Kalundborg Fjord inder, St. VSJ41008

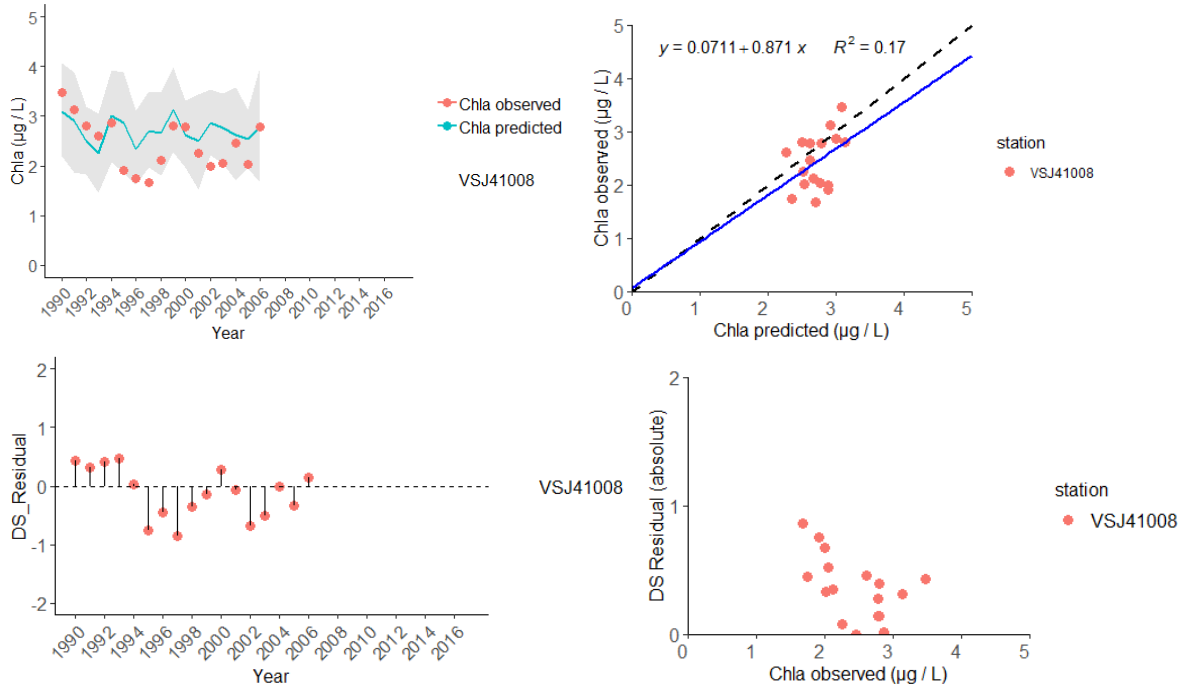

Figure A41.

Skælskør Fjord, St. VSJ51013

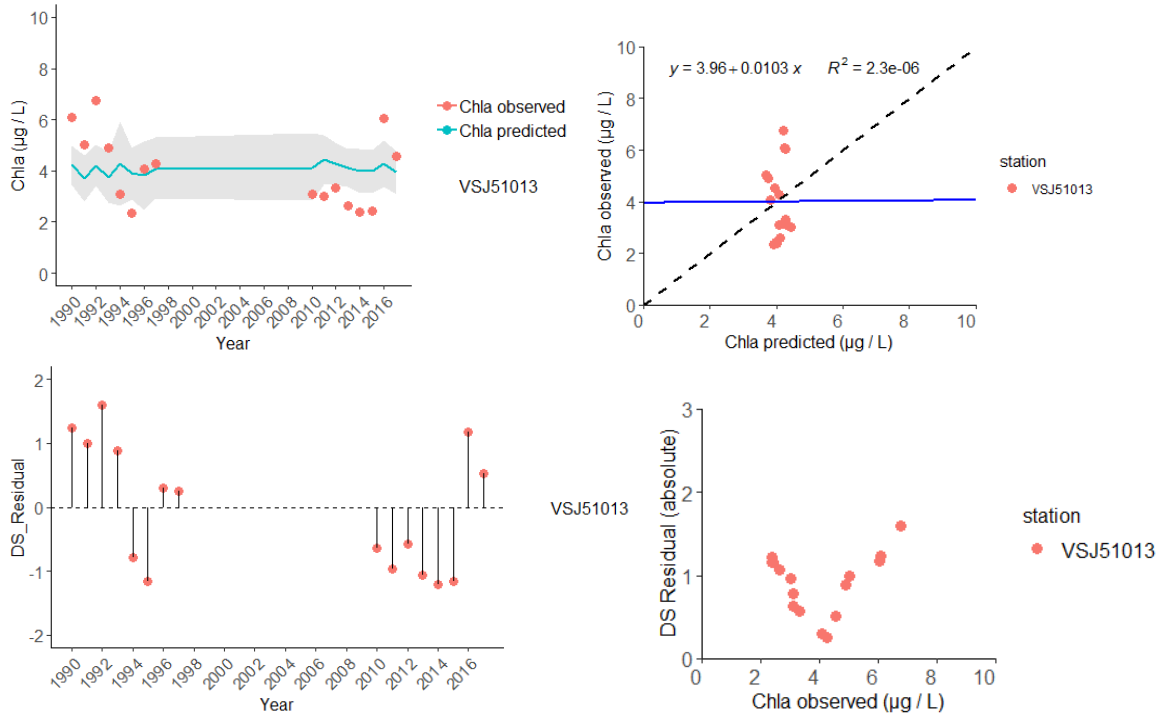

Figure A42.

Appendix B: Bayesian model evaluation for station-specific model

Hevring Bugt, St. ARH190004

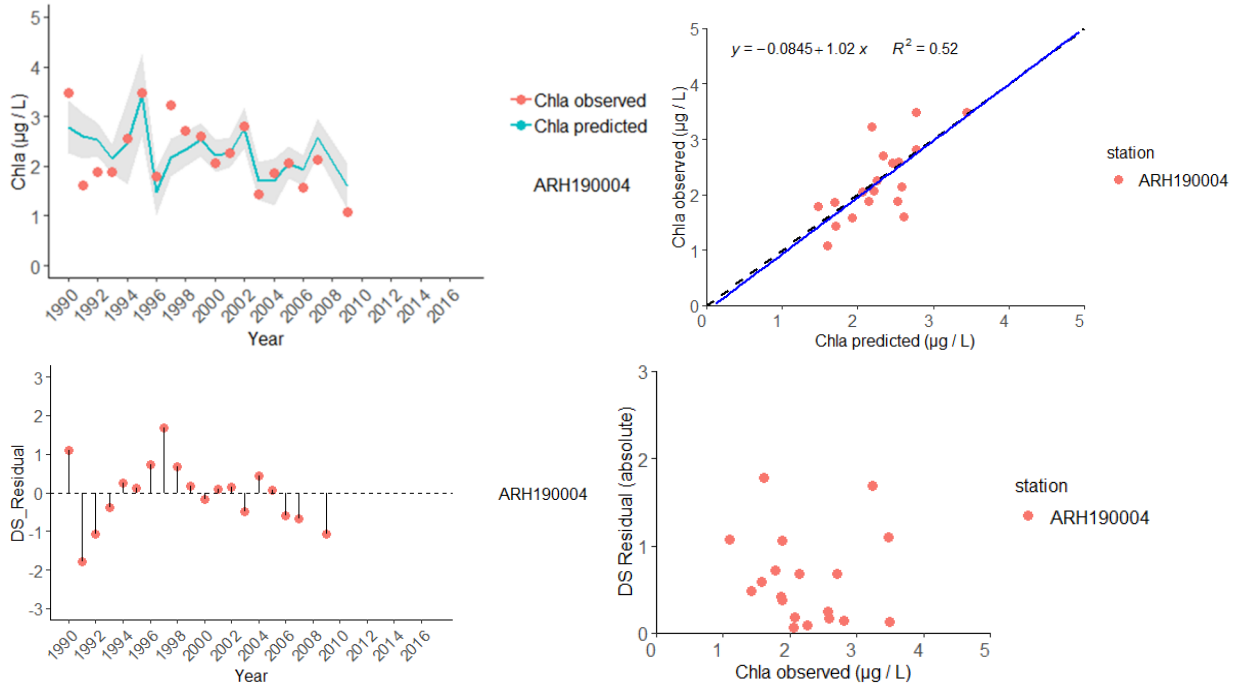

Figure B1.

Randers ml, St. ARH230902

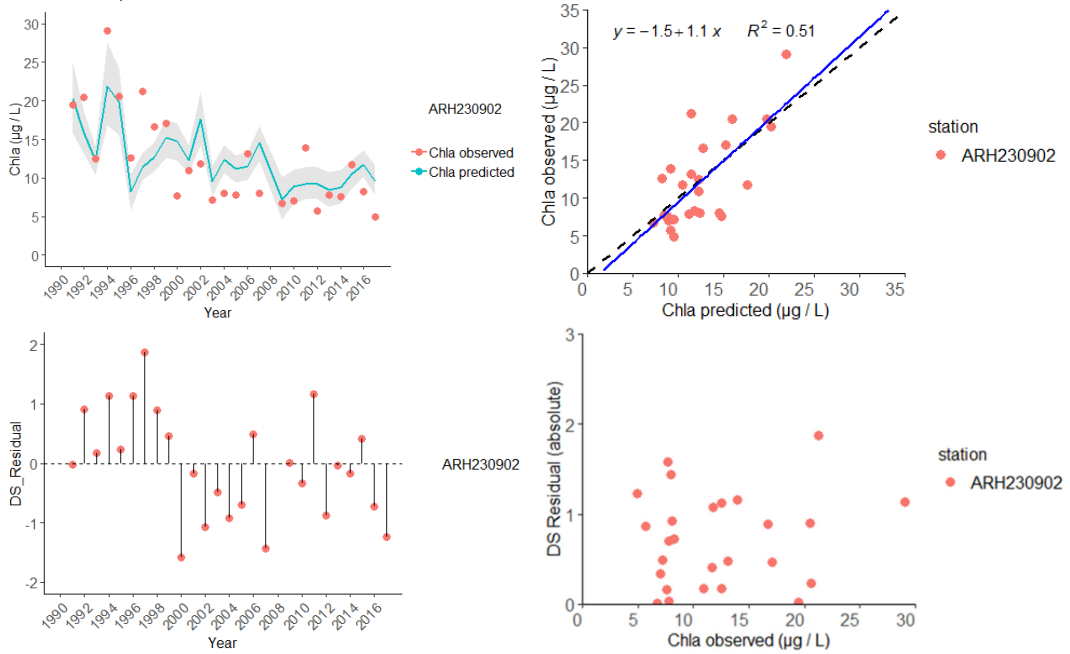

Figure B2.

