APPLICATION OF FREQUENCY-DEPENDENT NUDGING IN BIOGEOCHEMICAL MODELING AND ASSESSMENT OF MARINE ANIMAL TAG DATA FOR OCEAN OBSERVATIONS

by

Karl Bryan Lagman

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2013

DALHOUSIE UNIVERSITY

DEPARTMENT OF OCEANOGRAPHY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "APPLICATION OF FREQUENCY-DEPENDENT NUDGING IN BIOGEOCHEMICAL MODELING AND ASSESSMENT OF MARINE ANIMAL TAG DATA FOR OCEAN OBSERVATIONS" by Karl Bryan Lagman in partial fulfillment of the requirements for the degree of Master of Science.

Dated: June 28, 2013

External Examiner: ______________________________

Supervisor: ______________________________

Readers: ______________________________

____

____

DALHOUSIE UNIVERSITY

DATE: June 28, 2013

AUTHOR: Karl Bryan Lagman

TITLE: APPLICATION OF FREQUENCY-DEPENDENT NUDGING IN BIOGEOCHEMICAL MODELING AND ASSESSMENT OF MARINE ANIMAL TAG DATA FOR OCEAN OBSERVATIONS

DEPARTMENT OR SCHOOL: Department of Oceanography

DEGREE: M.Sc. CONVOCATION: October YEAR: 2013

Signature of Author

To Christiane, Karol and Renard...

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Numerical models are powerful and widely used tools for environmental prediction; however, any model prediction contains errors due to imperfect model parameterizations, insufficient model resolution, numerical errors, imperfect initial and boundary conditions etc. A variety of approaches is applied to quantify, correct and minimize these errors including skill assessments, bias correction and formal data assimilation. All of these require observations and benefit from comprehensive data sets. In this thesis, two aspects related to the quantification and correction of errors in biological ocean models are addressed: (i) A new bias correction method for a biological ocean model is evaluated, and (ii) a novel approach for expanding the set of typically available phytoplankton observations is assessed.

The bias correction method, referred to as frequency-dependent nudging, was proposed by *Thompson et al.* (Ocean Modelling, 2006, 13:109-125) and is used to nudge a model only in prescribed frequencies. A desirable feature of this method is that it can preserve high frequency variability that would be dampened with conventional nudging. The method is first applied to an idealized signal consisting of a seasonal cycle and high frequency variability. In this example, frequency-dependent nudging corrected for the imposed seasonal bias without affecting the high-frequency variability. The method is then applied to a non-linear, 1 dimensional (1D) biogeochemical ocean model. Results showed that application of frequency-dependent nudging leads to better biogeochemical estimates than conventional nudging.

In order to expand the set of available phytoplankton observations, light measurements from sensors attached on grey seals where assessed to determine if they provide a useful proxy of phytoplankton biomass. A controlled experiment at Bedford Basin showed that attenuation coefficient estimates from light attenuation measurements from seal tags were found to correlate significantly with chlorophyll. On the Scotian Shelf, results of the assessment indicate that seal tags can uncover spatio-temporal patterns related to phytoplankton biomass; however, more research is needed to derive absolute biomass estimates in the region.

LIST OF ABBREVIATIONS AND SYMBOLS USED

Abbreviations	Description
1D	one-dimensional
3D	three-dimensional
AR2	Second order auto-regressive process
CO_2	Carbon dioxide
CTD	Conductivity-Temperature-Depth sensor
DA	Data assimilation
hyper-pro	hyper-spectral sensor
MATLAB	MATrix LABoratory
ROMS	Regional Ocean Modeling System

Roman symbol	Description	Units
$a(550)$	Absorption excluding water at 550 nm	m^{-1}
chl	chl concentration	mg chl m^{-3}
Chl	chl concentration (model variable)	mg chl m^{-3}
chl_z	chl concentration at depth z	mg chl m^{-3}
e	Wavelength specific constant for bio-optical relationship	
E_d	Downwelling irradiance	$W \cdot m^{-2}$
F	Fluorescence	μg L^{-1}
H	Depth of water column	m
h_t	Mapping from x_t^f onto y_t	
K_d	Attenuation coefficient of downwelling irradiance	m^{-1}
$K_d(\lambda)$	Wavelength specific K_d	m^{-1}
K_t	Kalman gain matrix	
K_w	K_d for pure water	m^{-1}
$LDet$	Large detritus (model variable)	mmol N m^{-3}
mld	Mixed layer depth	m

Roman symbol	Description	Units
mld_{max}	Maximum mixed layer depth	m
mld_{min}	Minimum mixed layer depth	m
MSE	Mean-squared error	
n_d	Number of depth levels	
N_{K_d}	Monthly number of $K_d(550)$ estimates	km^{-2}
N_{seals}	Monthly unique seal visits	km^{-2}
n_t	Number of time steps in a year	
N_{totK_d}	Total number of $K_d(550)$ estimates	km^{-2}
$N_{totseals}$	Total unique seal visits	km^{-2}
NH_4	Ammonium (model variable)	mmol N m^{-3}
NO_3	Nitrate (model variable)	mmol N m^{-3}
o_{ij}	Observations at depth i and time j	
p_{ij}	Model prediction at depth i and time j	
$p(x_t)$	Prior distribution	
$p(x_t\|Y_t)$	Posterior distribution of x_t given Y_t	
$p(y_t)$	Marginal distribution	
$p(y_t\|x_t)$	Data distribution of y_t given x_t	
Phy	Phytoplankton (model variable)	mmol N m^{-3}
q	Normalized mixed layer depth	
$RMSE$	Root mean-squared error	
$SDet$	Small detritus (model variable)	mmol N m^{-3}
t	Time	day
T	Temperature	°C
v_t	Observation error at time t	
w_t	Model error at time t	
x_t	Ocean state vector at time t	
x_t^c	Climatology at time t	
x_t^f	Forecast of the ocean state at time t	
y_t,	Observations at time t	
Y_t	Set of observations up to time t	
z	Water depth	m

Roman symbol	Description	Units
Zoo	Zooplankton (model variable)	mmol N m^{-3}

Greek symbol	Description	Units
$\alpha(550)$	beam attenuation at 550 nm	m^{-1}
δ	stability parameter	
Δz	thickness of vertical grid	m
$\eta_1,\ \eta_2$	vertical diffusivity coefficients in and below the mixed layer	m^2 day^{-1}
$\eta_{winter},\ \eta_{summer}$	assigned vertical diffusivity coefficients in winter and summer	m^2 day^{-1}
γ	nudging parameter	
λ	wavelength	nm
ϕ	non-linear mapping from x_t to x_t^f	
$\phi_1,\ \phi_2$	coefficients of AR2 process at time t-1 and t-2, respectively	
σ_{K_d}	standard deviation of the slope from regression	m^{-1}

CHAPTER 1

INTRODUCTION

Many important biological processes in the ocean are driven by microscopic photosynthetic organisms called phytoplankton. Phytoplankton, which are responsible for nearly half of the planetary primary production (*Field et al.*, 1998), play an essential role in marine food webs and biogeochemical cycles (*Falkowski et al.*, 1998). Over recent years, different phytoplankton species and their habitat are being subjected to constant stress as a result of human impact to Earth's natural cycle. Unprecedented rise in anthropogenic CO_2 in the last 200 years has contributed to warming (*Crowley*, 2000) and acidification (*Doney et al.*, 2009) of the ocean, and anthropogenic contamination of coastal ecosystems through terrestrial and riverine sources have contributed significantly to increased occurrence of excessive production of algae, a process known as eutrophication (*Turner and Rabalais*, 1994; *Andersen et al.*, 2006; *Rabalais et al.*, 2009). All these have driven changes in phytoplankton productivity (*Behrenfeld et al.*, 2006; *Shi et al.*, 2010), population size (*Irwin and Oliver*, 2009; *Boyce et al.*, 2010; *Guinotte and Fabry*, 2010), phenology (*Edwards and Richardson*, 2004) and community composition (*Moran et al.*, 2010; *Hallegraeff*, 2010) which impact marine food web dynamics and ocean biogeochemical cycling. Knowledge of how these phenomena affect the phytoplankton community- in its important position at the base of the aquatic food chain- is vital.

Biogeochemical models are important tools to study phytoplankton processes and their role in food web dynamics and biogeochemical cycling. They help us build mechanistic understanding of important biogeochemical processes, particularly carbon, nitrogen and phosphorus cycles. Some examples of the application of biogeochemical models include the study of the role of nitrogen cycling on the nitrogen budget on continental shelf (*Fennel* *et al.*, 2006), the effect of phosphorus limitation on primary productivity (*Laurent et al.*, 2012), and the role of dynamics of carbon fluxes in fueling hypoxia in bottom waters (*Justic et al.*, 2002). Biogeochemical models are also used to reconstruct past and predict future scenarios including changes that impact marine organisms, particularly marine species associated with variability in phytoplankton's response to changing environment. For example, the biogeochemical model developed by *Thomas et al.* (2012) was used to predict that future warming is expected to cause poleward shifts in phytoplankton's thermal niche accompanied by a decline in their tropical diversity. A gradual shift toward smaller primary producers in a warmer ocean is predicted as smaller organisms are more able to tolerate increased temperature (*Daufresne et al.*, 2009; *Moran et al.*, 2010). Future changes in ocean carbonate chemistry are also predicted to affect phytoplankton's ability for photosynthesis, calcifications, and nitrogen fixation (*Rost et al.*, 2008). All these are likely to alter spatial and temporal distribution of primary and secondary pelagic production, affecting interaction between trophic levels.

The current generation of ocean biogeochemical models combine information about physical forcing, chemical cycling, phytoplankton physiology, and ecological structure to simulate the response of lower trophic levels to variable external forcing (*Gnanadesikan et al.*, 2011, and references therein). But since physical and biogeochemical processes in the ocean can be highly variable and involve interactions on multiple scales, baseline models often experience drifts and biases (*Gregg*, 2007; *While et al.*, 2010, reported biases in their prognostic runs). Many factors contribute to drift and bias in biogeochemical models including inadequate model resolutions, poor parameterizations of physical and biological processes and inaccurate boundary conditions. But since biogeochemical models are built around established ocean circulation models (e.g. *Haidvogel et al.*, 2008), the focus of improvement is usually on the biogeochemical aspects of the model. An example of increased biogeochemical resolution is the study of *Lehmann et al.* (2009) who considered two functional phytoplankton groups to represent different ecological regimes (diatoms and picoplankton) since a wide range of conditions is less likely to be captured by a single phytoplankton parameterization. Their model yields considerable improvement over other models with a single phytoplankton parameter (e.g. *Fennel et al.*, 2006). The use of multiple functional groups of planktons can allow further be used to examination complex ecological questions such as the competition patterns and structural shifts in planktonic community (e.g. *Ramin et al.*, 2008). It is however a daunting task to capture the contribution of all the relevant species in a model. Simpler models are still favored because they are easier to interpret and less computationally expensive. Using only one functional phytoplankton and zooplankton variable, *Mattern et al.* (2012) incorporated the effect of variation in species abundance in time and space by considering parameters that vary in time. Significant improvement over a single parameter model is reported since their parameterization possibly captures the different physiological characteristics of plankton species that dominate at different times of the year.

Biogeochemical models are only useful however if they properly represent the system of interest and results can be validated. It is for this reason that observations are inseparable from modelling. Validating a model with observations serves as the ultimate end that decides the fate of a model. In oceanography, observations usually come from cruises, stations, floats and satellite measurements. Cruises, stations and floats measurements are good sources of *in situ* data. They resolve the vertical dimension well but they suffer from limited spatial coverage. Satellite measurements on the other hand are excellent sources of spatial and temporal data, however can only measure the ocean surface.

Getting good agreement between model and observation is challenging. Due to the wide range and complexity of potential contributors to errors (e.g. physical, biogeochemical), general solutions to the problem are not available. Recently, there has been a growing number of studies that make use of observations to produce better models (another use of observations in ocean modelling). The strategy is to blend model dynamics with observations (hence called data assimilation) to improve a model's hindcast, nowcast and forecast abilities. Observations contain information about the true state of the ocean while models have no time and space limitation, thus when combined are thought to produce better estimates of the ocean state than using model or observations alone. Examples of such methods are illustrated by biogeochemical assimilation studies with models ranging from one-dimensional (*Allen et al.*, 2003; *Eknes and Evensen*, 2002) to three-dimensional (*Hu et al.*, 2012; *Natvik and Evensen*, 2003). Data assimilation (DA) opens up the possibility of examining incremental model improvements as well as to help identify, and ultimately reduce, model errors (e.g. *Nerger and Gregg*, 2008). DA can thus be viewed as an integral part of the model building process.

The process of combining model and data to produce better model estimates can be viewed as a two-step process. Since models oftentimes suffer from drifts and biases, the model forecast can be nudged first with observation climatology to suppress the bias before point observations are assimilated (see Figure 1.1). To introduce the concept of bias correction and relate it to data assimilation, the Bayesian framework is arguably the most useful in motivating the discussion.

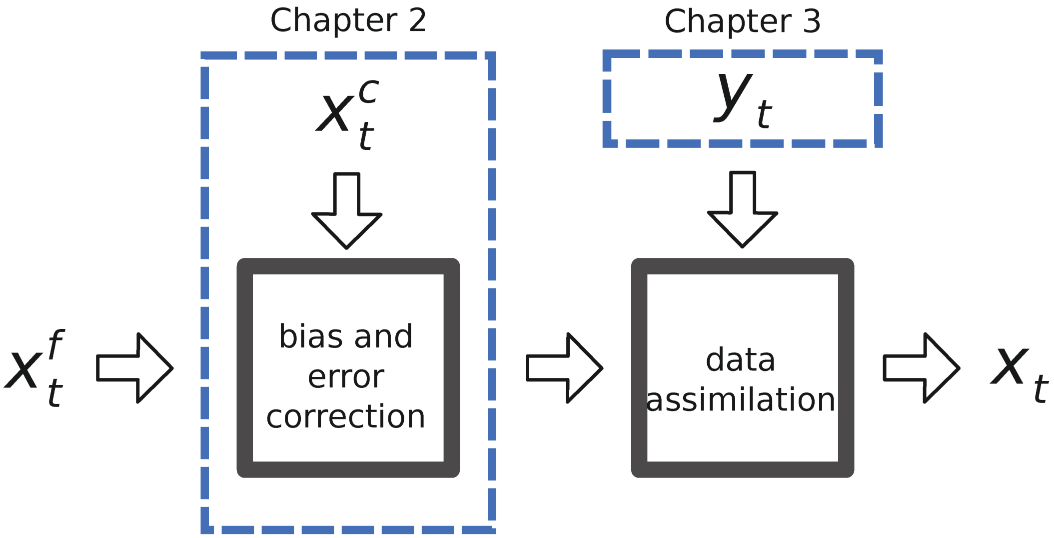


Figure 1.1: A two-step process in combining model and data to improve oceanographic model state estimation. The free run, x_t^f, is nudged with the climatology (step 1), x_t^c, to correct for drifts and biases. Observations, y_t, are then assimilated (step 2) to produce better model state estimate, x_t.

Lorenc (1986) was the first to study in the geosciences that DA equations can be derived from Bayesian probability. *Wikle and Berliner* (2006) showed that the Bayesian approach provides a basis to DA that it is analogous to the scientific method: one has a prior belief (information), collects data, and then updates the belief given the new data. To see this clearly, let x_t denote the unobservable quantities of interest and y_t denotes the observations. One can then formulate the data distribution as $p(y_t|x_t)$ from which one can obtain, by Baye's rule, the expression $p(x_t|y_t) = \frac{p(y_t|x_t)p(x_t)}{p(y_t)}$ provided $p(y_t) \in (0, 1]$. In the expression above, $p(x_t)$ is the prior distribution, which describes the *a priori* understanding of the unobservables, $p(y_t)$ is the marginal distribution and $p(x_t|y_t)$ is the posterior distribution, which is an update of our belief given the observations. This conditional density contains all the information on x_t given y_t including conditional mean of x_t and its covariance.

The Kalman filter, which is of the form $x_t = x_t^f + K_t(y_t - h_t(x_t^f))$, can provide an elegant way of updating the conditional mean and variance of x_t given y_t. In the above equation, K_t represents the nudging coefficients organized in the so-called Kalman gain matrix. The vector of observations, y_t, can be modelled as $y_t = h_t(x_t, v_t)$ where h_t is the observations operator which maps the model state, x_t, to y_t and v_t is the observation error. The one-step-ahead-forecast, $x_t^f = \phi(x_t, w_t)$ represents the dynamics of the model where w_t captures the uncertainty in the forecast. For problems of varying degree of complexity, different variants of Kalman filters are used, where the differences only affect the calculation of K_t. For example, a number of solutions to non-linear problems has been proposed which include the extended Kalman filter (e.g. *Grewal and Andrews*, 1993), the singular evolutive Kalman filter (*Pham et al.*, 1998) and the ensemble Kalman filter (*Evensen*, 1994).