Randers ydre, St. ARH230905

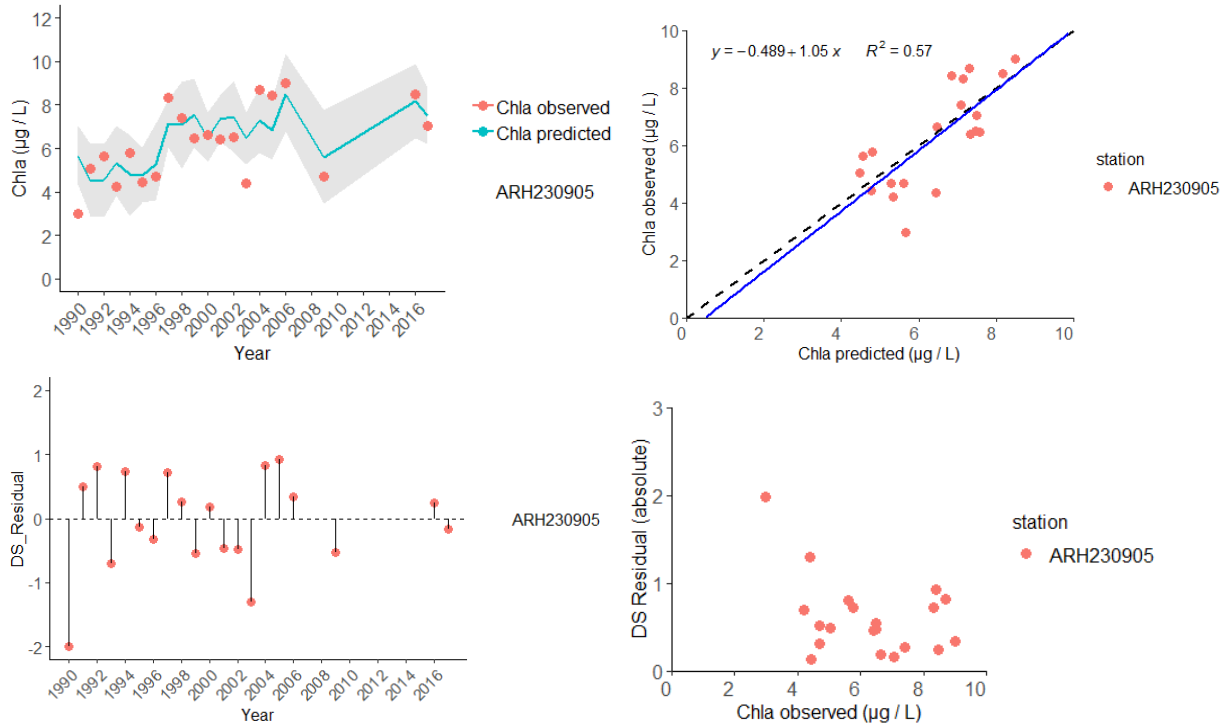

Figure B3.

Kerteminde Fjord/Kertinge Nor, St. FYN0018843

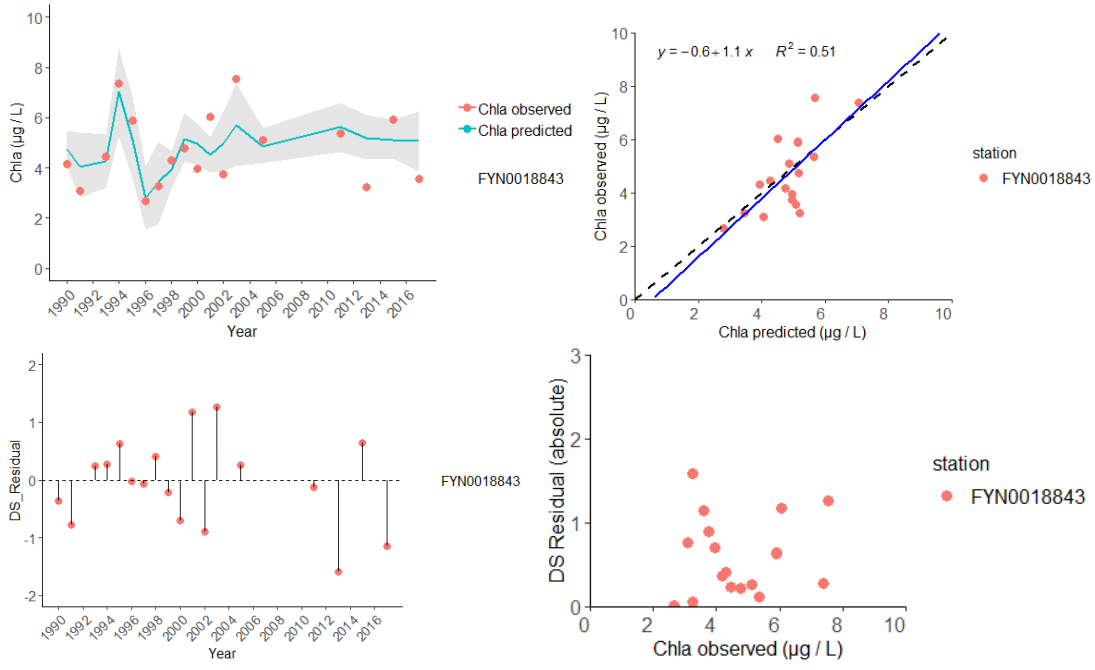

Figure B4.

Lillebælt nord - ved Fredericia, St. FYN6100051

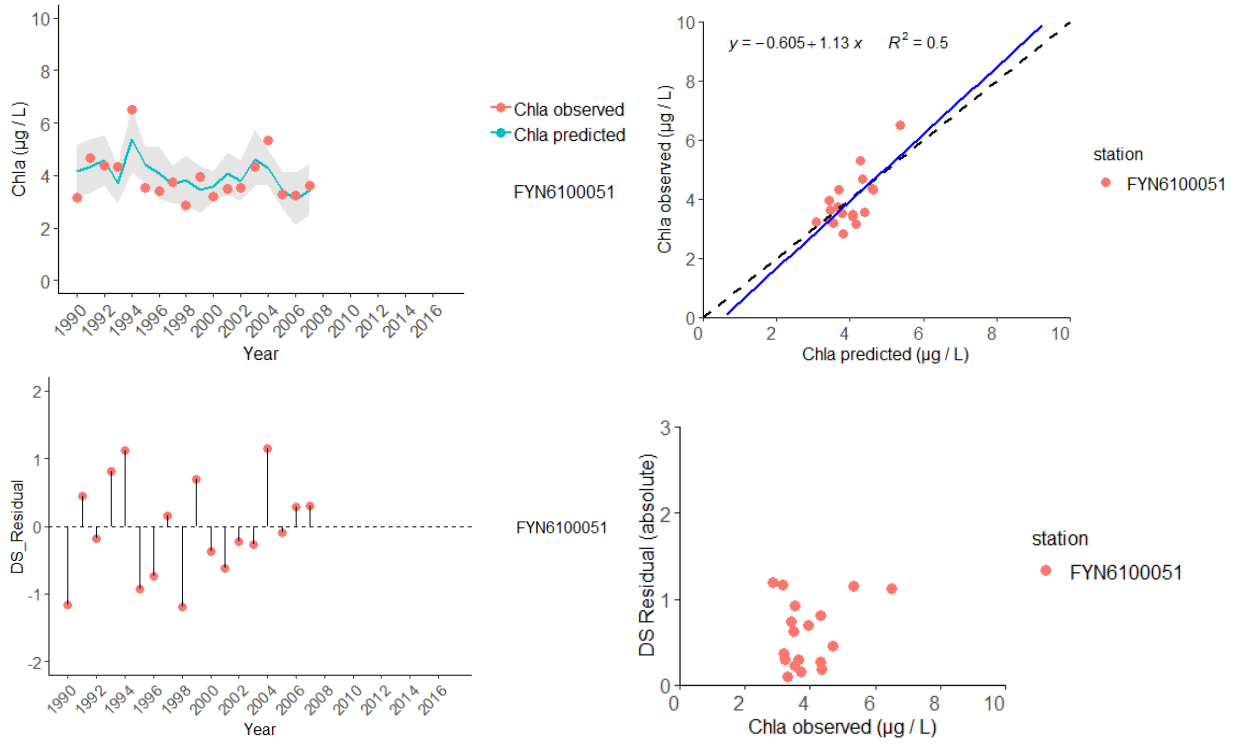

Figure B5.

Nord for Als (Lillebælt vest), St. FYN6200901

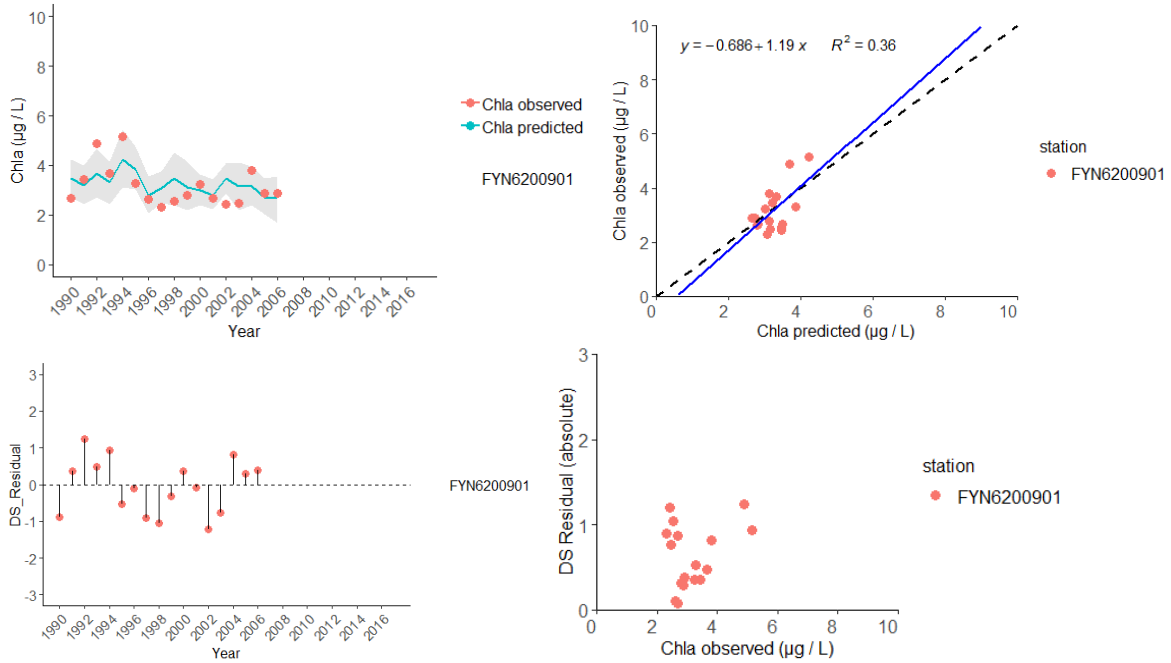

Figure B6.

Odense ydre, St. FYN6900017

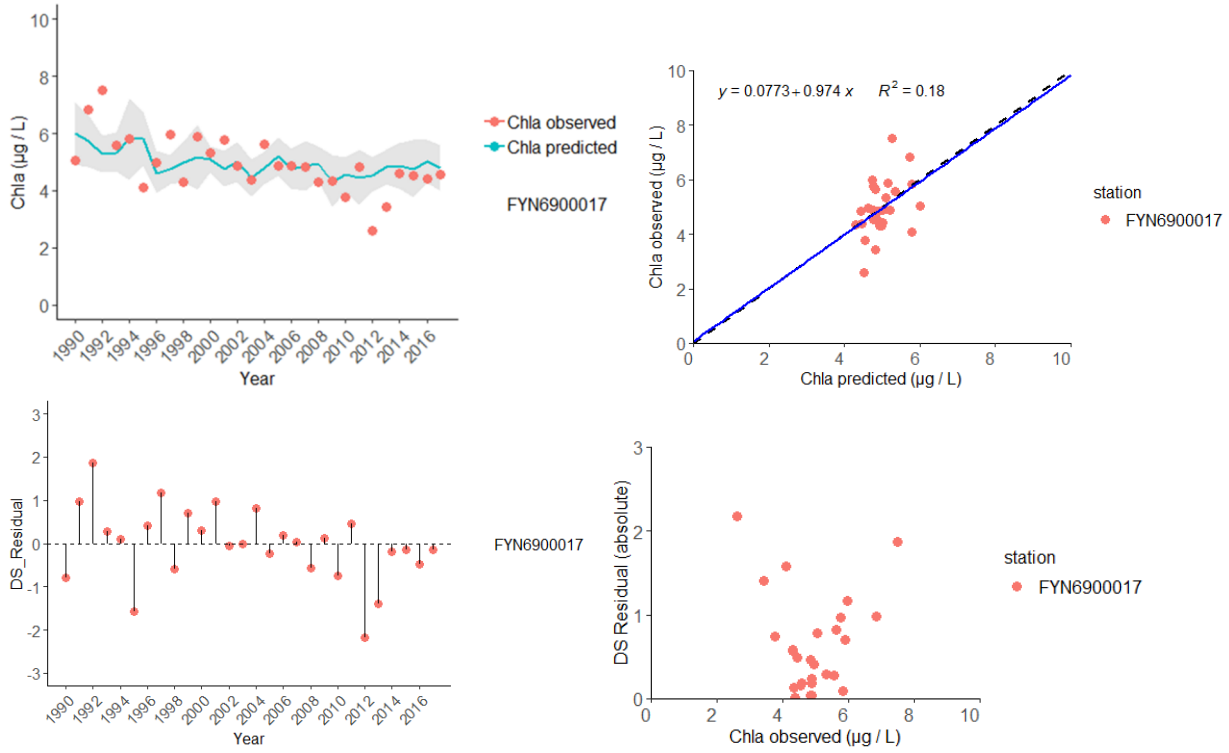

Figure B7.

Mariager Fjord, St. NOR5503

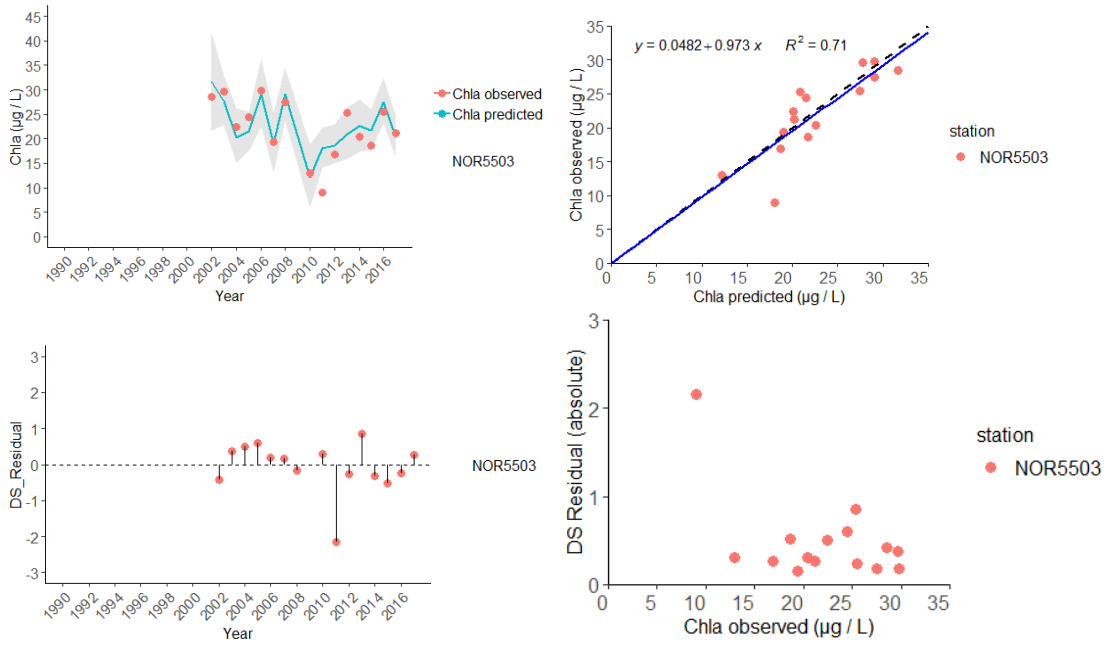

Figure B8.

Vadehavet Grådyb - Ho Bugt v Langli, St. RIB1610002

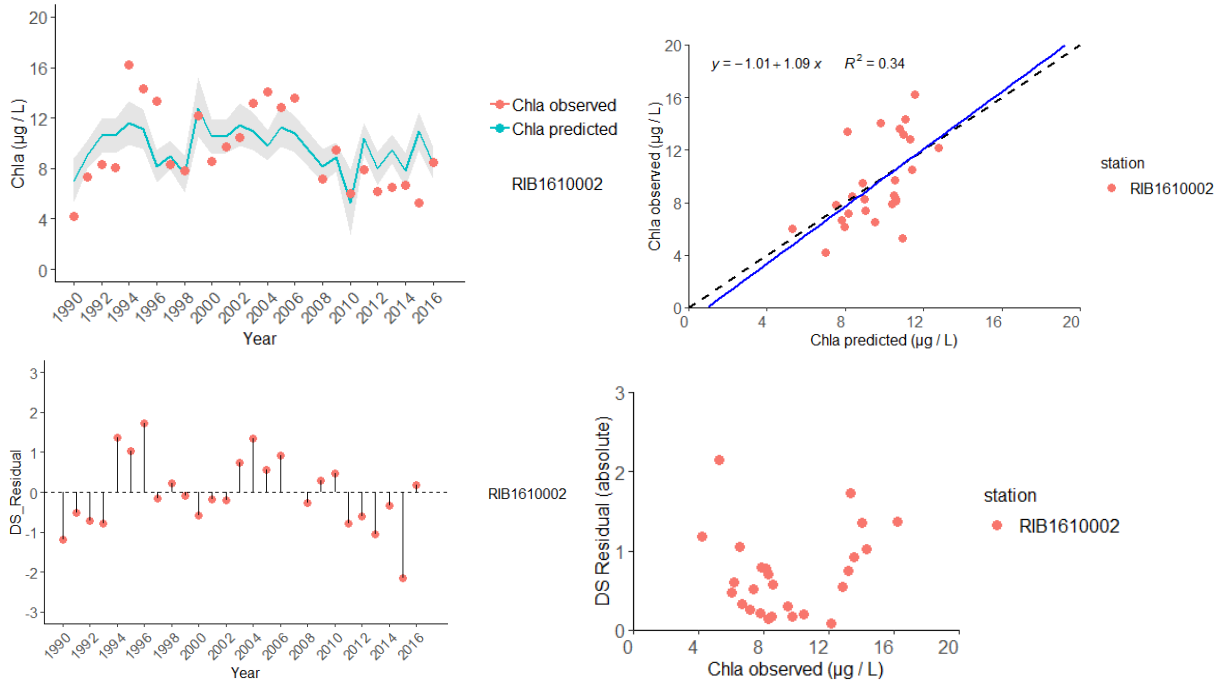

Figure B9.