Biogeochemical models oftentimes suffer from bias and drift. Although the Kalman filter can be used to nudge the biased models, this will be achieved in a suboptimal way since an important underlying assumption of Kalman filtering is that the prognostic model is unbiased. *Thompson et al.* (2006) proposed to suppress the bias and drift before observations are assimilated to ensure that the model's climatology doesn't deviate from the observed climatology. The model forecast will be nudged first with the observation climatology, x_t^c, to suppress the bias before point observations, y_t, are assimilated. This strategy is summarized in Figure 1.1.

A simplified approach to the Kalman filter is to simply prescribe K_t. If K_t is assumed to be diagonal then the nudging terms take the conventional form $\gamma(x_t^c - h_t(x_t^f))$ at observation points where γ^{-1} corresponds to a scalar relaxation time in model time steps. The main advantage of this simplified form is that it is easy to implement, robust and can keep the model arbitrarily close to the observations. A serious limitation however is the suppression of variability when nudging toward climatological observations, and the introduction of artificial phase lags in the model response (e.g. *Thompson et al.*, 2006, and references therein).

A gain and phase preserving scheme for combining models and observations has recently been introduced by *Thompson et al.* (2006) in an application for suppressing the bias and drift of basin-scale ocean circulation and has subsequently been applied to a number of atmospheric (*Meinke et al.*, 2006; *Radu et al.*, 2008) and ocean models (*Stacey et al.*, 2006; *Thompson et al.*, 2007). The basic idea behind this scheme is to nudge the model toward climatology in prescribed frequency bands. Outside of these bands the model's dynamics are unaffected by the nudging and the model can evolve uninhibited. Given this restriction of nudging to certain frequency bands, the method can be referred to as frequency-dependent nudging. In ocean models, the chosen frequencies that nudging is applied to are often the mean and annual cycle of a variable, which are typically well characterized in climatologies. The main advantage of this method compared to traditional nudging is that it allows high frequency variability to evolve freely even if nudging is strong.

To my knowledge, frequency-dependent nudging has not yet been applied to ocean biogeochemical models. The first goal of this thesis is to evaluate frequency-dependent nudging as one possible approach for bias correction in biogeochemical models (discussed in chapter 2, see Figure 1.1). Its utility in combining a biogeochemical model with an observation-based climatology without suppressing possibly important information contained in the high frequency variations present in biogeochemical models (e.g. blooms with typical scales of 1 week) will be assessed.

A critical part of data assimilation is the availability of high-resolution data (third step in Figure 1.1). Oceanographic data are usually obtained from satellites, ship-based sampling and, more recently, from *in situ* sensors on autonomous platforms and moorings. Recent technological progress in sensor development may even allow us to use pelagic animals as agents of data collection. Many electronic sensors on animals (henceforth referred to as tags) were not primarily intended for quantitative oceanographic measurements, but are capable of collecting high frequency data (up to every second). Considering the high mobility of pelagic animals and their need for frequent nutrition, biologically active regions could be represented better both temporally and spatially than through traditional *in situ* methods. Recent studies indeed show that electronic tags attached to pelagic animals can be reliable data sources (*Teo et al.*, 2009; *Boehlert et al.*, 2001). The second goal of this thesis is to assess data from electronic tags attached to grey seals, *Halichoerus grypus*, and their potential for ocean observations (discussed in chapter 3, see Figure 1.1).

The outline of this thesis is as follows. To address the first goal of this thesis, the bias correction step is discussed in chapter 2. Frequency-dependent nudging is first applied to an idealized signal with annual cycle and high-frequency variability. The application of frequency-dependent nudging to a 1D biogeochemical model is also described and discussed as well in this chapter. Generation of synthetic observations, climatology and prognostic runs are discussed, and results from conventional and frequency-dependent nudging are compared. The second goal of the thesis is addressed in chapter 3. Chapter 3 describes the methods implemented to validate and assess data from electronic tags for ocean observations. A baseline experiment at Bedford basin is described including comparison of estimates of physical and biogeochemical variables from electronic tags to the measured (using specialized sensors) values in the region. Results of estimation of the attenuation coefficient, $K_d(550)$, in Scotian Shelf waters are discussed as well. Results from both chapters are discussed and summarized in chapter 4, including overall conclusions of the whole thesis and suggestions for future work.

CHAPTER 2

APPLICATION OF FREQUENCY-DEPENDENT NUDGING TO BIOGEOCHEMICAL MODELS

In this chapter, the use of frequency-dependent nudging as one possible approach to bias correction for improving model state estimate is discussed. The strategy to assess its application to non-linear systems is summarized in Figure 2.1. A simulated reality is constructed using a complex realistic model (green box, Figure 2.1). The realistic model is used to generate the synthetic observations. The synthetic observations are the "measurable" quantities that represent the simulated reality. The goal is to match the prediction of the simple model with the synthetic observations.

In downscaling from a complex to a simple model, a lot of important information might be suppressed. This can result in biases and errors in the simple model's representation of simulated reality. To address this, the simple model will be nudged with the climatology calculated from the synthetic observations. This step makes sure that the prediction doesn't deviate from the observation climatology. A quantitative assessment between the simple model and potential improvements to be gained in using conventional and frequency-dependent nudging (red box, Figure 2.1) is discussed in this chapter.

To illustrate frequency-dependent nudging, details of the equations are discussed in section 2.1. Section 2.2 then discusses its application to an idealized signal consisting of a seasonal cycle and a high frequency variability. In section 2.3, its application to biogeochemical models is described. Finally, summary and conclusions for the chapter is given in section 2.4.

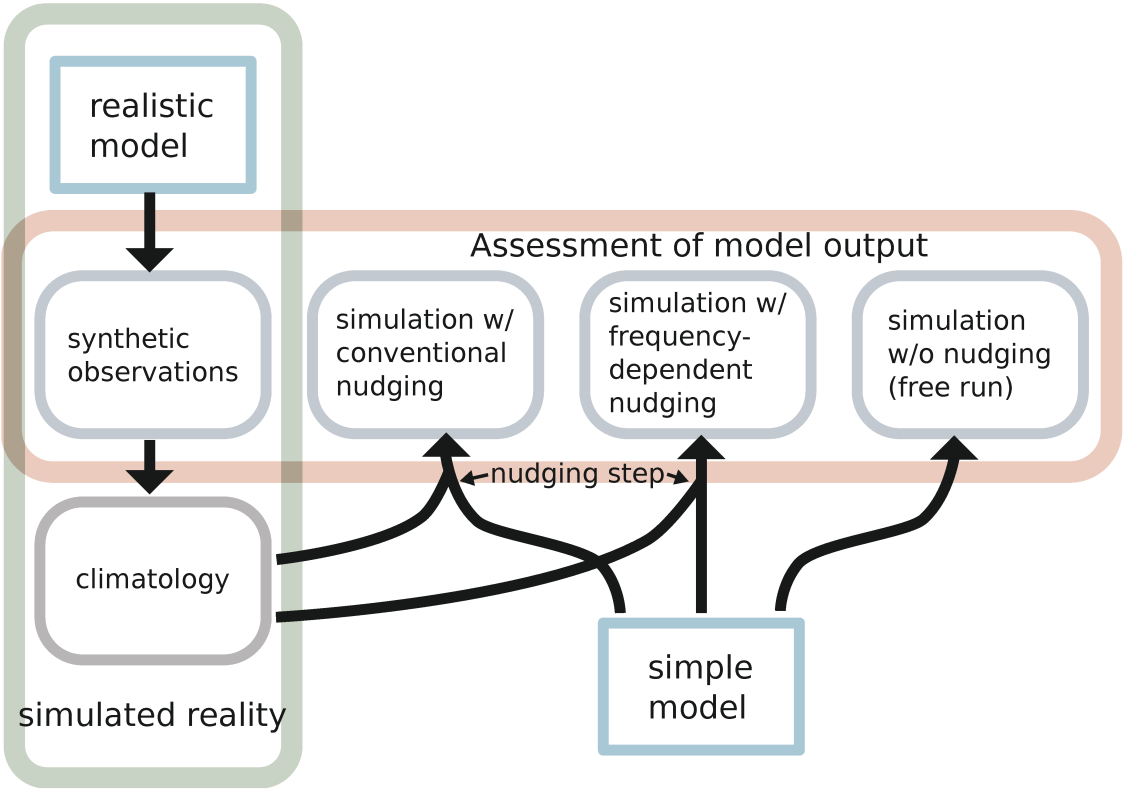


Figure 2.1: Strategy to asses the utility of frequency-dependent nudging to non-linear systems.

2.1 Frequency-dependent nudging

Thompson et al. (2006) proposed a nudging method that suppresses bias and drift while leaving the phase and the amplitude of the model response unaffected. The main innovation of their method over conventional nudging is that the model is nudged in frequencies that are well resolved in climatology, thus called frequency-dependent nudging. Frequency-dependent nudging is implemented as follows:

$$x_t = x_t^f + \gamma \langle x_t^c - x_t^f \rangle \tag{2.1}$$

The only difference between frequency-dependent nudging and conventional nudging is the presence of the angle brackets, $\langle \cdot \rangle$, which denotes nudging in specific frequency bands (0 and 1 cycle per year). A main advantage of frequency-dependent nudging over conventional nudging is that it allows high frequency variability to evolve prognostically while nudging to the climatology. This method is implemented sequentially, and thus observations can be assimilated after the suppression of bias and drift by the frequency-dependent nudging. The reader is referred to *Thompson et al.* (2006) for the details of the implementation of equation 2.1.

2.2 Application of frequency-dependent nudging to an idealized signal

To motivate the discussion about the application of frequency-dependent nudging, assume a system with a response described by a combination of an annual cycle and high frequency variability centred at 32 days (black line in Figure 2.2). This is a simulated reality (as in Figure 2.1) from which synthetic observations, y_t, can be "measured". In this example, further assume that all features of the simulated reality are captured by the synthetic observations. Now assume a simple model

$$x_t^f = \phi_1 x_{t-1}^f + \phi_2 x_{t-2}^f + w_t \tag{2.2}$$

to describe the synthetic observations. In the equation above, ϕ_1=-1.9, ϕ_2=0.94 and $w_t \backsim \mathcal{N}(0, \sigma_w^2)$. Figure 2.2 (green line, above panel) shows that the simple model captures only the high-frequency component of the synthetic observations, but that it is biased. In oceanography, models often suffer from biases due to the complex nature of interaction between components of different time scales (e.g. physical vs biogeochemical components). Thus the seasonal bias observed in the simple model can be thought of as some processes that are missing or misrepresented and can be corrected accordingly. To correct for the observed seasonal bias, data and model will be combined to improve model state estimation. In particular, the climatology (taken from synthetic observations, Figure 2.2) will be used to nudge the simple model to be able to recover the seasonal cycle present in the synthetic observations. Figure 2.2 (lower panel) shows the results of conventional and frequency-dependent nudging ($\gamma = 0.25$) when applied to the simple model. Results show that both nudging methods recover the seasonal cycle however high frequency variability is dampened in conventional nudging. On the other hand, the use of frequency-dependent nudging allows for the recovery of the seasonal cycle while preserving the high-frequency variability, resulting in a closer prediction of the synthetic observations.

Figure 2.3 shows the effect of the nudging strength, γ, on the nudging results. Stronger dampening of the high frequency variability is observed with conventional nudging for increasing γ. For $\gamma = 1$, a method known as direct insertion, the climatology simply replaces the model estimate. The phase of the signal from the simple model is altered as well. In frequency-dependent nudging, increasing γ has no effect on the the amplitude of the high frequency variability. Even at $\gamma = 1$, the high frequency variability is still persistent. The phase is preserved as well as γ increases. Furthermore, the seasonal cycle becomes more prominent as the γ increases.

2.3 Application to a vertically resolved biogeochemical model

In this section, the frequency-dependent nudging is applied to a one-dimensional (1D) biogeochemical model. A similar set-up as in Figure 2.1 is followed. A simulated reality from a three-dimensional (3D) biogeochemical model of *Bianucci et al.* (in prep) was generated using the Regional Ocean Modeling System (ROMS, http://myroms.org *Haidvogel et al.*, 2008). Synthetic observations were generated from “measurements” from 3 different stations shown in Figure 2.4 and were used to calculate a climatology for each station. A simple 1D model is developed to simulate the “measured” synthetic observations. Bias was observed in the simple 1D model since it cannot capture some of

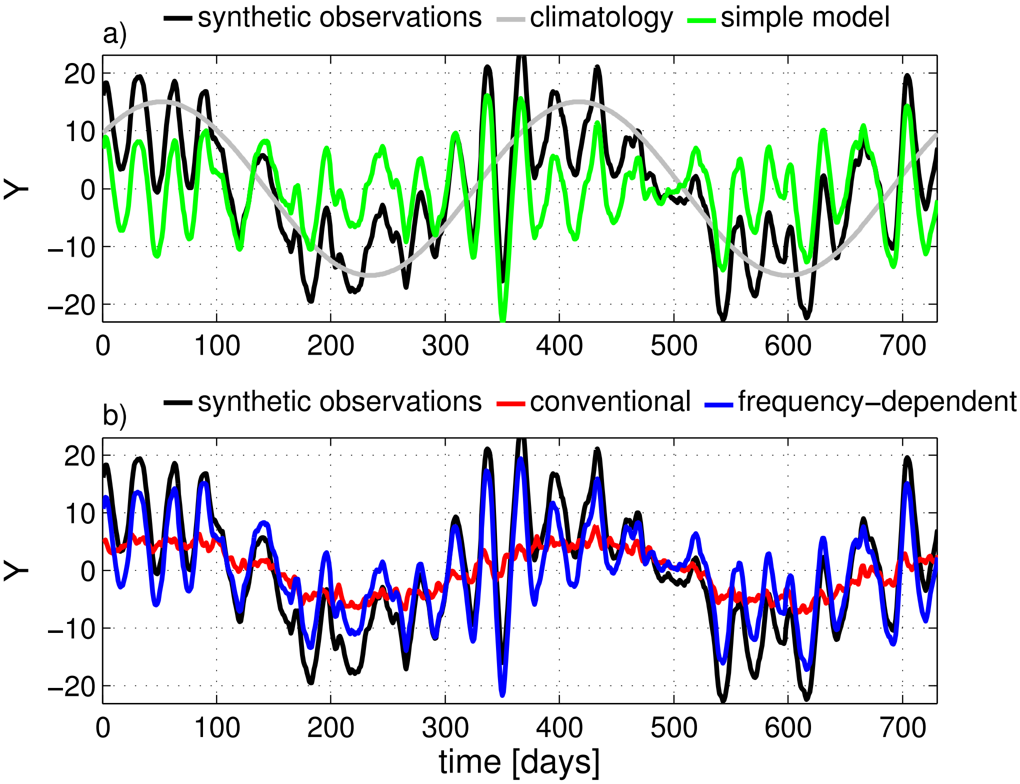


Figure 2.2: Comparison of the synthetic observations to the simple model and results from nudging. (a) The synthetic observations (black line) are generated as a sum of a realization from an AR2 process (ϕ_1=-1.9 and ϕ_2=0.94) with a sampled resonance centred on period of 32 days and a sine function with an amplitude of 30 units and a period of 1 year. The climatology (grey line) is generated using only the sinusoidal component of the synthetic observations. The simple model (green line) is generated using only the AR2 component of the synthetic observations. (b) A nudging strength of $\gamma = 0.25$ is used to generate results from nudging.