Vadehavet Knudedyb, St. RIB1620014

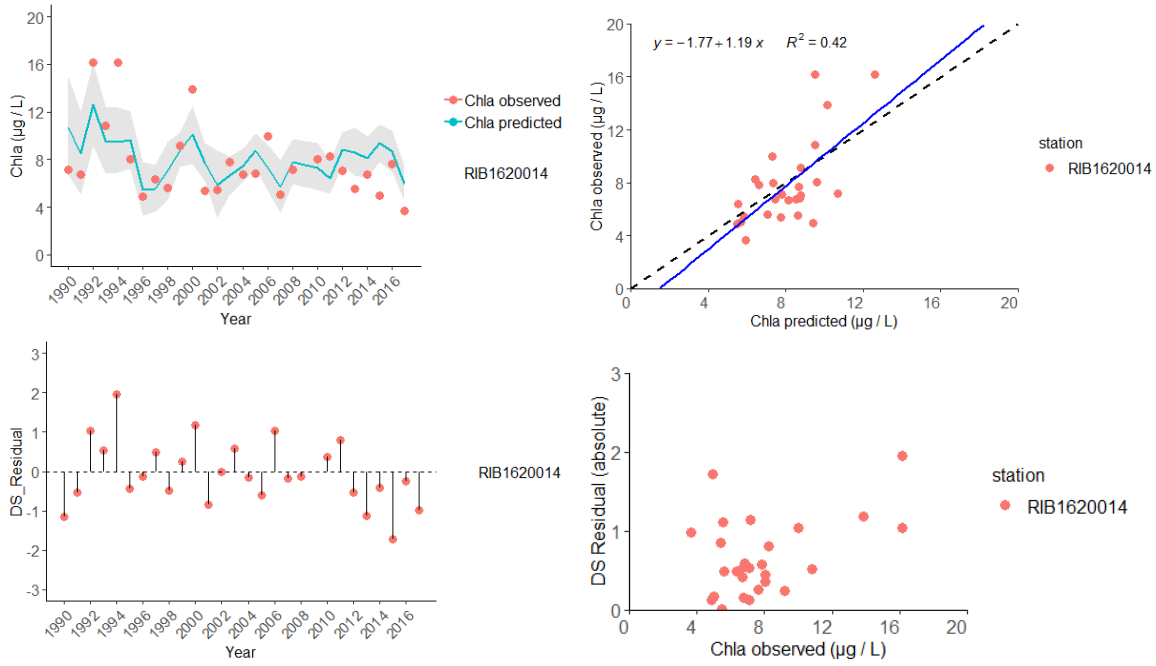

Figure B10.

Ringkøbing Fjord nord, St. RKB1

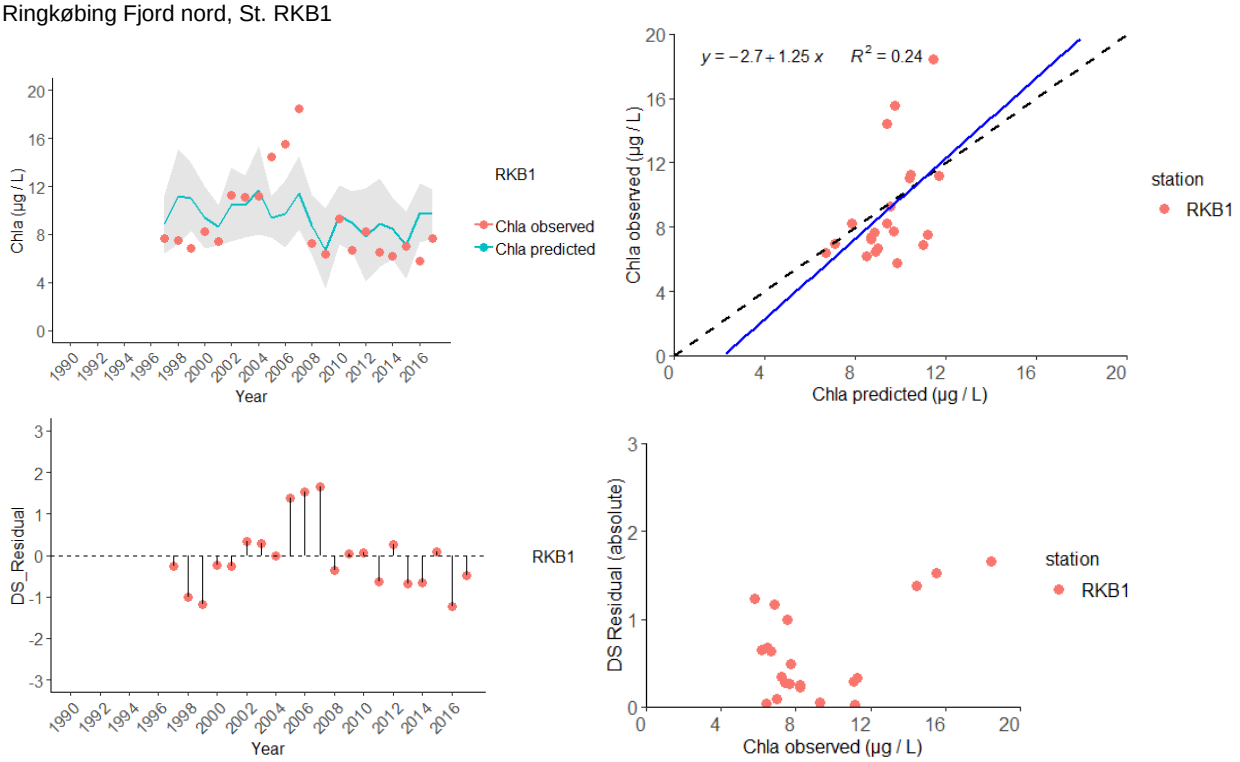

Figure B11.

Nissum Fjord, St. RKB22

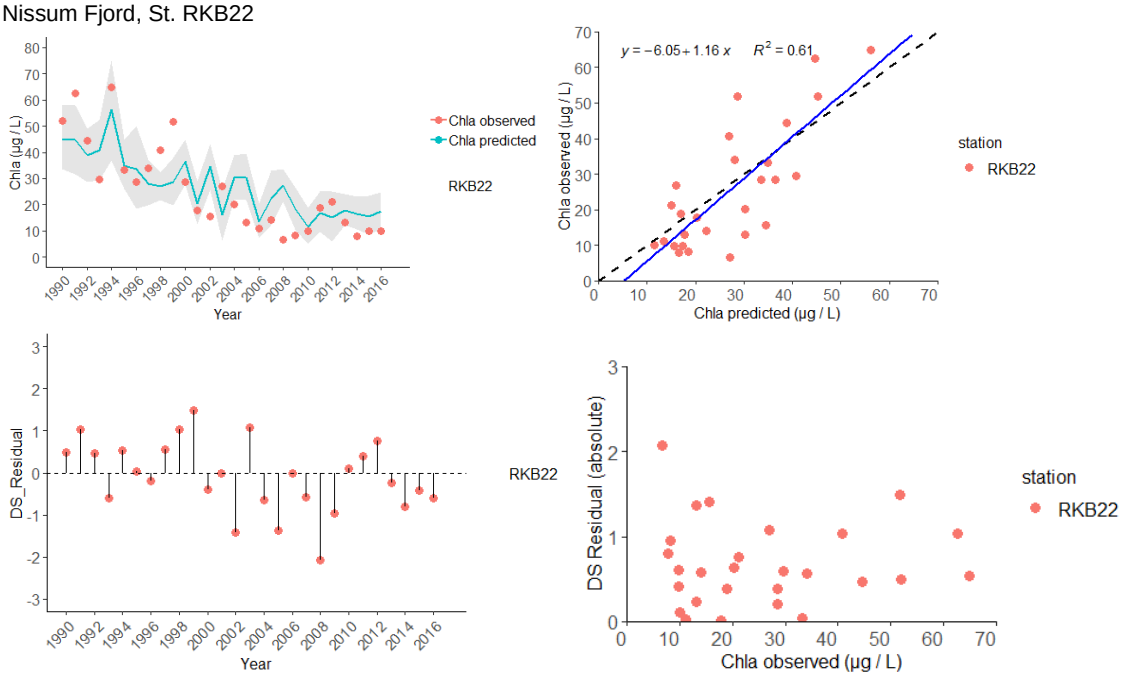

Figure B12.