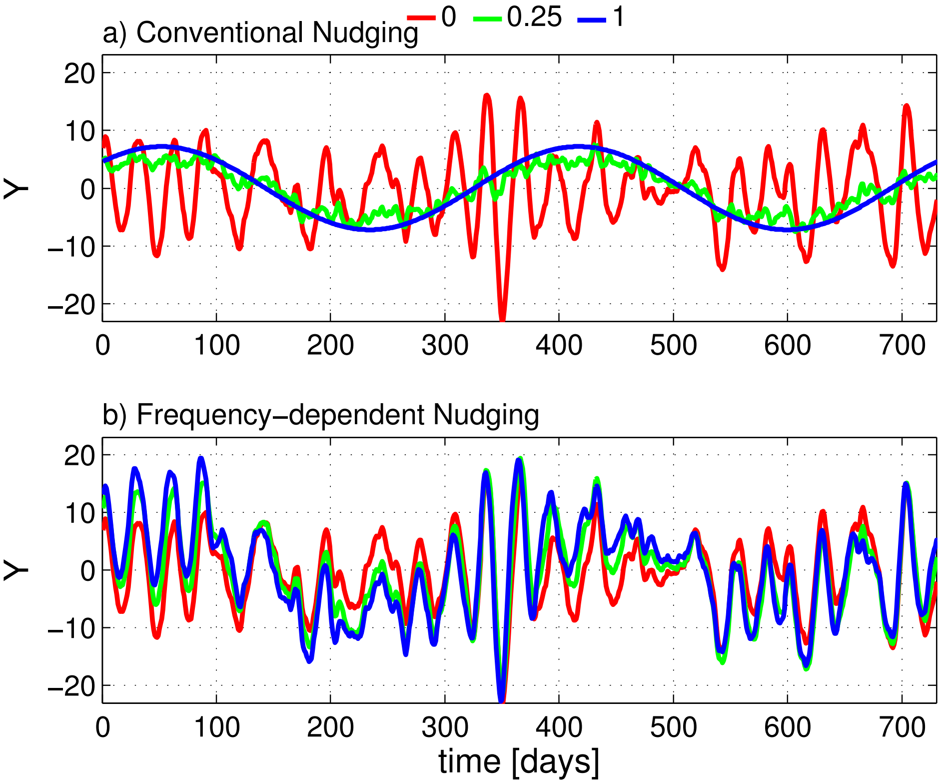


Figure 2.3: Sensitivity analysis of the effect of the nudging strength, γ, on conventional (upper panel) and frequency-dependent nudging (lower panel).

the physical process in present only in 3D systems (e.g. horizontal dynamics). The bias was corrected by nudging the climatology to the simple model. A more detailed description of the 3D and 1D model, and quantitative comparison of the results from conventional and frequency-dependent nudging, are described in the subsequent sections.

2.3.1 The 3D ROMS model

The 3D model is an implementation of the Regional Ocean Modeling System (ROMS, http://myroms.org *Haidvogel et al.*, 2008) coupled to the biogeochemical model of *Fennel et al.* (2006, 2008), forced by atmospheric reanalysis fields of *Large and Yeager* (2004) and is described in detail in *Bianucci et al.* (in prep). The model domain includes the Grand Banks, the Gulf of St. Lawrence, the Scotian Shelf and the Gulf of Maine (Figure 2.4) and is nested within the larger-scale physical model of *Blanco and Sheng* (2012). The model's horizontal resolution is $\backsim$10 km with 30 sigma-layers that are stretched to allow for higher resolution near the surface; it was run for 5 years from 1 January 1999 to 31 December 2004 and the model state was saved every 5 days.

The biogeochemical component of the model is a relatively simple representation of the marine nitrogen cycle and includes two species of dissolved inorganic nitrogen (nitrate, NO_3, and ammonium, NH_4), one functional phytoplankton group, Phy, chlorophyll, Chl, as a separate state variable to allow for photoacclimation, one functional zooplankton group, Zoo, and two pools of detritus representing large, fast-sinking particles, $LDet$, and suspended, small particles, $SDet$. The main processes described in the model are 1) temperature, light- and nutrient-dependent phytoplankton growth with ammonium inhibition of nitrate uptake, 2) zooplankton grazing represented by a Holling-type III parameterization, 3) aggregation of phytoplankton and small detritus to fast sinking large detritus, 4) photoacclimation (i.e. a variable ratio between phytoplankton and chlorophyll), 5) linear rates of phytoplankton mortality, zooplankton basal metabolism, and detritus remineralization, 6) a second order zooplankton mortality, 7) light-dependent nitrification (i.e. oxidation of ammonium to nitrate), and 8) vertical sinking of phytoplankton and detritus. In the implementation of the model, sinking organic matter instantaneously remineralizes once it reaches the bottom. For further details on model justification, equations and parameters the reader is referred to *Fennel et al.* (2006, 2008).

The ROMS model was used to generate the synthetic observations and climatology. The biogeochemical state variables were extracted from the 5-year simulation for the stations shown in Figure 2.4. The first year was discarded as spin-up, the following 4 years were interpolated onto the equidistant vertical layers of the simple 1D model (details of the model are discussed in the next section) and a synthetic observation was produced by harmonic regression using the first 15 harmonics. The resulting synthetic observation is shown in Figure 2.6 (left panels) for station 1 and Figures A.1 and A.6 for stations 2 and 3. Similarly, a climatology is extracted as well from the ROMS model by harmonic regression using only the first harmonic. The resulting climatology is shown in Figure 2.6 (middle panels) for station 1 and Figures A.1 and A.6 for stations 2 and 3.

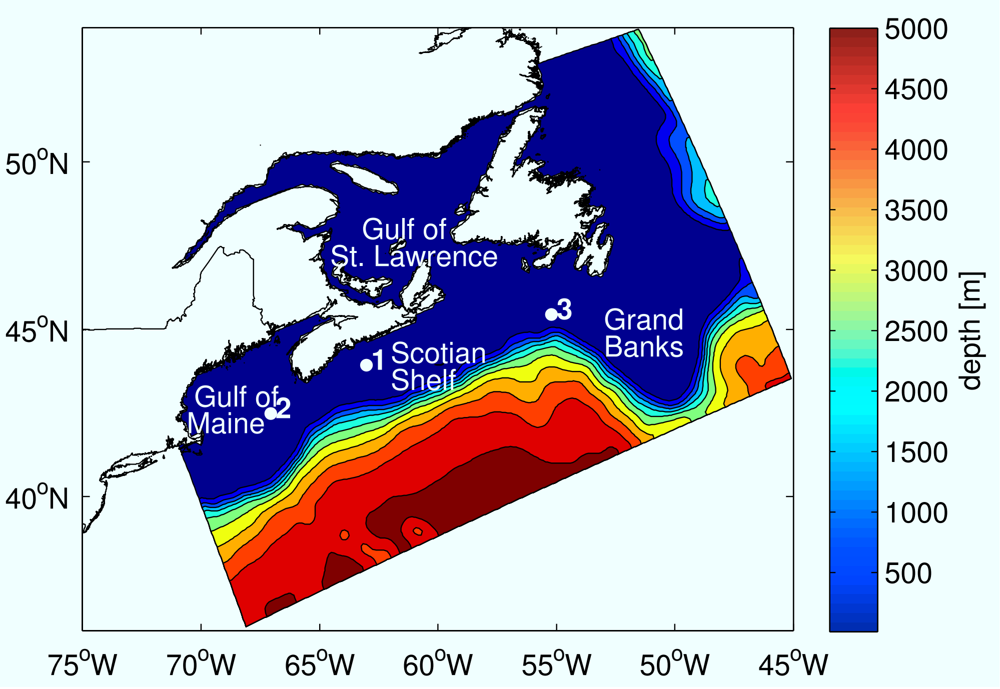


Figure 2.4: Model domain of the ROMS implementation of *Bianucci et al.* (in prep). The colormap shows the bathymetry of the domain. The white dots show the locations where model data were extracted.

2.3.2 The simple 1D model

The simple model has $H/\Delta z$ equidistant layers (H is the water depth, Δz is the layer thickness) and uses a highly simplified physical framework with an imposed, time-varying mixed layer coupled with the biogeochemical model of *Fennel et al.* (2006) (Figure 2.5). Vertical mixing is prescribed with the help of two mixing coefficients, one for the mixed layer (η_1) and one below the mixed layer (η_2), which vary seasonally as described further below.

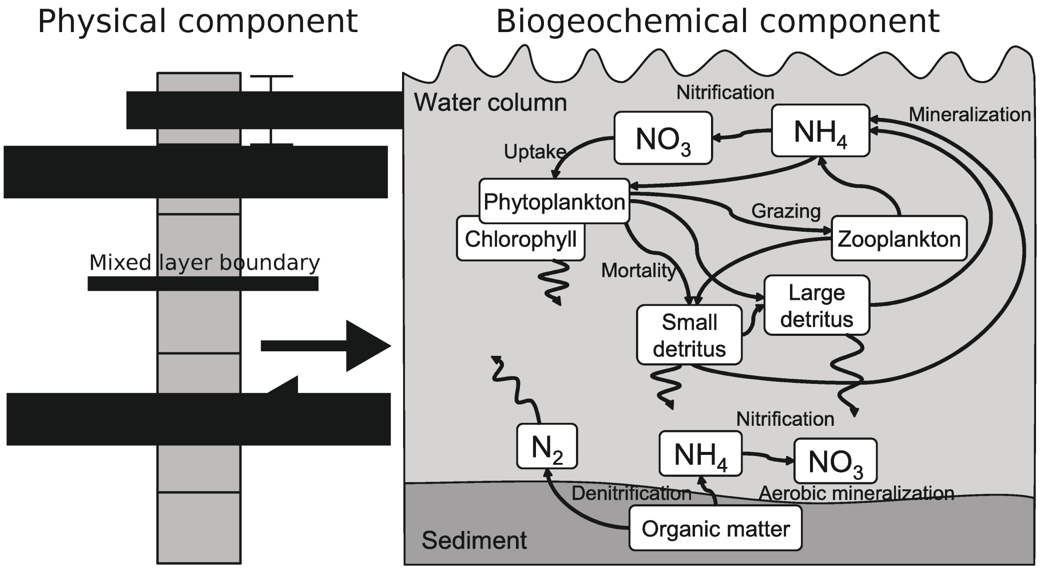


Figure 2.5: Schematic of the physical and biogeochemical (from *Fennel et al.*, 2006) component of the model.

The mixed layer depths were determined from the 5-day snapshots of temperature and salinity in the ROMS model. A mixed layer climatology was used and was calculated by harmonic regression using the first 4 harmonics. The climatological mixed layer depths are shown in Figure 2.6 for station 1 and Figures A.1 and A.6 for stations 2 and 3.

In addition to the imposed mixed layer, the model is forced with incoming solar radiation that attenuates vertically and drives phytoplankton growth but does not affect mixing. The incoming solar radiation for the 3 stations was defined using shortwave radiation estimates from the National Center for Environmental Prediction (NCEP, http://www.ncep.noaa.gov/) and is identical to that used in ROMS. It is important to note that NCEP estimates are daily-integrated values (i.e. without diurnal cycle), while ROMS imposes a diurnal cycle internally within its biological module by re-scaling the daily integral of incoming solar radiation according to astronomical formulae. The time step of the ROMS model is about a minute, which ensures that the diurnal cycle is resolved very well. The simple 1D model has a much longer time step of 6 hours, which would lead to a poor resolution of the diurnal cycle. Hence, two time steps were simply designated as night (setting incoming solar radiation to zero) and the daily-integrated solar radiation is distributed over the other two time steps (designated as day). This ensures that the simple 1D model receives that same daily integral of solar radiation as the ROMS model. Again, a climatology of solar radiation was produced by harmonic regression.

Starting from defined initial conditions for the biogeochemical variables (taken from the ROMS model) the simple 1D model is integrated forward in time by consecutively applying a vertical mixing step and a biological update step. Vertical mixing is calculated using the Crank-Nicolson scheme. Biological sources and sinks are calculated implicitly to prevent the model from overshooting into physically unrealistic negative concentrations during biological updates.

Here the model is run with a vertical resolution of 5 m and a time step of 6 hours, which makes it computationally efficient, and applied to three locations that lie within the domain of a 3D, physical-biological model for the northwestern North Atlantic shelves (Figure 2.4). This allows the use of the simulated physical and biological fields from the 3D model to drive the simple 1D model described above (using simulated mixed layer depths and water temperatures).

In initial 1D simulations, it became clear that the choice of vertical diffusivities critically determines the vertical structure of the biogeochemical variables. A range of diffusivities was tested by applying the physical model component to a temperature variable, which was initialized with the ROMS profile and clamped to the ROMS climatology at the surface and bottom. These tests indicated that vertical diffusivities should vary seasonally, presumably because of changes in stratification and wind mixing. A simple way to parameterize this effect is to allow the diffusivity in each layer to vary with mixed layer depth (mld) as follows:

$$\eta = (1 - q)\eta_{winter} + q\eta_{summer} \tag{2.3}$$

where q is the normalized mixed layer depth: $q = (mld_{max} - mld)/(mld_{max} - mld_{min})$ and mld_{min} and mld_{max} are the minimum and maximum values of the climatological mixed layer, respectively. Values of η_{winter}=70 and η_{summer}=10 m^2 day^{-1} were chosen for the upper layer and were reduced by a factor of 10 for the lower layer, because they resulted in simulated temperature evolutions that matched the ROMS temperature closest.

The 1D models were initialized on 1 January from the synthetic observations, forced with the climatological mixed layer depths and solar radiation and integrated for 15 years until a cyclic steady state was achieved. The final year of simulation is shown in Figures 2.6, A.1 and A.6 (right panels). Important features of a North Atlantic bloom cycle are captured by the model. For example, depletion of nutrients is observed around early April. This is followed by the peaking of the concentration of phytoplankton by mid-April. Sufficient sunlight at this time coupled with sufficient nutrients provide an ideal condition that accelerates the growth of phytoplankton. Similarly, the presence of phytoplankton supports zooplankton growth. The phytoplankton bloom depletes because most of the nutrients are already used and because of zooplankton grazing. Starting at the time of the phytoplankton bloom, an increase in concentration of detritus can be observed as well. This is because of the contributions from the loss terms in phytoplankton and zooplankton i.e. mortality and aggregation in phytoplankton, and mortality and sloppy feeding in zooplankton. The simple model is however limited in describing the synthetic observations because 3D physical processes cannot be represented with the use of a 1D model (Figure 2.6a versus 2.6c). These leads to drifts and biases in the concentration of the biogeochemical variables.

There are significant systematic differences between the synthetic observations from ROMS and the evolution simulated by the simple 1D model (Figure 2.6a versus 2.6c). Most notably, simulated nitrate concentrations are much higher at depth compared to the synthetic observations, and zooplankton concentrations are much lower compared to the synthetic observations. Furthermore, the peak concentration of the phytoplankton spring bloom is much lower in the simulation and the concentrations of detritus are smaller as well. This simple 1D model is thus a good test case for a model that could be improved by nudging.

2.3.3 Hybrid frequency-dependent nudging of the 1D model

The simple 1D model is nudged to the climatology (Figure 2.6b) to correct for the biases observed in Figure 2.6c. The climatology is obtained using a similar procedure used to calculate the synthetic observation, however only the mean and annual cycle are used in the reconstruction. In this study, the 7 biogeochemical variables were used to nudge the simple 1D model.

Initial simulations reveal that use of equation 2.2 in nudging introduces unrealistic variabilities which cause the model to eventually blow up. To get around this problem, a hybrid of conventional and frequency-dependent nudging is introduced as follows:

$$x_t = x_t^f + \gamma[(1-\delta)\langle x_t^c - x_t^f\rangle + \delta(x_t^c - x_t^f)] \tag{2.4}$$

where $\langle\cdot\rangle$ means that the nudges are filtered in time and are only applied at climatological

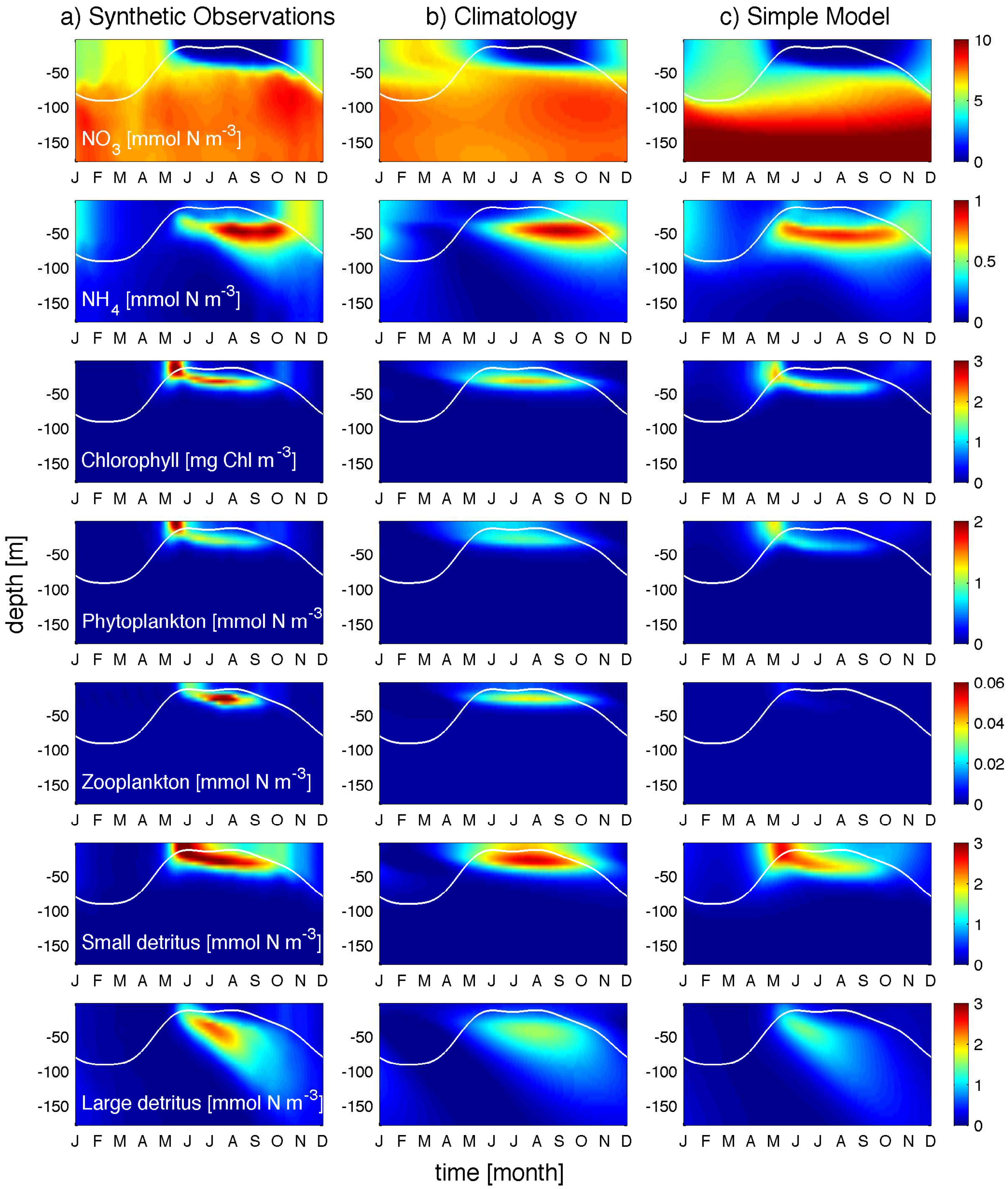


Figure 2.6: Seasonal cycle of the biogeochemical variables for the synthetic observations and climatology, and the result from the simple 1D model for station 1. a) Output from the ROMS simulation of *Bianucci et.al.* (in prep) is regressed using 15 harmonics. This is the benchmark against which all other runs will be compared. b) Similar to (a) however only the mean and 1 harmonic (annual cycle) are used in the regression. c) Output from the simple 1D-model. The biogeochemical component is similar to the 3D ROMS model, the physical processes are kept simple however. The physics in the model is governed only by a difference in vertical diffusivity coefficient below and above the mixed layer.

frequencies (mean and annual cycle). In the hybrid form, weak conventional nudging is applied in the background to suppress the noise that are otherwise amplified due to the nonlinear interaction between the model and frequency-dependent filter. The strength of the additional conventional nudge is controlled by the stability parameter, δ, which I took to be 0.1. A main advantage of the hybrid nudging over conventional nudging is that it allows high frequency variability to evolve almost prognostically while nudging to the observed climatology. The hybrid is implemented sequentially, and thus observations can be assimilated after the suppression of bias and drift (refer to Figure 1.1).

In the present implementation, the bandwidth of the filter is chosen by setting dt/(4×365). This is equivalent to a spin up time of 4 years. Details of how the nudging parameters are chosen are discussed later.

2.3.3.1 Measuring model fit

To quantify the error in the predictions of the synthetic observations, let o_{ij} denote the synthetic observations for a given variable at depth i and time j, and let p_{ij} denote the corresponding prediction. The following scalar quantity is used to measure the difference between the observed and predicted seasonal cycles:

$$MSE = \sum_{i=1}^{n_d} \sum_{j=1}^{n_t} \left(o_{ij} - p_{ij}\right)^2 / \left(n_d n_t\right) \tag{2.5}$$

where n_d is the number of depth levels and n_t is the number of time steps in a year. Thus, MSE is a scalar, defined for each of the seven variables, that measures how well the synthetic observations of a given variable can be predicted.

2.3.3.2 Choice of the nudging parameter

Specifying the correct value of the nudging strength is critical in obtaining a good model state estimate (as seen in Figure 2.3) . Both model and observations contain errors and must be weighted accordingly in the assimilation step. In Kalman filters, this is captured in K_t that dynamically takes into account the changes in covariances between variables at each time step. The method is less clear however for a static, single valued nudging parameter. The approach taken in this thesis is to determine the nudging parameter that gives the best agreement between the synthetic observations and the nudged runs. A sensitivity analysis of the model state to the nudging parameter is carried out to find optimal values of the nudging parameter.

The impact of the conventional and hybrid frequency-dependent nudging is shown in Figure 2.7 which shows the MSE for each variable as a function of the nudging coefficient, γ. The MSEs for each variable have been normalized by the MSE calculated from the climatology (i.e., by the MSE of the conventionally nudged run with γ=1). The fit of the simple model improves with nudging for all variables with increasing γ but the outcome slightly degrades again once γ increases from the optimum values toward 1. Results from the hybrid are generally better (lower values of MSE) at estimating the synthetic observations than results from conventional nudging.

No single optimal nudging strength however is found across all seven variables but rather a range of qualified values that will give the nudged results a significant improvement over the simple model. Figure 2.7 suggests further that a larger range of the nudging parameter is available to the hybrid than to conventional nudging. Values of γ_{con}=0.02 and γ_{spec}=0.025 were used for conventional and hybrid frequency-dependent nudging, respectively.

2.3.3.3 Nudged runs

Nudging results in Figure 2.8b and 2.8c show considerable improvement over the simple 1D model (Figure 2.6c) (see table 2.1 as well). The concentration of NO_3 through out the water column as well as other variables now closely resembles that of the synthetic observations. Nudging improved the supply of nutrients throughout the whole water column, corrected the concentration of phytoplankton and zooplankton, and improved the vertical distribution of detrital deposits. Nudging permitted the model to overcome the shortcomings of using simple physics and diurnal cycle. Results from the hybrid frequency-dependent nudging, however predict the synthetic observations better than conventionally nudged results. Results from Figure 2.8 show that concentrations of the 7 biogeochemical variables are in closer agreement to the synthetic observations than the conventionally nudged results. Table 2.1 shows that for most of the variables (except NH_4 for station 1 and 2), the calculated mean squared error (MSE) for the hybrid frequency-dependent nudging is lower than those from conventional nudging. This is not surprising since more frequency is allowed to pass through in hybrid frequency-dependent nudging, thus approximating the synthetic observations better. This can be seen well in Figures 2.9 and 2.10 where time series of the surface concentration and at 30 m for different runs are compared. Conventionally nudged results generally follow the shape of the climatology

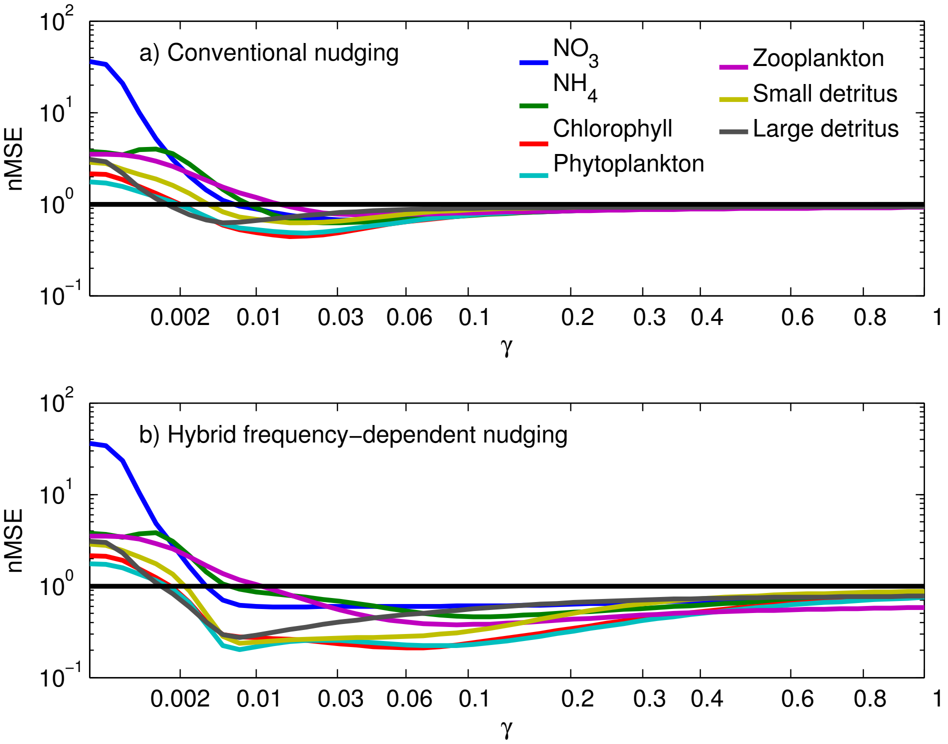


Figure 2.7: MSE as a function of the nudging coefficient γ for the conventional and hybrid frequency-dependent nudging for station 1. The values have been normalized ($nMSE$) by the MSE values for climatology (MSE_{SC}). Each of the seven lines corresponds to a specific biogeochemical variable. The x-plotting positions were stretched to vary as γ^3 in order to expand the plot near γ=0.

Table 2.1: Mean squared errors (MSE) from the difference of the synthetic observation and the different runs considered. MSE is calculated for the whole water column and each biogeochemical variables for the three stations. Cases where MSE from conventional nudging is higher than hybrid frequency-dependent nudging is highlighted in grey.

	Runs	NO_3	NH_4	Chl	Phy	Zoo	$SDet$	$LDet$
Station 1	Prognostic	24.48	4.3	1.60	1.42	3.29	2.26	1.61
	Conventional	0.74	0.65	0.45	0.49	0.82	0.62	0.78
	Hybrid frequency-dependent	0.62	0.70	0.24	0.25	0.58	0.27	0.41
Station 2	Prognostic	25.45	5.80	1.76	1.63	3.31	2.55	1.92
	Conventional	0.98	0.79	0.60	0.63	0.78	0.70	0.76
	Hybrid frequency-dependent	0.75	0.90	0.17	0.16	0.37	0.22	0.33
Station 3	Prognostic	3.31	2.16	0.56	0.66	1.40	0.86	0.33
	Conventional	0.71	0.93	0.50	0.51	0.33	0.70	0.73
	Hybrid frequency-dependent	0.38	0.51	0.22	0.20	0.11	0.24	0.18

Units: NO_3 [mmol N m^{-3}], NH_4 [mmol N m^{-3}], Chl [mg Chl m^{-3}], Phy [mmol N m^{-3}], Zoo [mmol N m^{-3}], $SDet$ [mmol N m^{-3}] and $LDet$ [mmol N m^{-3}]

(thus, suppressing other frequencies). Consider for example the phytoplankton; although both nudging methods correctly predict the timing of the peak, the seasonal cycle from conventional nudging is much wider and centered at May. In contrast, the seasonal cycle of phytoplankton from hybrid captures the correct shape of the synthetic observation indicating more frequency was used for its prediction. The extra frequencies allowed in the hybrid resulted in prediction of the correct timing of nutrient depletion and consequently plankton blooms and detrital deposits. Similar results are also obtained for stations 2 and 3 (see Figure A.1-A.10).

2.4 Summary and Conclusions

A computationally efficient and effective scheme for suppressing bias and drift in a 1D biogeochemical model was evaluated. The basic idea behind this scheme is to nudge the model toward climatology in prescribed frequency bands. Outside of these bands the models dynamics are unaffected by the nudging and the model can evolve prognostically.

To assess the scheme, results from hybrid frequency-dependent nudging are compared against data taken from 3 locations within a 3D physical-biogeochemical model for the

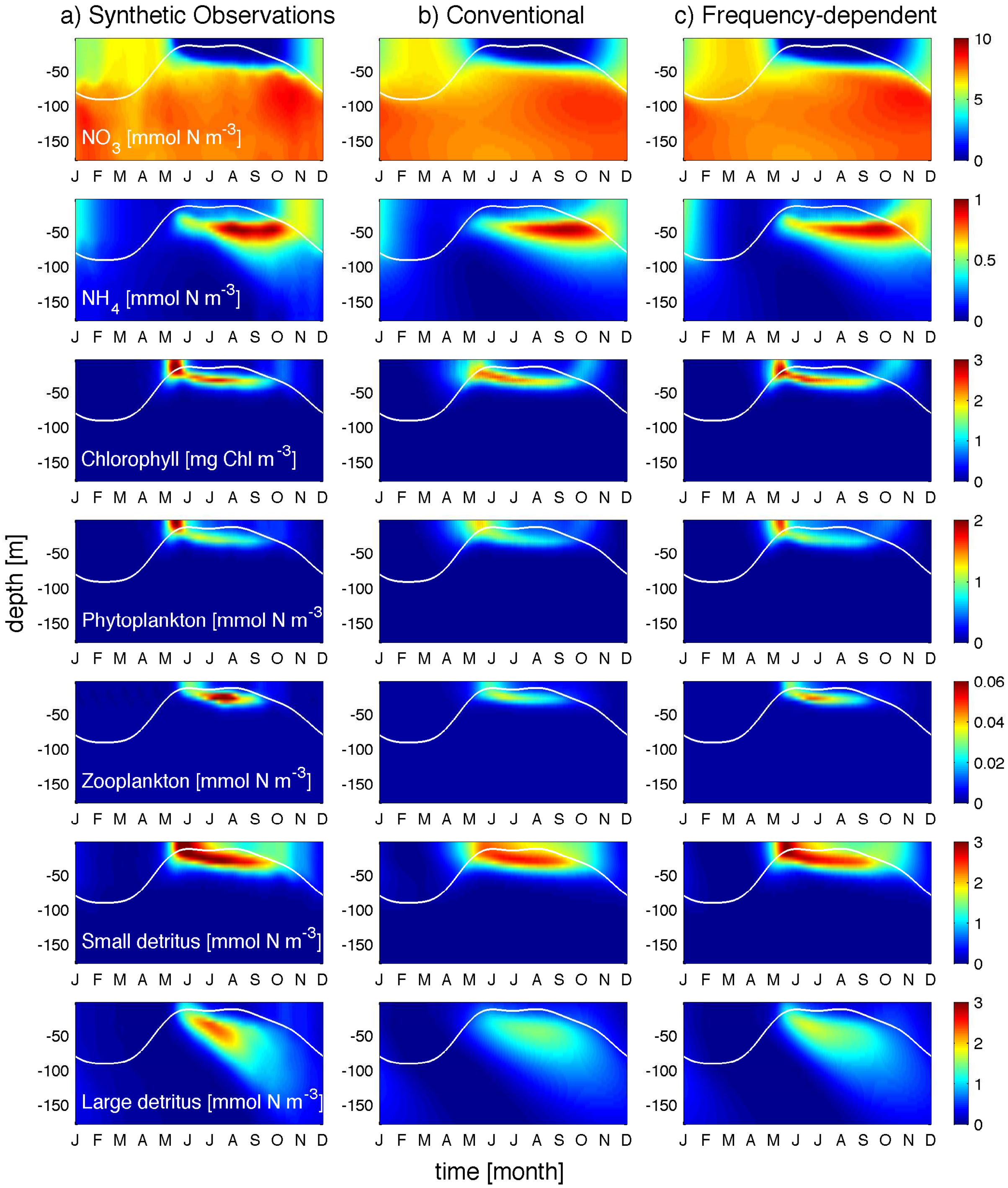


Figure 2.8: Seasonal cycle of the biogeochemical variables for the synthetic observations, and the result from conventional and hybrid frequency-dependent nudging for station 1. a) Similar to Figure 2.6a. b) Output from conventional nudging using γ=0.020 (equivalent to a nudging strength of 12.5 days). c) Output from hybrid frequency-dependent nudging using γ=0.025 (equivalent to a nudging strength of 10 days).