Lister Dyb, St. SJY1

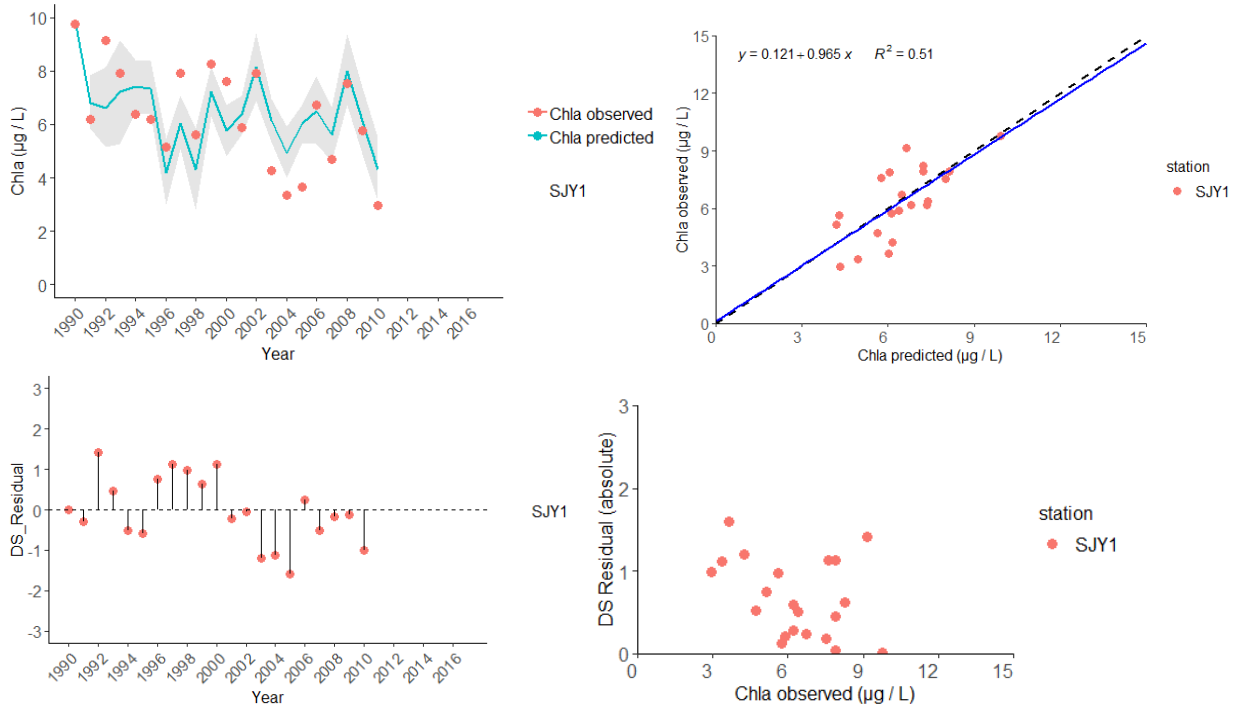

Figure B13.

Als Fjord, St. SJY13B

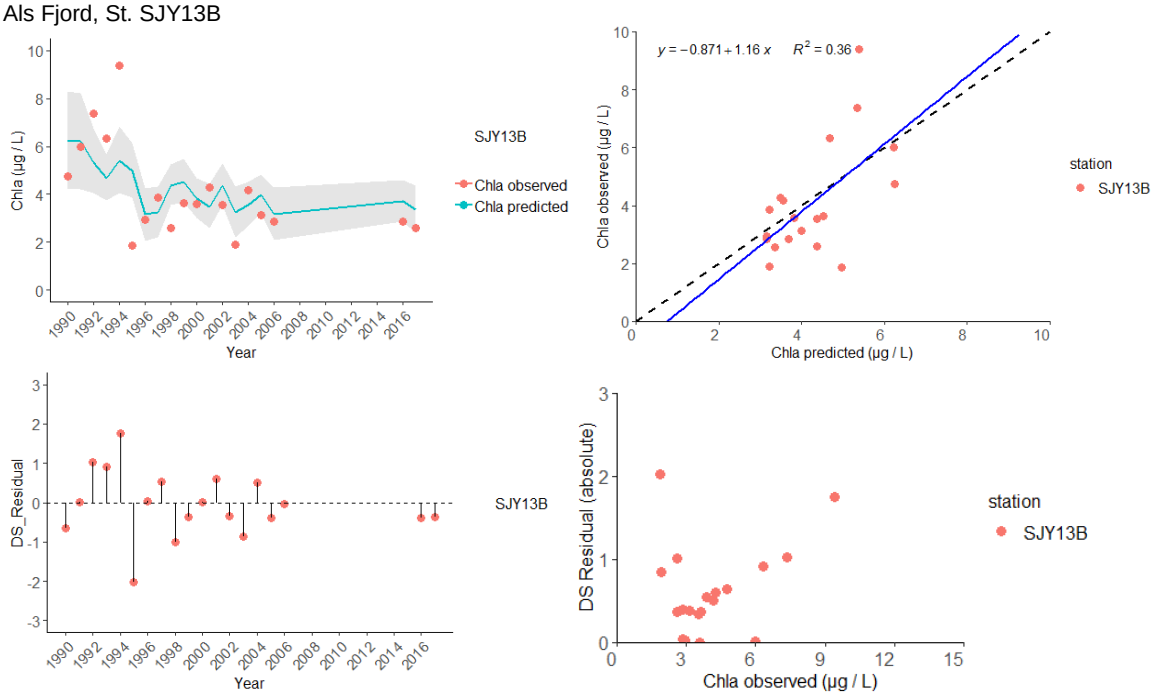

Figure B14.

Genner Bugt, St. SJY19

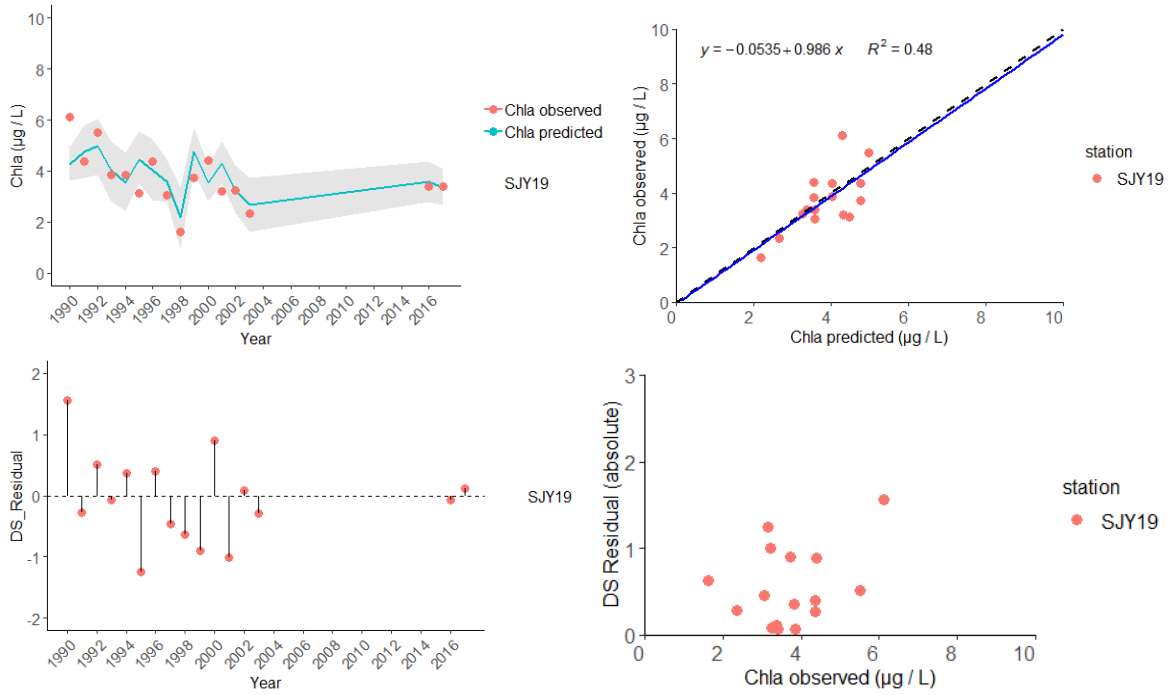

Figure B15.

Lister Dyb v Rømø Havneby, St. SJY3

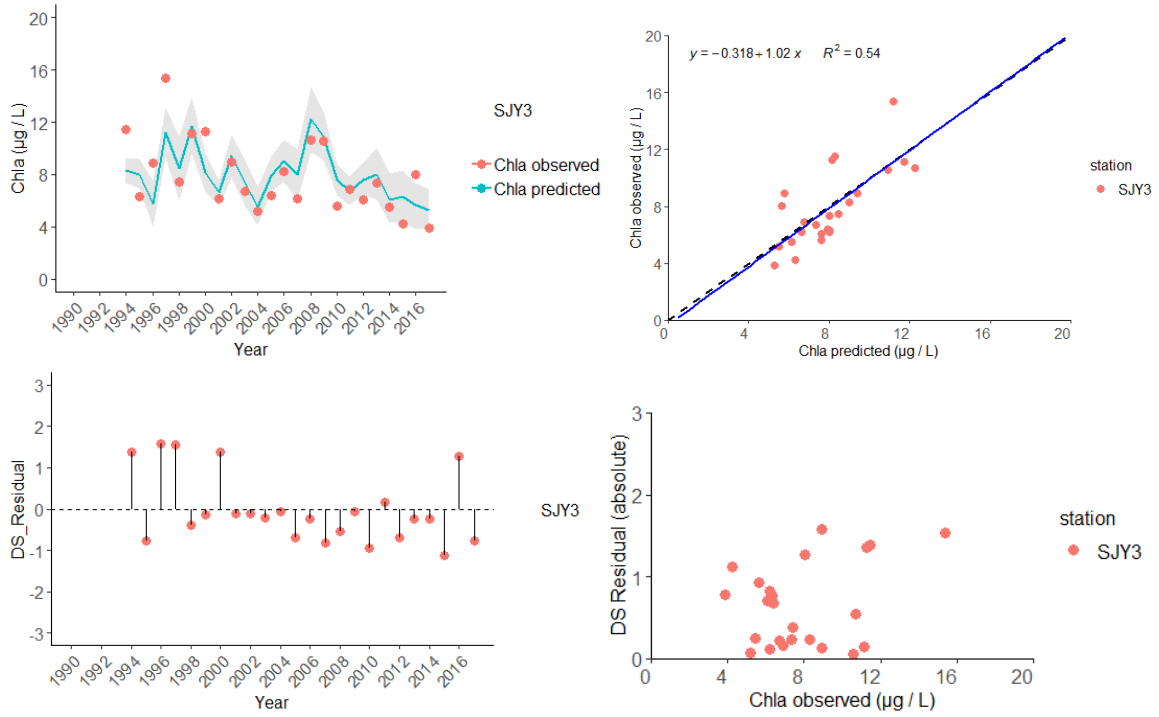

Figure B16.