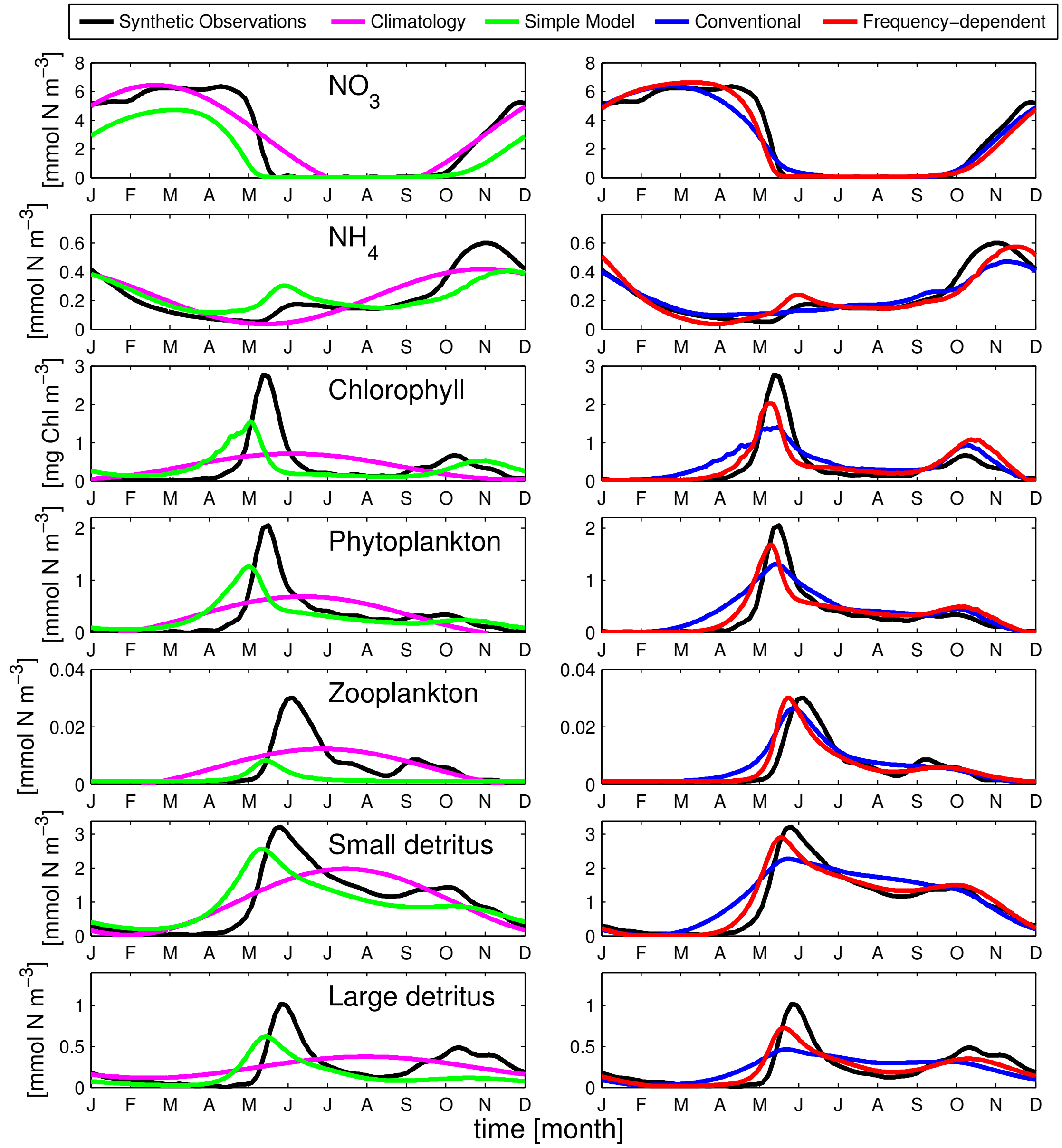


Figure 2.9: Comparison of the seasonal cycle of the different runs at the surface. The lines show the daily average of the concentration the biological variables. Comparison of the synthetic observation (black line) with climatology (magenta line) and simple 1D model (green line) are shown on the left. Comparison of the synthetic observation with results from conventional (blue line) and hybrid frequency-dependent nudging (red line) are shown on the right.

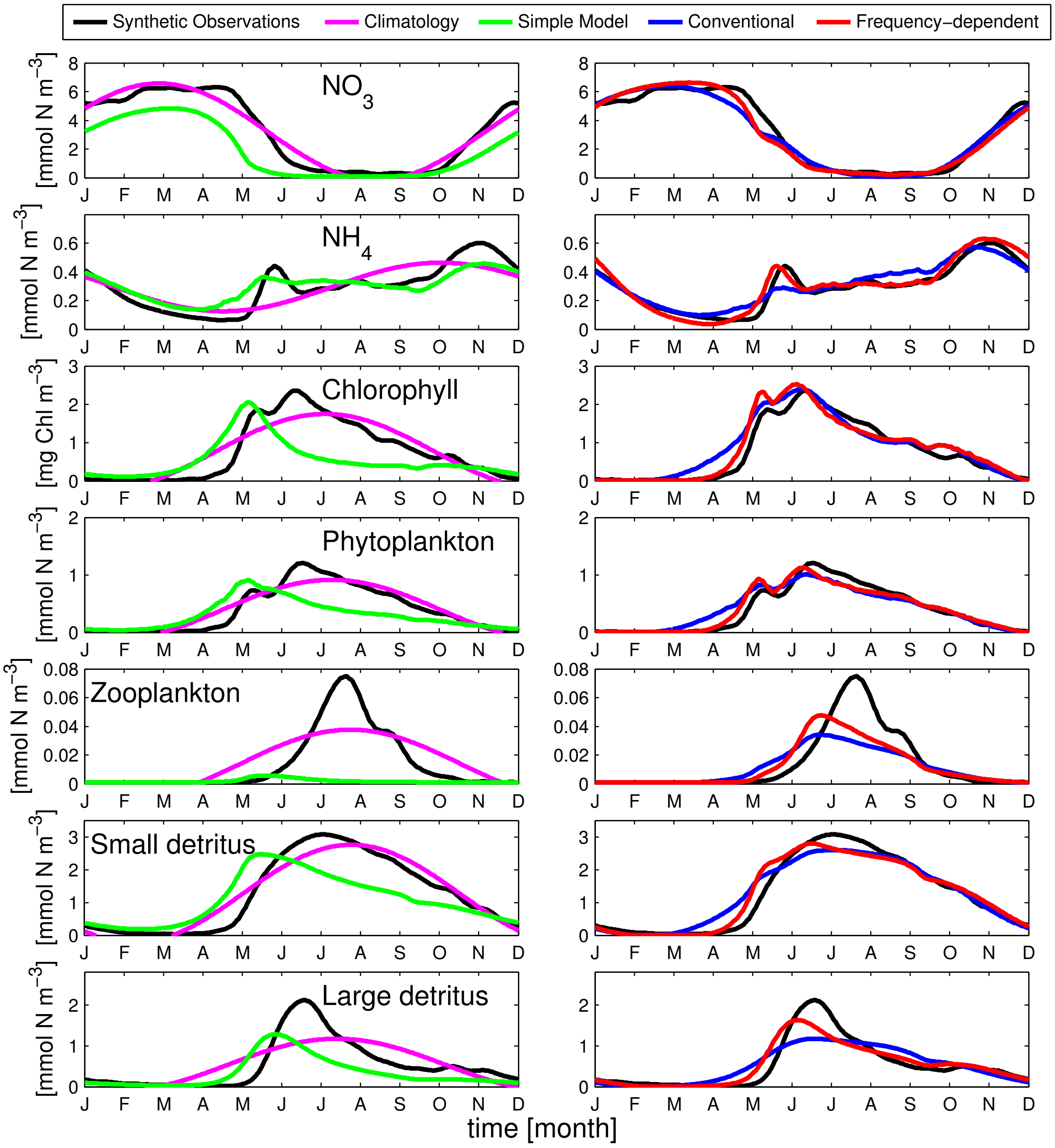


Figure 2.10: Similar to Figure 2.9 but time series taken at a depth of 30 m.

Northwestern North Atlantic, and compared with results from conventional nudging. Overall results are encouraging: although both nudging schemes improve biogeochemical estimation as compared to the simple 1D model, results from the hybrid frequency-dependent nudging perform better since they can capture the main features of biogeochemical dynamics. Specifically, the timing of NO_3 depletion, and phytoplankton and zooplankton blooms are well represented as compared to results from conventional nudging. The duration of the bloom is reasonable as well ($\backsim$ 2 weeks); the better fit is due to more harmonics being allowed in hybrid frequency-dependent nudging. In contrast, conventionally nudged runs overestimate the early NO_3 depletion, and phytoplankton and zooplankton blooms. This affects the timing of the dynamics of other variables as well. The overestimation is caused by the strong nudging ($\backsim$12.5 days) which suppresses higher frequency (more than 1 cycle per year) in the model dynamics; conventional nudging "forces" the model to follow the shape of the climatology.

Although overall results from hybrid frequency-dependent nudging are better, it is worth noting that there are some discrepancies when compared to the synthetic observations. Figures 2.9 and 2.10 show that concentrations of phytoplankton and zooplankton blooms are not as peaked as the synthetic observations. The dampening effect of conventional nudging in the background might have contributed to this result. There is a slight delay in the timing of the blooms as well ($\backsim$ 1 day). This suggests that the model might benefit from increased nudging strength, γ. This is subtle however as no optimum γ exist for all the variables. Increase in γ increases the effect of conventional nudging to hybrid frequency-dependent nudging as well, which can affect how the hybrid method handles high frequencies. Another possible reason for the discrepancies observed could be due to the handling of allowed frequencies in a discretized model. Although in theory, frequency-dependent nudging allows all frequencies (other than mean and annual cycle) to evolve uninhibited, limitations in numerical resolution (finite time duration) might cap the maximum number of frequencies allowed (this exist in other frequency based methods as well e.g. discrete fourier transform). This can cause a slight underestimation of the system being approximated.

Thompson et al. (2006) pointed out an important caveat when using the scheme. Although hybrid frequency-dependent nudging can eliminate bias and drifts, it does so at the expense of adding artificial sources and sinks of biogeochemical variables. This suggests that the model and its forcing are incorrect in some fundamental ways, especially in a 1D model where physical dynamics are clearly lacking. It was suggested that nudges may be useful instead in diagnosing problems with the model, which when fixed is likely to lead to a more realistic biogeochemical model. This is in agreement with the idea expressed by *Nerger and Gregg* (2008) who estimated their model bias by the assimilation of satellite ocean chlorophyll data into a global model.

CHAPTER 3

DEVELOPMENT OF A BIOMASS PROXY FROM ELECTRONIC TAG DATA

As described in Figure 1.1, observations can be assimilated after the bias correction step. Observations can either come from satellite measurements, ship-based sampling and *in situ* autonomous platforms and moorings. In this chapter, the potential of using observations (with focus on estimation of phytoplankton biomass) from electronic tags attached to marine animals will be assessed. These are potentially very good sources of oceanographic data due to their wide spatial coverage and high temporal resolution.

Assessment of physical quantities measured like temperature, depth and irradiance are straightforward since these variable are directly measured by the electronic tags. This is not the case for biogeochemical variables however since these quantities are not measured. An indirect relationship between phytoplankton biomass and changes in physical quantities like light attenuation needs to be established first to be able to assess biogeochemical quantities.

To cover the important concepts that are needed for this chapter, the next sections will be presented as follows: In section 3.1, the behaviour of light under seawater and its relation to other materials present in seawater is discussed. The relationship between light attenuation and phytoplankton biomass, which is the central theory that is important for this chapter, is discussed as well. Furthermore, the use of attenuation coefficient of downwelling irradiance, K_d, to derive bio-optical relationship in different water bodies is described and justified. In section 3.2, the pattern of the Grey seals' (*Halichoerus grypus*) movement is described. This is important as it provides insights to the spatio-temporal pattern uncovered in electronic tags data along Scotian Shelf. The data obtained from electronic tags are also described in this section. Since the electronic tags are not originally intended for quantitative measurement of irradiance but rather to geo-locate seals by determining dawn and dusk (personal communication with Heather Baer, support coordinator Wildlife Computer), a calibration experiment at Bedford Basin is discussed in section 3.3, including comparison of the estimated $K_d(550)$ to potential indicators of phytoplankton biomass (i.e. chl, fluorescence, absorption, attenuation). In section 3.4, the results of estimation of attenuation coefficient, $K_d(550)$, and the bio-optical relationship along Scotian Shelf is presented. Finally, a summary and conclusion for the chapter is given in section 3.5.

3.1 Light in sea water

Incident light from the sun progressively attenuates in seawater. It is influenced by the scattering and absorption properties of seawater itself and the various substances present in it. The properties only depend on the content of the water, and thus are referred to as inherent optical properties (IOPs; *Kirk*, 1994).

The reduction of light energy entering the ocean is mostly due to absorption by water molecules. This is not surprising since for every 100 H_2O molecules, there are only 3-4 molecules of other substances (*Wozniak and Dera*, 2007). Water itself is highly absorbing at wavelengths below 250 nm and above 700 nm (*Mobley*, 2010), where the absorbed energy is converted to non-radiant form and given off as heat. As more photons are absorbed with depth, less light penetrates deeper down the water column.

Scattering is associated with the particles present in water. Scattering causes light to change its direction of propagation when it encounters an obstacle. Incident light coming from the sun is reduced before it enters the seawater because part of it is reflected back to the atmosphere. Once light enters the water column, particles scatter it in all directions with the majority being scattered in the forward direction. Light can either be reflected or refracted at the interface of two media. Light can also be diffracted. The mere proximity of the particle to a light beam's path can cause the light to change direction because light waves bend around corners (*Davies-Colley et al.*, 1993). Furthermore, scattering also contributes to further absorption of the light because changing directions increases the path length that the light must travel, thus the probability of being absorbed increases as well (*Gallegos*, 2001).

IOPs give us information about the type and concentration of water constituents. Most of the bulk properties of a water body can be deduced from measurements of absorption coefficients and volume scattering functions. However, the difficulty associated with measurement of IOPs in the early days of optical oceanography led to the use of apparent optical properties (AOPs, *Mobley* (2010)). *Mobley* (2010) defines AOPs as (i) properties that depend both on the medium (IOPs) and on the directional structure (spatial distribution) of the radiance distribution, and that (ii) display enough regular features and stability to be useful descriptors of a water body. Radiometric variables such as upwelling, E_u, and downwelling irradiances, E_d, are commonly used measurements to derive AOPs such as irradiance reflectance and various irradiance attenuation functions.

The diffuse attenuation coefficient (K_d) is one of the most common AOPs being used. It provides a direct measure of the penetration of light in the ocean since absorption and scattering effects are assumed to be captured by light attenuation with depth. K_d depends on both the IOPs and spatial distribution of the radiance distribution which satisfies the definition of AOP, and it remains the same under similar water conditions (i.e. displays enough regular features and stability) despite the sensitivity of the measured E_d to varying environmental conditions. This can be seen in the equation below:

$$K_d = -\frac{1}{z_2 - z_1} ln\left(\frac{E_{d_2}}{E_{d_1}}\right) \tag{3.1}$$

where E_{d_1} and E_{d_2} are irradiances measured at depths z_1 and z_2, respectively (*Kirk*, 1994). In the equation above, if E_{d_1} suddenly changes because of a changing angle of the incident light, presence of clouds in the sky or surface waves, the effect will be cancelled out since E_{d_2} will experience a proportional change thus leaving K_d unchanged.

Estimates of K_d alone can give enough information for water mass classifications. High K_d means that there are abundant attenuating agents present in the water, and thus the water is more turbid. Coastal regions are usually associated with higher K_d as compared to open ocean waters (*Mobley*, 2010).

Classification of water masses into two types were first introduced by *Morel and Prieur* (1977). Spectral reflectance and attenuation functions were used to group water bodies as case I and case II. Water bodies where contributions from phytoplankton control the optical properties are classified as case I waters. In this scenario, the measured absorption and scattering due to phytoplankton is high compared to other substances.

Case I waters can range from very clear (oligotrophic) to very productive (eutrophic) water, depending on the phytoplankton concentration. Case II waters are those where contributions from phytoplankton and other substances influence the optical property of water. These substances include coloured dissolved organic matter (CDOM), dissolved organics, and various suspended sediments. Case II waters are generally found in coastal zones influenced by land (*Mobley*, 2010).

To relate the estimated AOP (K_d) to the biological constituents present, *Morel* (1988) proposed a bio-optical model, which allows the optical properties and the solar radiation propagation within the ocean to be predicted from the phytoplanktonic pigment content. Case I waters were considered since chlorophyll (chl) retrieval is relatively straightforward because only phytoplankton influence the optical properties. chl concentrations range from very low values ($\backsim 10^{-1}\ mg\ m^{-3}$) in oligotrophic waters to high values in coastal upwelling areas ($\backsim 10^{1}\ mg\ m^{-3}$). Regions with high chl concentrations are associated with high K_d estimates while measurements from oligotrophic waters are associated with low K_d estimates. Higher phytoplankton concentration means that more phytoplankton-associated materials (i.e. debris and excreted organic matter) are present as well, which contribute further to absorption and scattering of light particles. A relationship between phytoplankton biomass (and its associated materials) and attenuation coefficient for case I waters, first deduced by *Morel* (1988) and confirmed again by *Morel and Maritorena* (2001), was shown to follow the general power law form

$$K_d(\lambda) = K_w(\lambda) + \chi_e(\lambda) \cdot chl^{e(\lambda)} \tag{3.2}$$

where K_w is the wavelength-specific diffuse attenuation of downward irradiance due to pure sea water, and χ_e and e are wavelength-specific fitted constants.

Bio-optical relations for case II waters are more complicated because of the presence of other substances whose optical signatures are independent from those of chl. *Morel and Maritorena* (2001) point out that bio-optical models based on chl for case II waters are insufficient and must be supplemented by accounting for the presence of suspended solids and colored dissolved organic matter (CDOM), not co-varying with the algal biomass. This is necessary since CDOM contributes to absorption of light while suspended solids contribute to scattering and further absorption of light (*Lewis*, 2011).

In case I water, chl is commonly used as an index to quantify algal content, and more generally the bio-optical state of the region. The relationship between K_d and chl given by *Morel* (1988) in equation 3.2 captures the bio-optical dynamics in bodies of water that are far from land influence. Results using K_d as proxy for phytoplankton biomass for case I waters yield generally robust results in estimating chl concentration (*Antoine et al.*, 2005; *Teo et al.*, 2009). Bio-optical models for case II waters are different as they often involve complex relations that give considerable weight to other substances.

In coastal regions like the Gulf of Maine, the New England continental shelf and the Scotian Shelf, optical conditions generally resemble that of case II waters. Surface water in these regions is case II due to particles, CDOM and dissolved and detrital materials (*Balch et al.*, 2004; *Green et al.*, 2003). These substances have significant absorption and scattering properties. During fall and spring however (possible bloom periods), pigment material dominates the optical environment and hence permits the use of pigment algorithms (*Topliss et al.*, 1991). In addition, *Topliss et al.* (1991) point out that the long-term trend in turbidity in Scotian Shelf waters is related to the general level of biological material present.

3.2 Seal tag data

3.2.1 Spatio-temporal pattern of the grey seals' movement

Considerable interest has been given to understanding the spatial and temporal pattern of grey seals' (*Halichoerus grypus*) foraging behaviour because of increasing evidence that marine mammals can have significant top-down effect on ecosystem functioning (*Bowen*, 1997). *Breed et al.* (2009) suggested that patterns in foraging behavior of grey seals could be explained by contributions from demographic, environmental, and seasonal factors. Depth (likely related to abundance of food source) is the number one factor that seals consider in foraging. Grey seals are shelf animals that forage on shallow banks to conserve energy, and acquire food (*Harvey et al.*, 2012).

Effects of sex differences were observed to be important in determining the diving behaviour with males diving 5-15 m deeper than females. Males dive consistently throughout the day, whereas females showed strong diurnal patterns in dive depth, duration and frequency. Females however had longer dives and spent more time at depth. The pattern of foraging of female grey seals is consistent with early accumulation of body energy to support pregnancy and the subsequent lactation period during which females fast, while

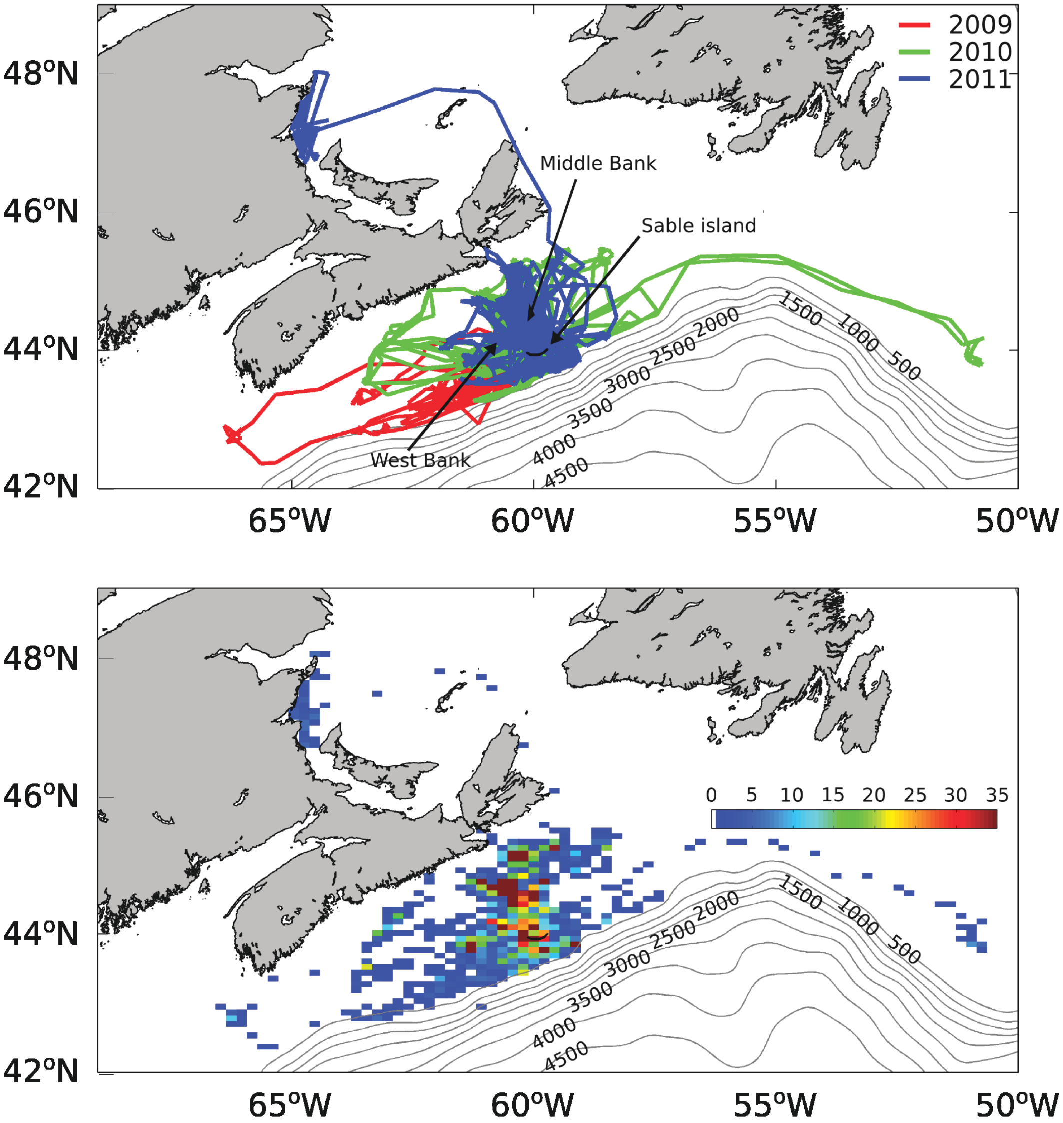


Figure 3.1: Spatio-temporal patterns in grey seals (*Halichoerus grypus*) movement. a) Tracks of 45 grey seals for the years 2009-2011. The legend shows the year in which electronic tags were deployed. For the red tracks, electronic tags were deployed in October 2009 and recovered from December 2009-February 2010; (green tracks) deployed on September 2010 and recovered from December 2010-January 2011; (blue tracks) deployed on June 2011 and recovered from December 2011-January 2012. b) Mean daily position density of 45 grey seals for the years 2009-2012. Each bin is approximately 225 km^2. The gray lines show the bathymetry.

males suggest a gradual accumulation of body energy (*Beck et al.*, 2003).

Diving patterns differ as well seasonally with both sexes foraging in deeper waters in the winter and shallower water in summer. Deeper dives require more energy and leave less stored body oxygen and are therefore less efficient (*Thompson and Fedak*, 2001); moving to deeper waters in winter suggests that prey are to be found in deeper waters. The scattered nature of foraging locations were also observed in winter suggesting that prey are less predictable in winter as well. On the contrary, shallower water in the summer and fall is more accessible to prey and more focused patterns of foraging are observed (*Breed et al.*, 2009). Several studies support this observation and points out that in the western Atlantic, many fish species migrate to deeper water during winter and to shallow banks during summer and fall to remain in warmer water (*Perry and Smith*, 1994; *Swain et al.*, 1998). In summer and fall, spatial patterns of foraging are observed to be focused in the Middle Banks (see Figure 3.1). Both sexes remain inshore near haulout sites in Sable Island and spend a larger proportion of their time ashore. Few foraging trips were observed as well. Despite the reduction in foraging effort, increased body mass were observed in both sexes (*Beck et al.*, 2003). This, and the intense use of Middle Banks and surrounding areas, suggest a highly predictable prey resource over this bank in the summer and fall (*Breed et al.*, 2009). The spatio-temporal patterns associated with grey seals movement are strongly influenced by bathymetry and foraging effort focused in shallower areas of the shelf that changes dynamically through the year and between sexes.

3.2.2 Data from electronic tags attached to grey seals

Electronic tags (MK10 loggers from Wildlife Computer, USA) were deployed by the group of S. Iverson (http://fatlab.biology.dal.ca/) and D. Bowen (http://bowenlab.biology.dal.ca/) in three different months in the post-moulting (moulting ends around May) periods for 3 different years; October 2009, September 2010 and June 2011. The seals were captured on Sable Island using hand-held nets and anaesthetized (see *Bowen et al.* (1992, 1999) for details). Once immobilized, electronic tags were glued to the pelage at the back of the animal using epoxy resin. In October 2009, 15 electronic tags were attached to Grey seals (7 males, 8 females) and only 13 were recovered (5 males, 8 females) in the pre-breeding periods December 2009-February 2010. Deployment and recovery of electronic tags were all done on Sable Island. In September 2010, 20 electronic tags were attached to Grey seals (6 males, 14 females) and were all recovered between December 2010-January 2011.

Data from sealID 5954 (male seal, see table 1 in appendix) is dropped however because the GPS file is inaccessible and might be corrupted. In June 2011, 20 electronic tags were attached to Grey seals (20 females) and only 16 were recovered in December 2011-January 2012. Only 13 data sources were used however since the file from sealID 6120 is not available and GPS data from sealID 137 and 10331 are corrupted (see table 1 in appendix).

The electronic tags record depth, temperature, light level and GPS locations. The depth and temperature sensors were calibrated to provide an accuracy of 1%. Depth and light level are temperature-compensated to provide consistent readings through temperature variations (*Wildlife Computer USA*, 2012). The electronic tags were set to sample depth (0 to 1000 m $\pm$ 0.5 m), water temperature (-40°C to +60°C $\pm$ 0.05°C), and light level every 10 seconds. Light level is measured as irradiance at a wavelength of 550 nm with a logarithmic range from 5×10^{-12} W$\cdot$ cm^{-2} to 5×10^{-2} W$\cdot$ cm^{-2}. Sensor measured light level are converted to irradiance (W$\cdot$ m^{-2}) using the relation provided by Wildlife Computer (see Figure B.1 in the appendix). The electronic tags are not intended for quantitative measurement of irradiance but rather to geo-locate seals by determining dawn and dusk (personal communication with Heather Baer, support coordinator Wildlife Computer). Calibration experiments will however be discussed in section 3.3 to show that measured irradiance from electronic tags are comparable with measurements from specialized irradiance sensors.

Figure 3.2 shows a time series of dive and temperature profiles, and irradiance (W$\cdot$ m^{-2}) for one seal for one day. In this example the seal dove approximately 180 times per day (top). Because of the seals' diving pattern and frequency, and the sensors' sampling rate the temperature profile (middle) of the seal's dive cycles can be reconstructed. The time series of the temperature profile suggests that the well-mixed layer is about 40-50 meters deep, the region where the attenuation coefficient, $K_d(550)$, is calculated. The time series of irradiance (bottom) shows a clear distinction between day and night (dawn and dusk). The fluctuations that are observed in the time series correspond to the seals' diving behaviour.

Pre-processing of the irradiance data reveals that there is a difference between the ascent (blue circles) and descent (green circles) phase of the seal dives in terms of the irradiance measured (Figure 3.3). To calculate the attenuation coefficient from the irradiance profile, the following data were used: (i) from the ascent phase, (ii) between 10:00 to 14:00 hours,

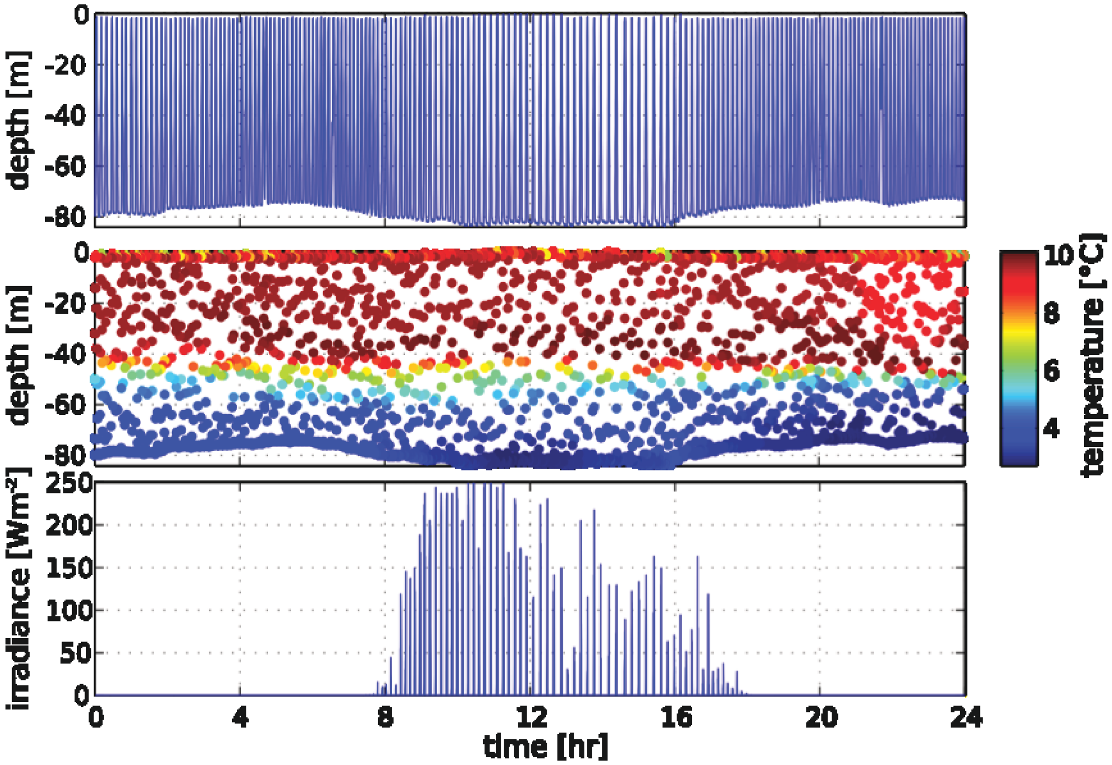


Figure 3.2: Time series of the dive and temperature profiles, and irradiance as measured by the seal tags. Data is taken from measurements in October 22, 2009 of a female grey seal with sealID 4267 at a daily mean position of $43.98°N$, $59.98°W$.