Flensborg Fjord inder, St. SJYKFF2

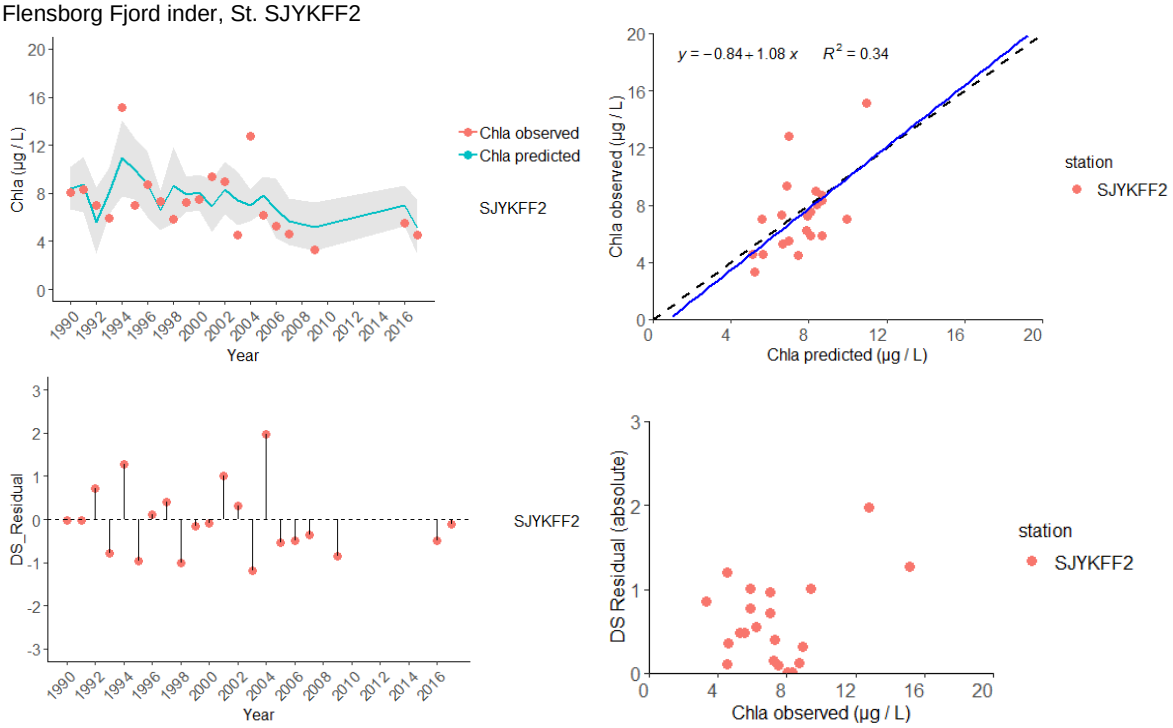

Figure B17.

Kolding Fjord, St. VEJ0003350

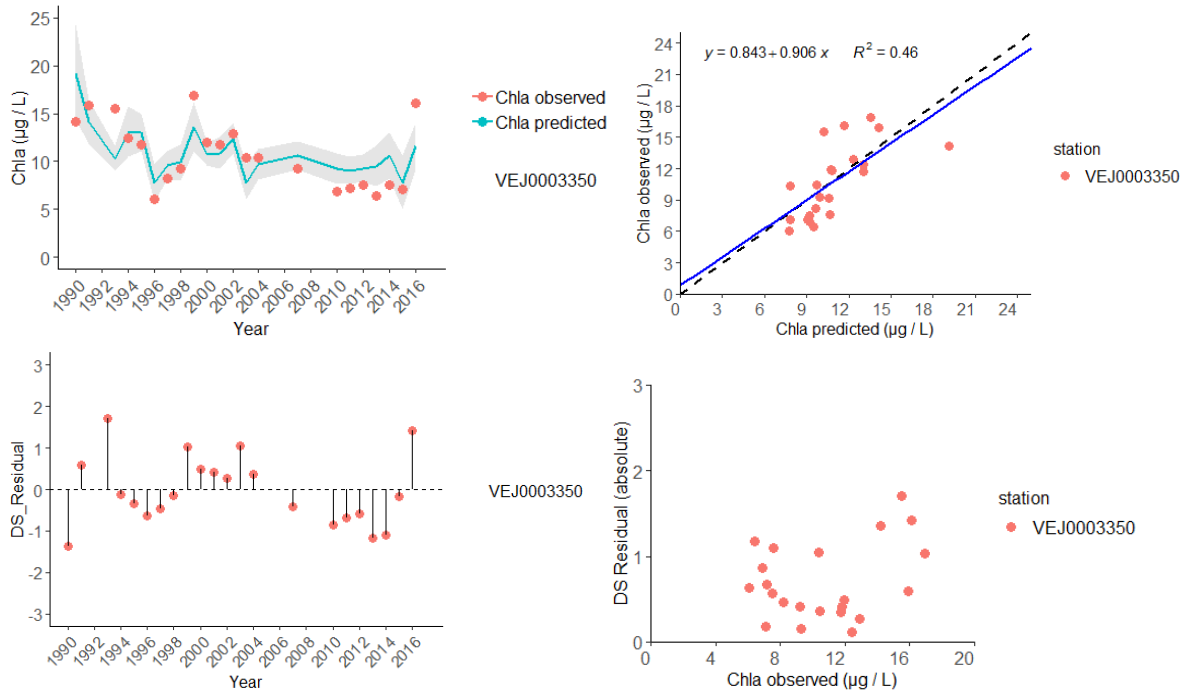

Figure B18.

Vejle Fjord, St. VEJ0004273

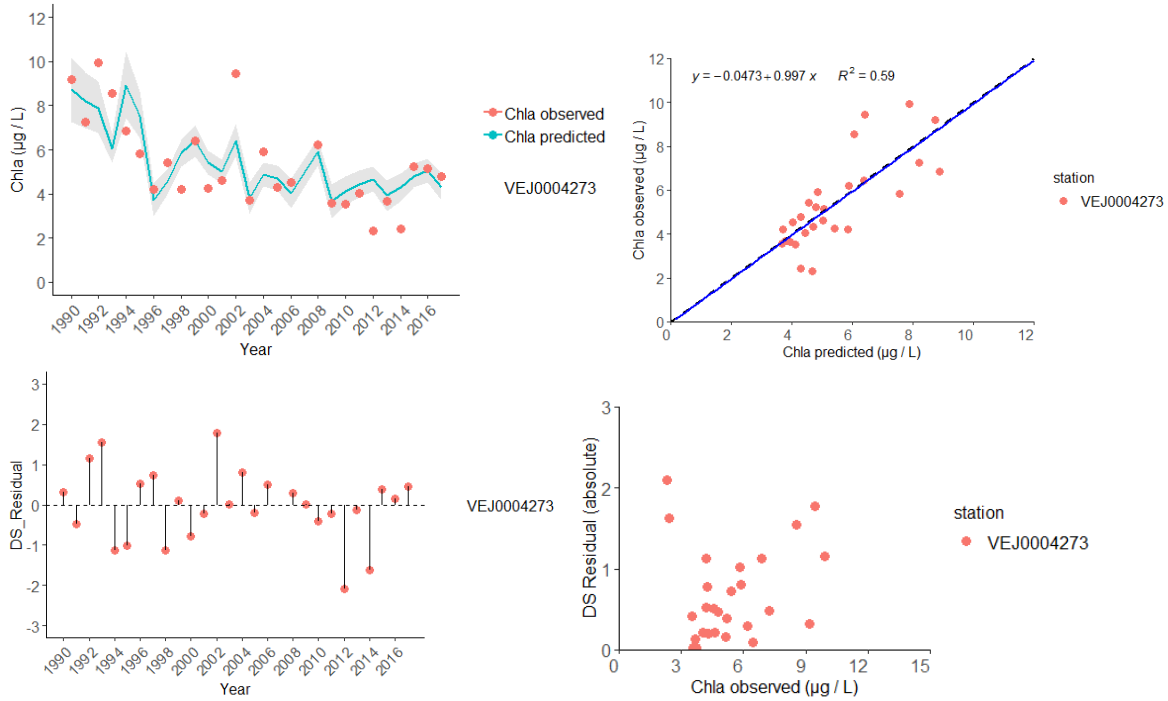

Figure B19.

Horsens inder, St. VEJ0005790

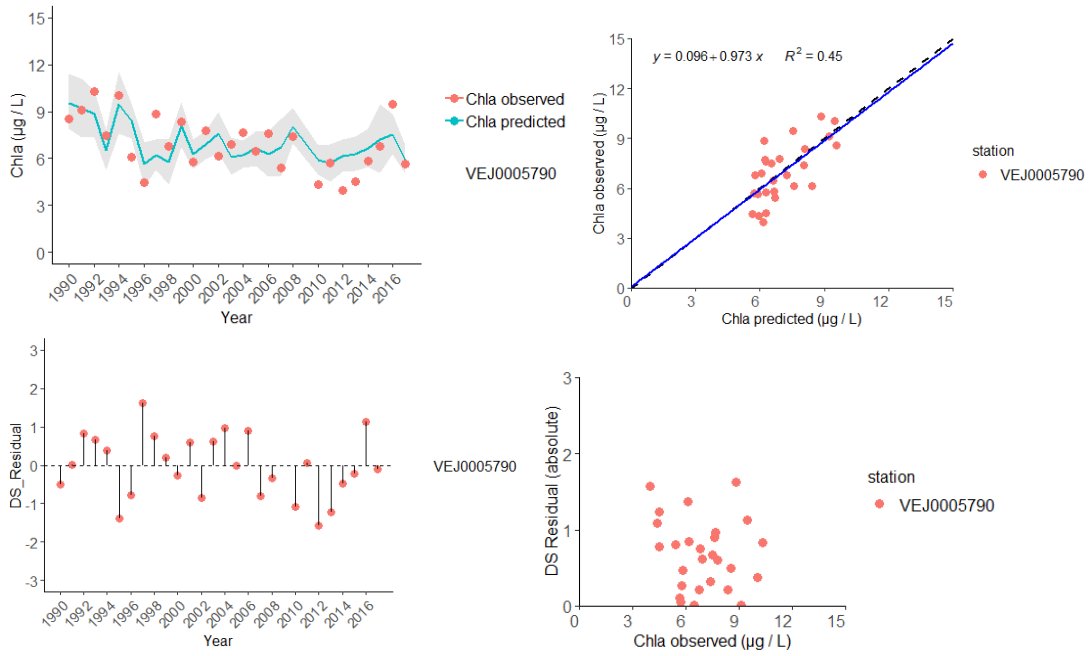

Figure B20.

Hjarbæk Fjord, St. VIB3729-00001

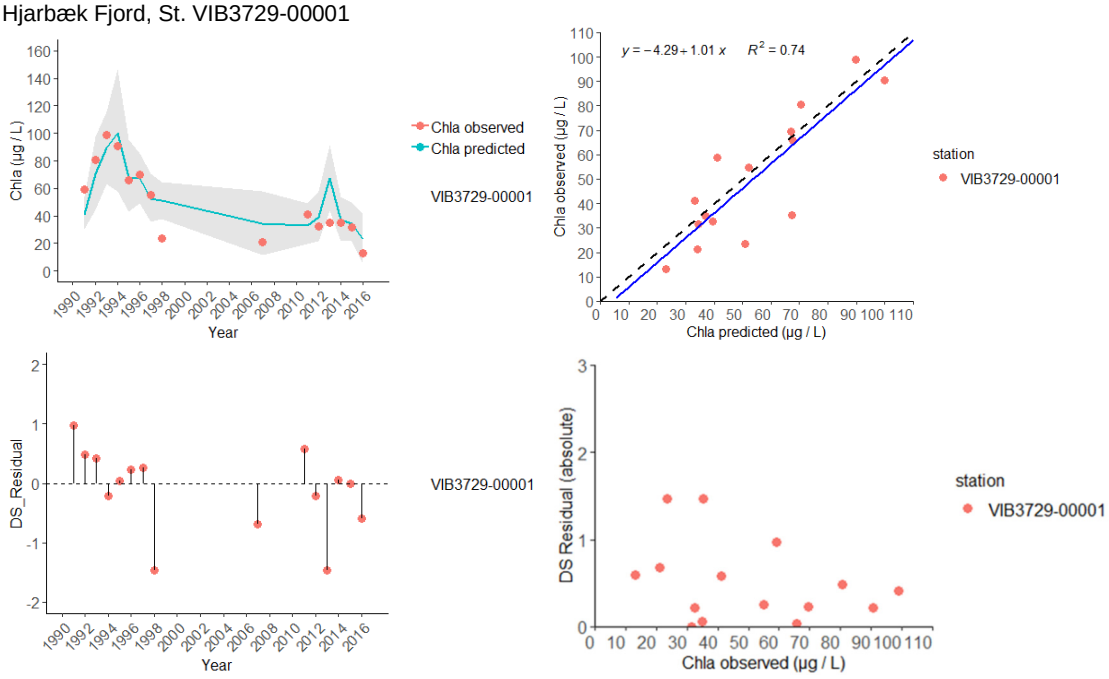

Figure B21.

Isefjord dybt basin, St. VSJ10003

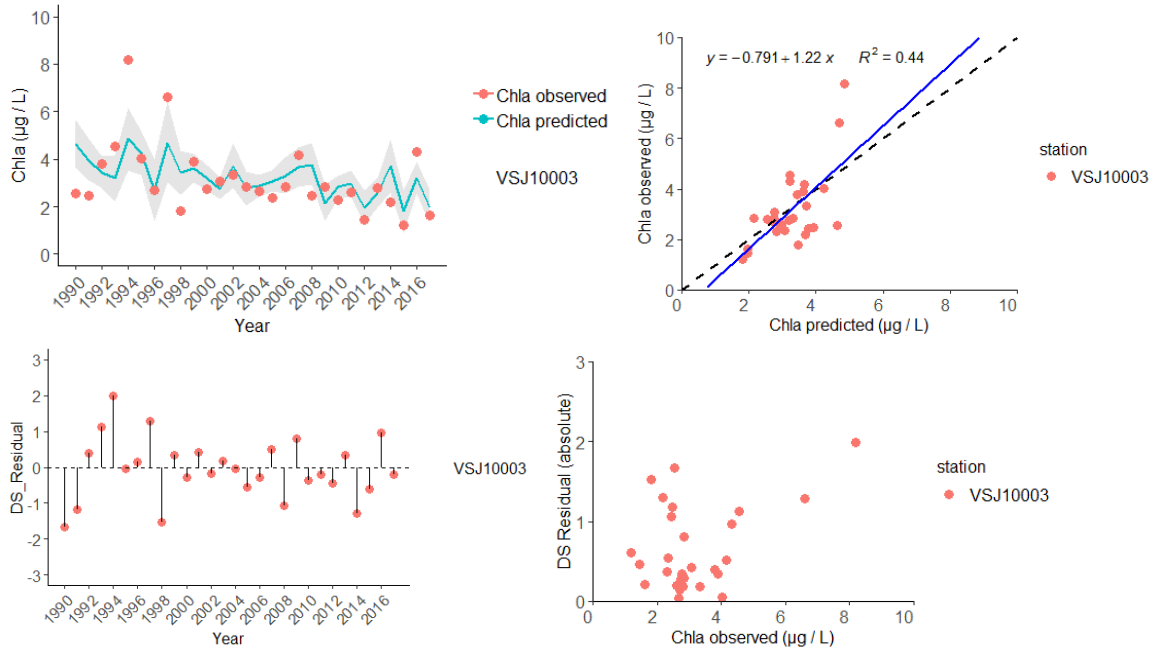

Figure B22.

Isefjord inderbredning, St. VSJ10006

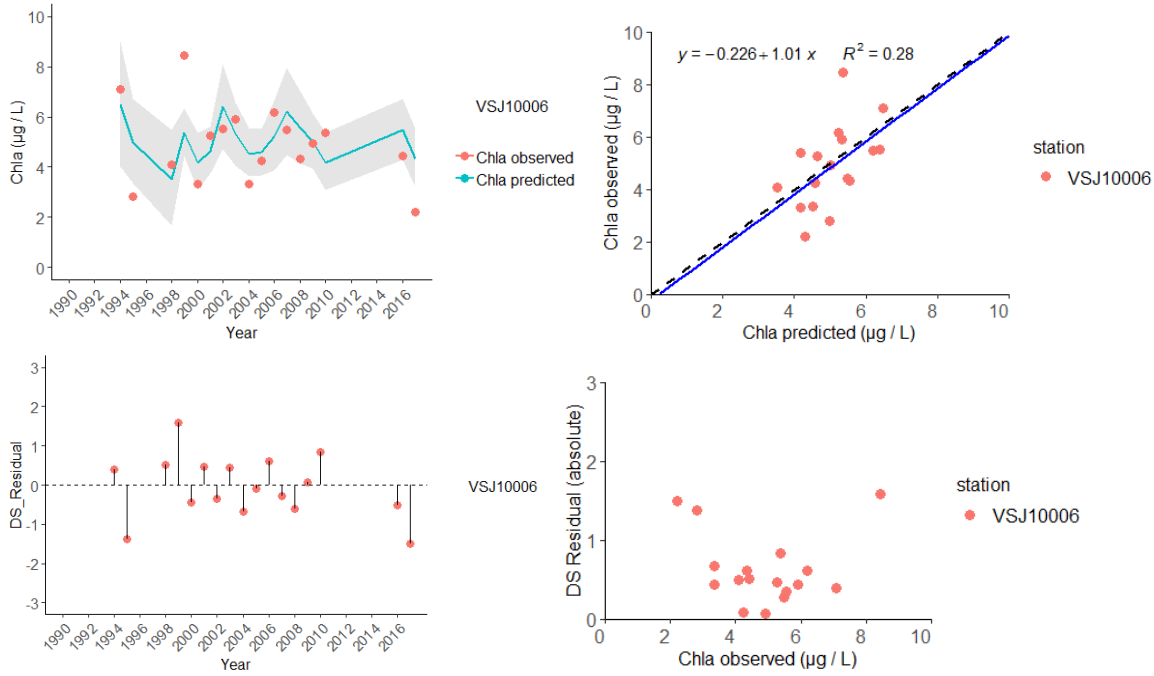

Figure B23.