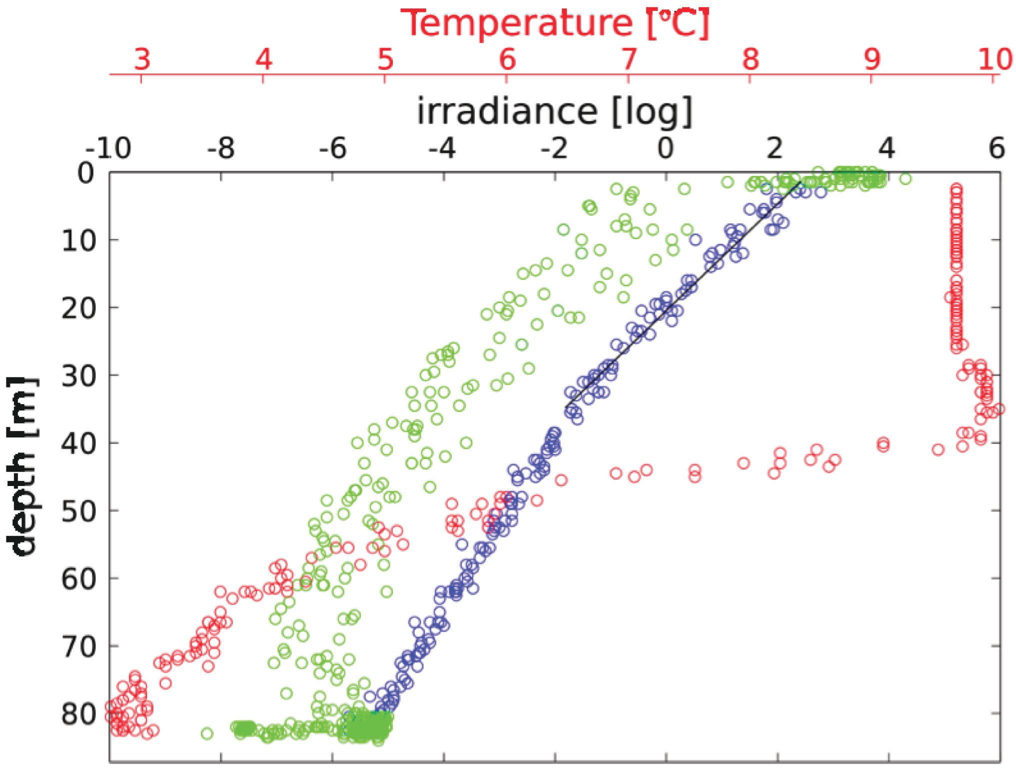


Figure 3.3: Log-irradiance (blue and green circles) and temperature (red circles) profiles for the ascent (blue circles) phase of the seal dives from Figure 3.2. The profiles were plotted using data between 10:00 and 14:00, when the sun is near normal to earths surface. The green circles correspond to irradiance profiles where the seal is in descent phase, is at depths shallower than 5 m (accumulation of points near at depths 0-5 m), and when the seal has minimum vertical displacement (usually happens in the bottom). The green circles were dropped in the estimation of attenuation coefficients since they produce larger errors (points are more spread out) as compared to the estimates from the ascent phase (blue circles). The black line is the regression line in the mixed layer. Measurements from October 22, 2009 of the female grey seal with sealID 4267 at a daily mean position of $43.98°N$, $59.98°W$..

(iii) when seal is in vertical motion, and (iv) deeper than 5 meters. Condition (i) assures that the light sensor is facing upwards, the choice of time in (ii) allows for enough data points to use when sunlight is near normal to the surface, (iii) guarantees that changes in light condition with depth are being captured and (iv) since unrealistic noise are observed near the surface which might be due to surface waves or seal behaviour near the surface (e.g. rolling). From the processed data, the attenuation coefficient is calculated by regression (black line) of the data points in the mixed layer. The temperature profile (red circles) is plotted as well, which sets the limit for regression since the attenuation coefficient is only estimated in the mixed layer (reasons are discussed in section 3.4).

Considering the high frequency capability of tags to collect data, coupled with the high mobility of grey seals and their need for frequent nutrition, biologically active regions could be sampled better both temporally and spatially than through traditional in situ methods. Oceanographic variables in these regions can be sampled directly like temperature, salinity and density. Other variables such as chl and phytoplankton concentration can be sampled indirectly by using a bio-optical relation between light attenuation and water constituents. Recent studies indeed show that electronic tags attached to pelagic animals can be an excellent and reliable oceanographic data sources (*Teo et al.*, 2009; *Boehlert et al.*, 2001).

The data that can be inferred from electronic tags can vary in consistency however. There are days when seals go out to the sea, but do not dive deeper than 10 m (set diving threshold) and thus no useful data can be generated. In some cases, seals move back and forth around Sable Island. All these can contribute to temporal patchiness that can affect the quality of the time series data generated. In October 2009-January 2010, a total of 60% useful data were generated from 13 seals. This increased to 62% from 19 seals from September 2010-February 2011. Data collection from June 2011-January 2012 generated the least useable data set- a total of 43% from 15 seals. Despite having only 55% (from 2009-2012) of the seals' itinerary convert to useful data, the regions sampled are well represented.

3.3 Calibration of electronic tags in Bedford Basin

3.3.1 Comparison of electronic tag data with CTD and the hyper-pro measurements

Absorption of light by different phytoplankton species has been well documented (*Kirkpatrick et al.*, 2000; *Barlow et al.*, 2002; *Ciotti et al.*, 2002). Regardless of the phytoplankton species, absorption spectra are dominated by blue ($\backsim$440 nm) and red ($\backsim$675 nm) wavelengths wherein absorption in the blue region is around 60$\%$ greater than at red. Presence of pigments other than chl will cause broadening peaks and appearance of additional absorption spectra. Absorption at 550 nm however remains very low. This apparently is a source of concern (will be referred to as C1) for this study since the light level sensor operates at 550 nm. Hence, the calculated K_d might not represent the actual biomass present along the Scotian Shelf because of limited information that can be deduced in using 550 nm.

The next concern (will be referred to as C2) in using electronic tags is the reliability of the data collected. According to *Wildlife Computer USA* (2012), the manufacturers of the tag, the light sensors are not designed for quantitative measurement of irradiance. The light sensors are used only to determine dawn and dusk to aid in estimating the geolocation of the seals. This explains why 550 nm is the wavelength of choice since green wavelengths are least absorbed and thus can penetrate the deepest in coastal waters (*Jerlov*, 1976).

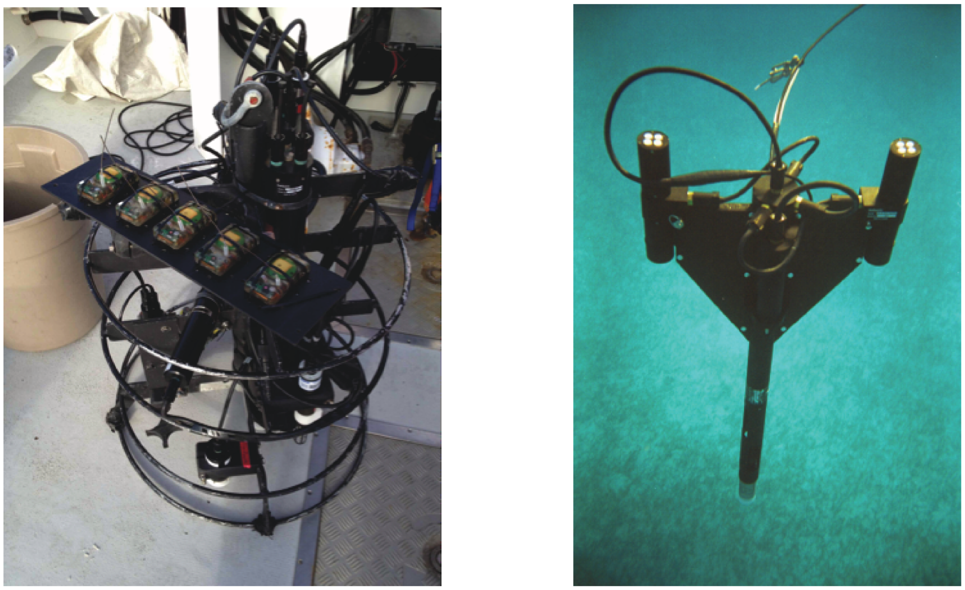
The Bedford Basin experiment was designed to learn more about the potential problems described above and to allow for proper calibration of the seal tags. In C1, the chl estimates derived from $K_d(550)$ from electronic tags are compared to actual chl measurements in the region. The direct comparison should address whether irradiance estimates from electronic tags can be used to estimate phytoplankton biomass. A caveat to this experiment however is that the relationship between $K_d(550)$ and chl that will be derived for Bedford Basin might not necessarily hold true for the Scotian Shelf. Although both waters can be categorized as case II, Bedford Basin is surrounded by land mass and receives abundant contribution of CDOM and suspended particles from the surrounding industrial activities. For C2, direct comparison of the irradiance profiles should suffice to assess the reliability of irradiance sensors. It is also the goal of this experiment to determine if measurements between seal tags are comparable.

In this experiment, the electronic tags along with the conductivity-temperature-depth sensor (CTD) and the hyper-spectral profiler (hyper-pro) were deployed to measure irradiance, depth and temperature in Bedford Basin (Figure 3.4). The electronic tags were mounted on the CTD (Figure 3.4a) and lowered at a station in Bedford Basin every week starting February 29, 2012 by J. Cullen's lab (www.ceotr.ca, I participated twice in the deployment and helped design the calibration experiment). Data were collected at a one second interval and lowered at a speed of 1 ms^{-1}. The hyper-pro was lowered into the water 5-10 minutes after the seal tag-CTD measurements.

Deviations in the measured depth occur among the seal tags and when compared to the CTD (baseline measurement). To calculate this deviation, the temperature profiles from electronic tags and CTD were interpolated first to a vertical resolution of 5 cm. The temperature profiles from electronic tags were then decremented or incremented by 5 cm and compared to the temperature profiles from the CTD. The depth adjustment that gave the minimum squared difference between the temperature profiles from electronic tags and CTD were reported as depth deviations. Visual comparison of the corrected temperature profiles in Figure 3.6 shows that the calculated depth deviations are reasonable.

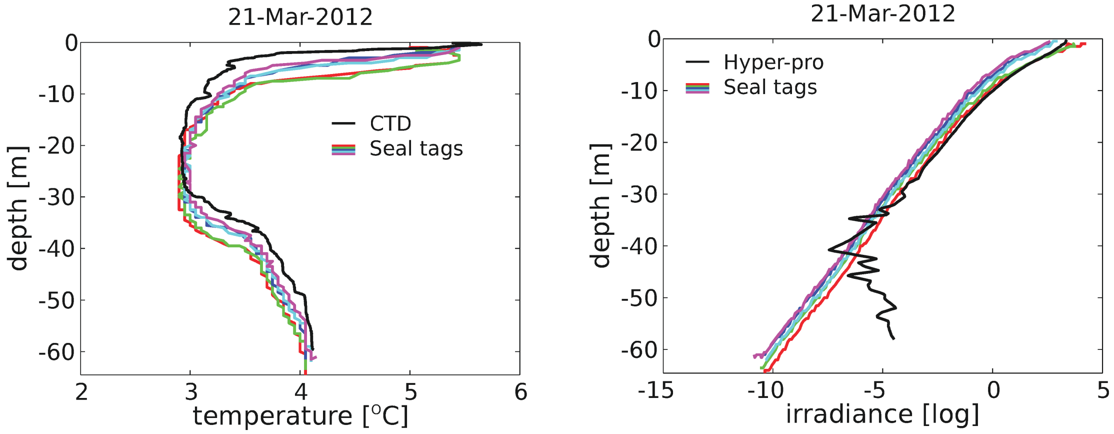
To calculate the depth deviation between hyper-pro measurements and electronic tags data, a similar method to the previous paragraph was used with two important modifications: i) log-irradiance profiles were minimized in place of temperature profiles and ii) a mean of the depth adjusted irradiance profiles from electronic tags was compared to the hyper-pro measurements. The mean of the depth adjusted irradiance profiles from electronic tags was used since the only interest in this calculation is to determine how much the deviation there is to the depth measurements of the hyper-pro. Visual comparison of the corrected irradiance profiles in Figure 3.7 shows that the calculated depth deviations are reasonable.

Figure 3.5 shows the evolution of depth offsets for several sampling days. There appears to be no systematic drift in depth offsets, suggesting a constant systematic error (that can be corrected). Differences in the capability of the pressure transducer (which is used to measure depth) might account for this (a similar problem was also experienced by *Beck et al.* (2003) who pointed out this possible reason). The deviation in depth was removed by adding the calculated differences to measurements from seal tags. There appears to be a drift in the temporal offsets as well calculated by adjusting the clocks of the tags to agree with CTD. The temporal drift ranges from 10-120 seconds, but this error is of less concern for the purpose of estimating attenuation coefficients.



(a) CTD

(b) the hyper-pro



(c) Temperature profile

(d) Irradiance Profile (550 nm)

Figure 3.4: Temperature, depth and irradiance sensors used to calibrate measurements from electronic tags and samples of the profiles measured. (a) CTD with 5 seal sensors mounted on the board, which measure temperature and irradiance. (b) The the hyper-pro sensor, which measure irradiance and depth. (c) Temperature profiles taken on March 21, 2012. (d) Log-irradiance profiles taken on March 21, 2012. The offsets among measured temperature and irradiance is due to depth discrepancies.

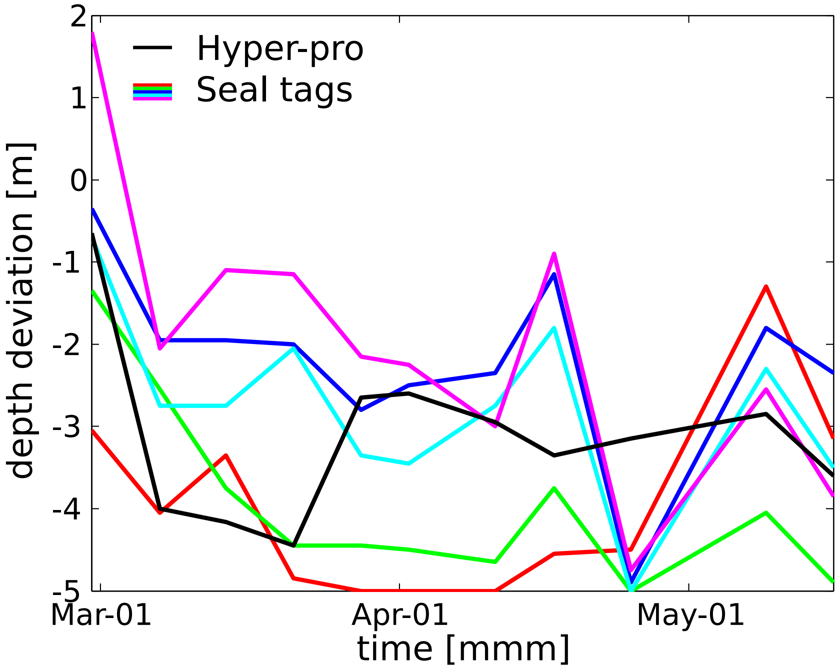
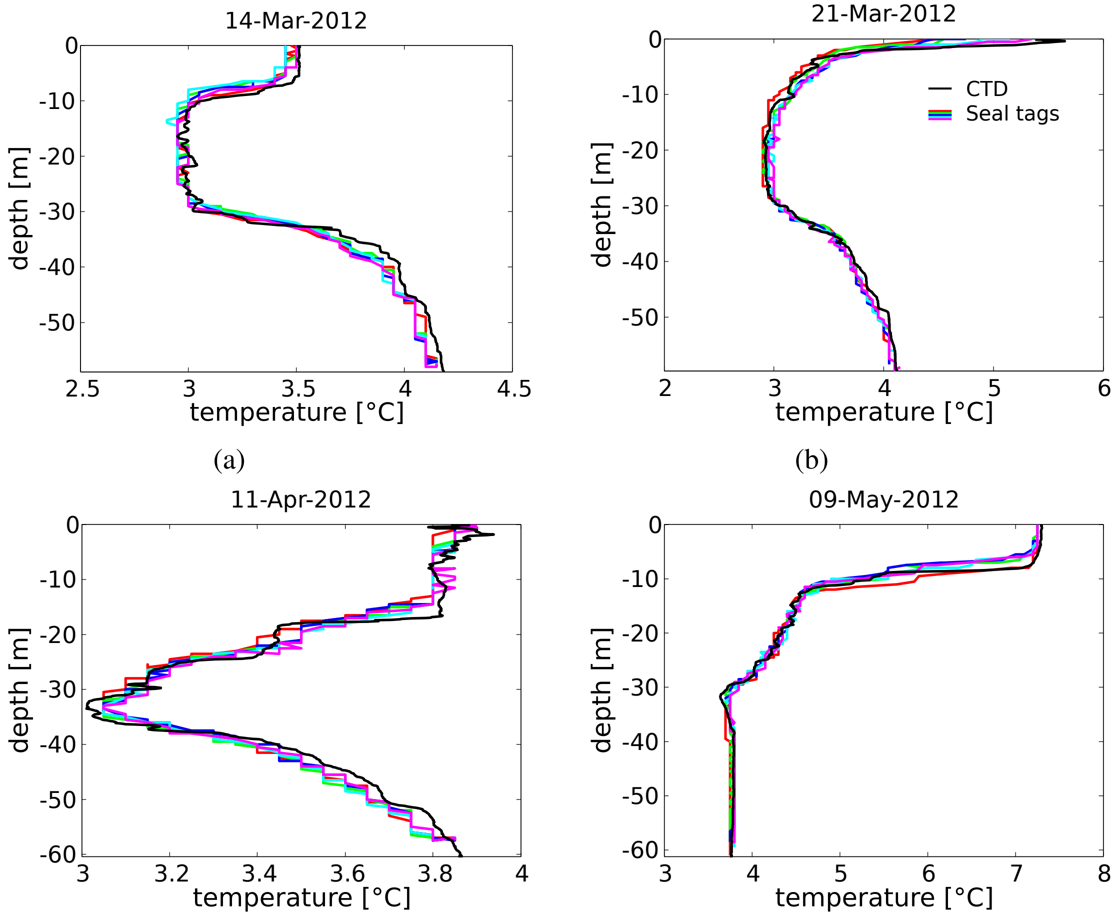


Figure 3.5: Deviation of measured depth of the electronic tags (coloured lines) and the hyper-pro (black line) from CTD measurements. The plotted lines represents the difference between values from electronic tags and the hyper-pro to CTD measurements for each collection day from February 20, 2012 to May 16, 2012.