Kalundborg fjord yder, St. VSJ41007

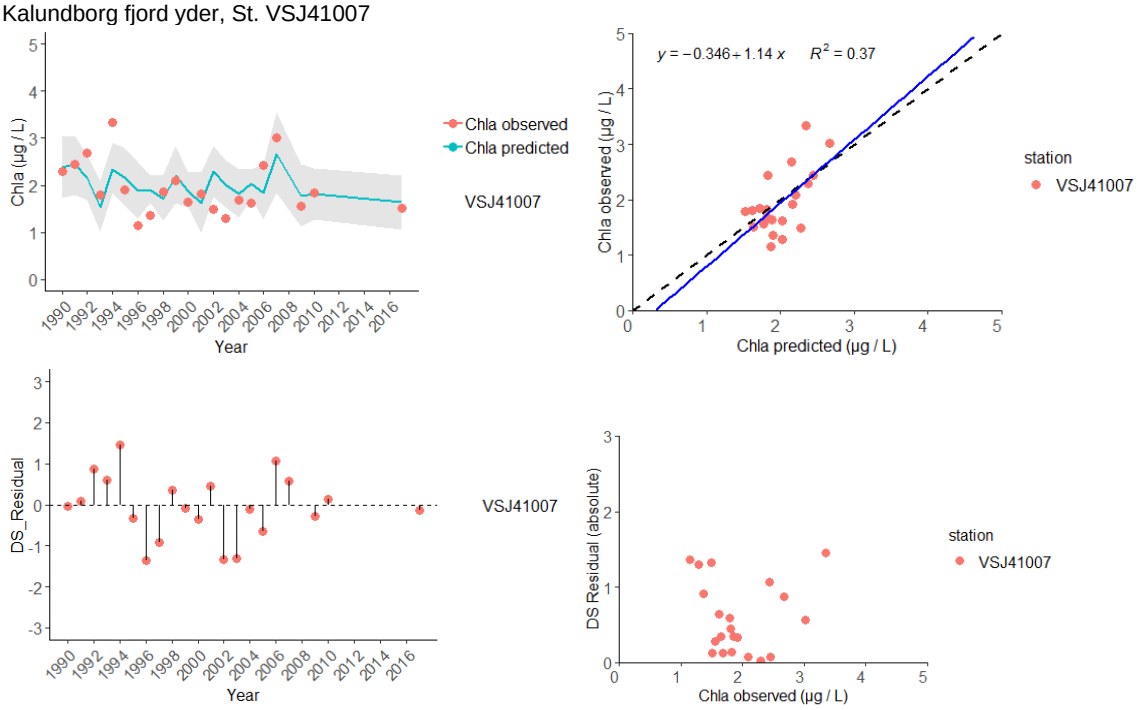

Figure B24.

Kalundborg Fjord inder, St. VSJ41008

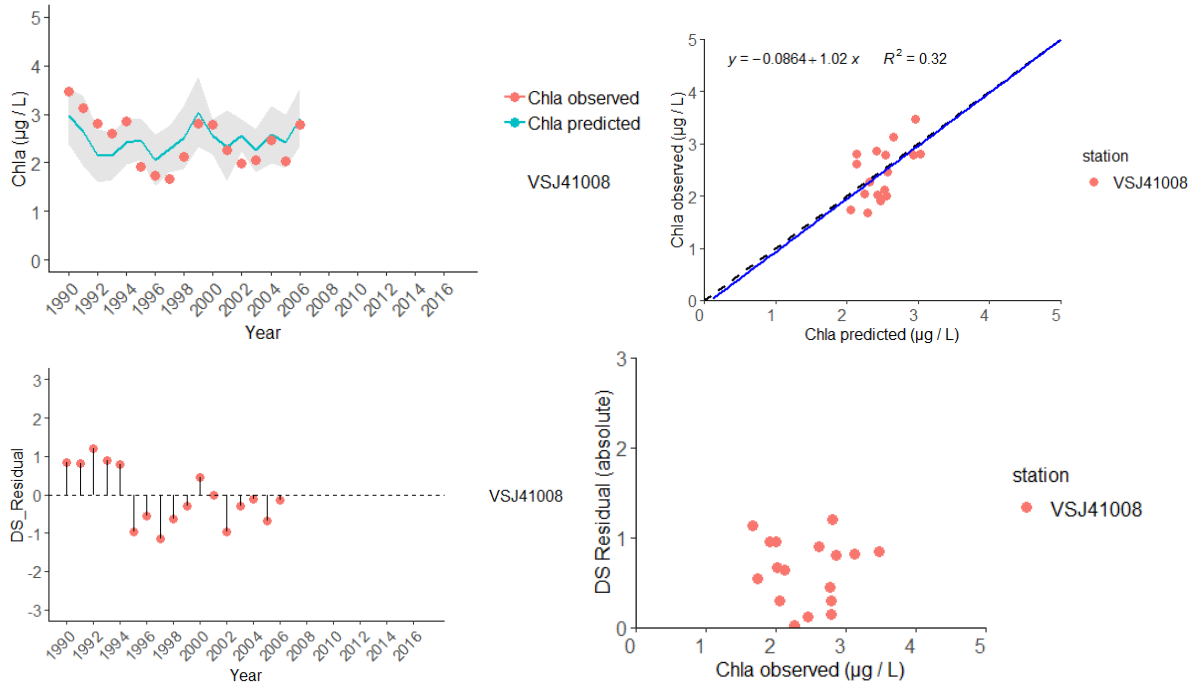

Figure B25.

Skælskør Fjord, St. VSJ51013

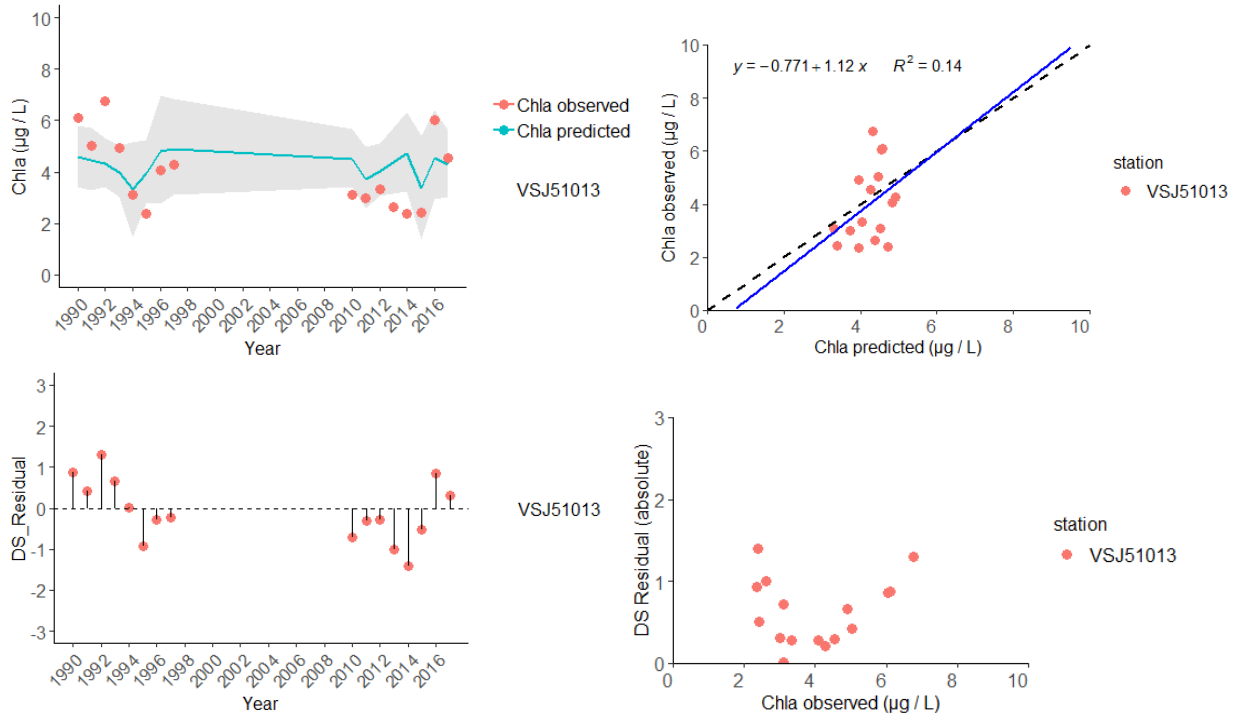

Figure B26.

MODELLING CHLOROPHYLL-A CONCENTRATIONS IN DANISH COASTAL WATERS USING A BAYESIAN MODELLING APPROACH

Documentation report

In this study a Bayesian group model and 46 single station models were developed for estimation of chlorophyll a concentration, nutrient input relationship, in Danish waterbodies. The models relationships were derived to support the Danish EPA, setting input targets to fulfil the EU water framework directive. As expected nutrient input was the most common predictor of chlorophyll a concentration in single model stations (31 of 46) and temperature the second most abundant predictor. In the group model nutrient input, water temperature, salinity and water column stability, were the best predictors of chlorophyll a concentration in the 42 coastal monitoring stations, in the model.

ISBN: 978-87-7156-639-0
ISSN: 2244-9981