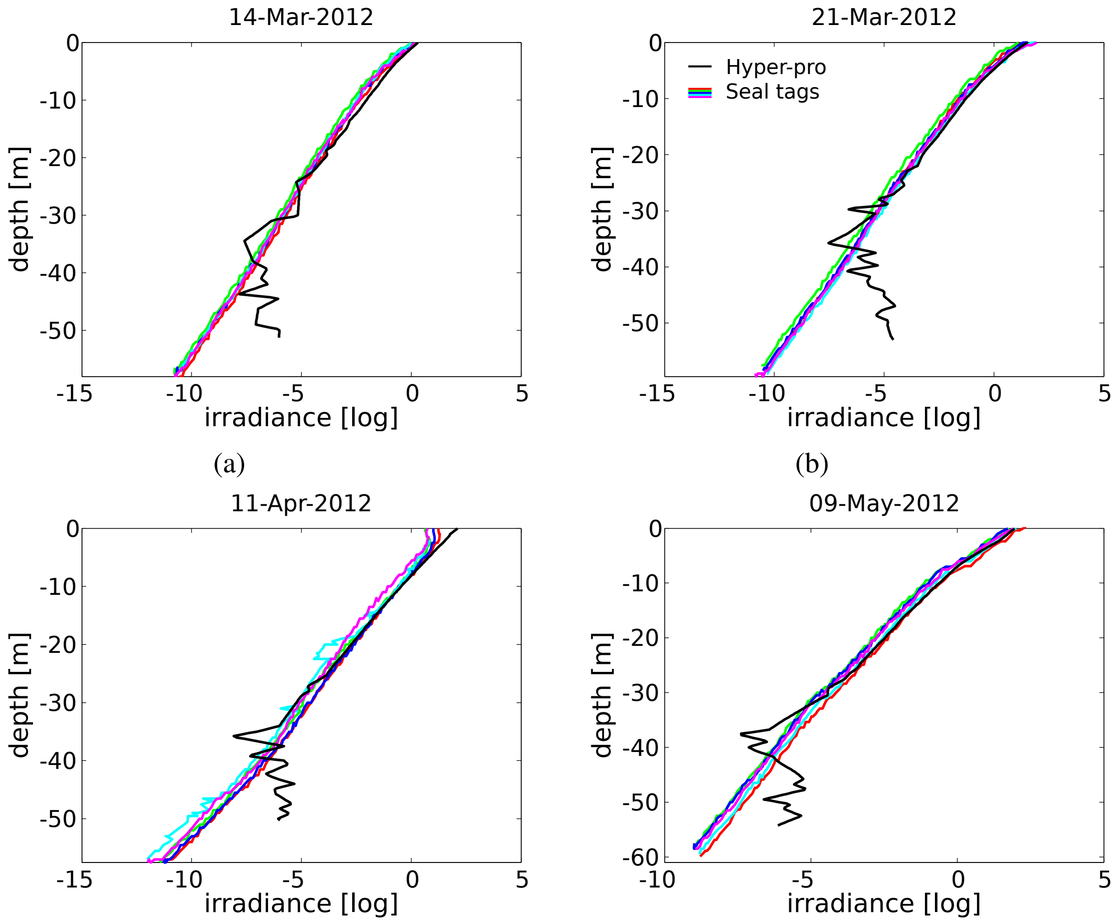
(a)

(b)

(c)

(d)

Figure 3.6: Comparison of temperature profiles from CTD (black line) and seal tags (multi-coloured lines) after adjusting for depth differences. Different Figures correspond to different dates.



(a) (b)

(c) (d)

Figure 3.7: Comparison of log-irradiance profiles at 550 nm from the hyper-pro (black line) and seal tags (multi-coloured lines) after adjusting for depth differences. Different Figures correspond to different dates.

Figures 3.6 and 3.7 show that profiles of temperature and log-irradiance (at 550 nm) obtained from seal tags are comparable to the temperature and irradiance profiles from the CTD and the hyper-pro after adjusting for depth differences. The hyper-pro's light sensor capability however is limited at low ambient light corresponding to log-irradiances of -6 and below (John Cullen, personal communication) which causes unreliable comparisons at depths deeper than 30-40 m.

3.3.2 Estimation of chl in Bedford Basin

Calculation of measured chl-a in the upper 10 m. Chlorophyll-a (will be referred as chl-a_i, where i is the sample number) was measured from bottle samples taken on the same date as CTD and the hyper-pro were deployed in Bedford Basin. Three bottle samples were taken at 1, 5, 10 and 60 m. Since only $K_d(550)$ in the upper 10 m is considered in the analysis in this thesis (details are given below), measurements at 60 m were not used. To get a single value of measured chl-a to represent the upper 10 meters, the following procedures was used.

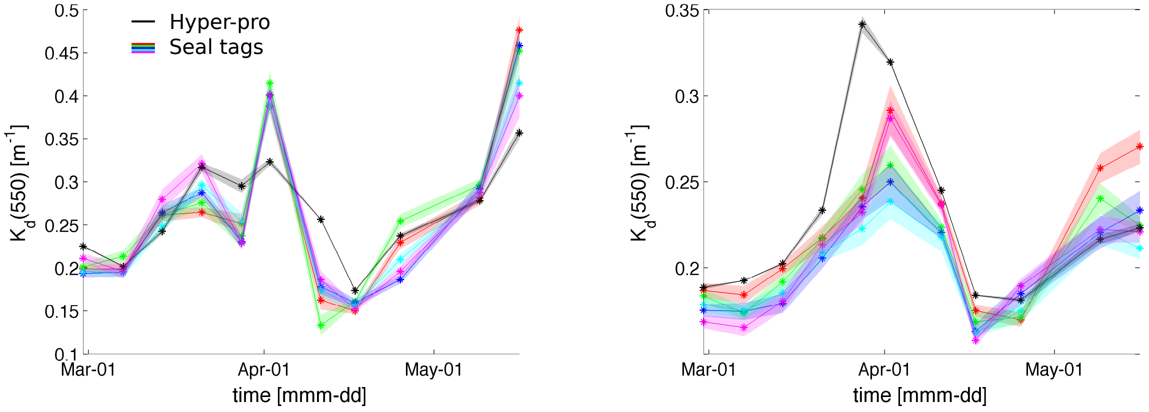
i) The mean, chl_{z_k}, and standard deviation, $\sigma_{chl_{z_k}}$, of 3 measurements, j, of chl-a_{jk} at each level k were calculated. ii) A representative value of chl-a in the upper 10 meters was then calculated using a weighted mean as follows:

$$chl\text{-}a = \sum_{k=1}^{3} w_{z_k}\, chl_{z_k} \tag{3.3}$$

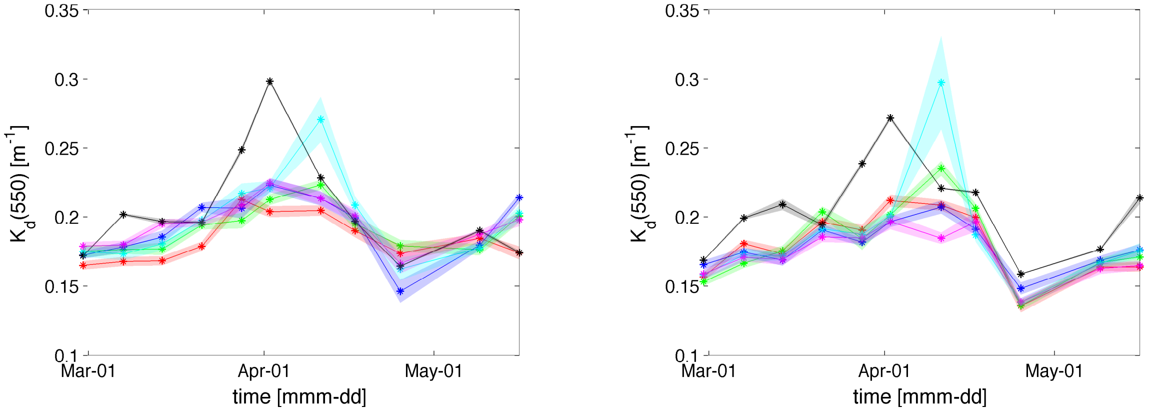
where chl-a is the weighted mean of chl_{z_k} in the upper 10 meters and $w_{z_k} = \frac{2}{9}, \frac{1}{2},$ and $\frac{5}{18}$ are the weights for depths z_k=1, 5 and 10 m, respectively. The weights, w_{z_k}, were calculated by vertically integrating chl_{z_k} values in the top 10 m, and dividing it by the total depth where measurements were made. Standard deviations for chl-a were calculated by assuming the errors at the three depths are independent, $\sigma_{chl\text{-}a} = \sqrt{\sum_{k=1}^{3} w_{z_k}^2 \sigma_{chl_{z_k}}^2}$.

Estimation of chl from $K_d(550)$ estimates. Attenuation coefficients, $K_d(550)$, were calculated from regression (type-I) of log-irradiance profiles that were corrected for depth differences. Figure 3.8 shows the time series of the calculated $K_d(550)$ at each 10 m interval at 5 m increments up to a depth of 35 m (see table 3.1 for values in the upper 10 m). In general, there is good agreement for $K_d(550)$ estimates from the hyper-pro and different seal tags. The RMSEs (difference between seal tags and the hyper-pro estimates in a least square sense) from 0-30 m have values around 0.04 m^{-1} which is around 20% of the values of the $K_d(550)$ estimates in the depth range (see table 3.2). This small $RMSE{:}K_d(550)$ ratio can be seen as well in Figure 3.8 where estimates at depth from 0-30 m have better fit as compared to estimates at below 30 m. The ratios of $RMSE{:}K_d(550)$ begin to increase for depths deeper than 30 m. Values of $RMSE$ around 0.11 m^{-1} are calculated for depths deeper than 30 m, which is around 66% of the value of the estimated $K_d(550)$ at this depth range. This is not surprising since ambient noise dominates the log-irradiance data measured from the hyper-pro (see log-irradiance profiles in Figure 3.7).

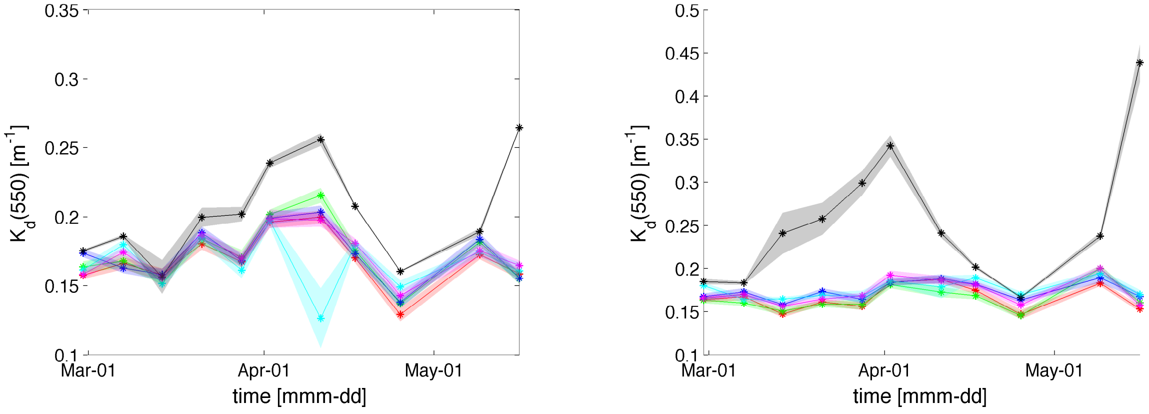
The time series reveals strong temporal patterns especially in the upper 10 m. Figure 3.9 shows that $K_d(550)$ in this region is higher as compared to other levels indicating that bio-optical conditions at Bedford Basin are strongly influenced by attenuating constituents in the upper 10 m. This observation is consistent with findings of *Roy* (1989) who reports higher concentrations of measured chl near the surface ($<$ 15 m) from February to April in Bedford Basin. Increasing $K_d(550)$ is observed from February 29-April 2, 2012 suggesting that there is a gradual build up of light attenuants. This is followed by a decrease in $K_d(550)$ up to April 17 and then a subsequent increase in $K_d(550)$ up to May 16, 2012. Figure 3.10 shows that the pattern observed from February 29-April 17, 2012 is closely related to the observed pattern in measured fluorescence (F), absorption ($a(550)$) from all constituents except water, beam attenuation ($\alpha(550)$) and chl a (see appendix B.1 for calculation) suggesting that estimated $K_d(550)$ in this period is due to phytoplankton bloom. From April 17-May 16, 2012, $K_d(550)$, fluorescence and beam attenuation have increasing trends. At the same period, the measured chl-a and absorption increases initially however they eventually tapers off suggesting that light attenuation at 550 nm in this period is mostly due attenuating agents other than phytoplankton.



(a) $0 - 10\ m$ (b) $5 - 15\ m$



(c) $10 - 20\ m$ (d) $15 - 25\ m$



(e) $20 - 30\ m$ (f) $25 - 35\ m$

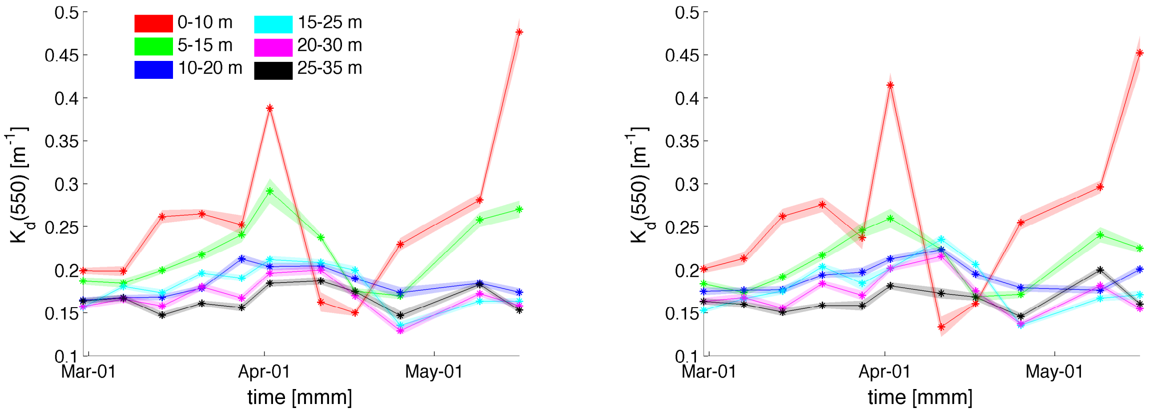
Figure 3.8: Time series of the estimated $K_d(550)$ from the hyper-pro (black line) and seal tags (multicolor) in Bedford Basin. Data points were obtained from February 29 - May 16, 2012. (a)-(f) Estimates at different depths. The light shade for each line represents the standard deviation calculated from the slope ($K_d(550)$) of the regression.

Table 3.1: Calculated values of depth deviations (d [m]), $K_d(550)$ [m^{-1}] estimates and their standard deviation ($\sigma_{K_d(550)}$ [m^{-1}]) for the duration of Bedford Basin experiment for seal tags and the hyper-pro. Depth deviations were calculated using equations **??** and **??**, $K_d(550)$ and $\sigma_{K_d(550)}$ were obtained using type-I regression. Multicolored texts represent estimates from seal tags (see Figure 3.8 for the corresponding plots).

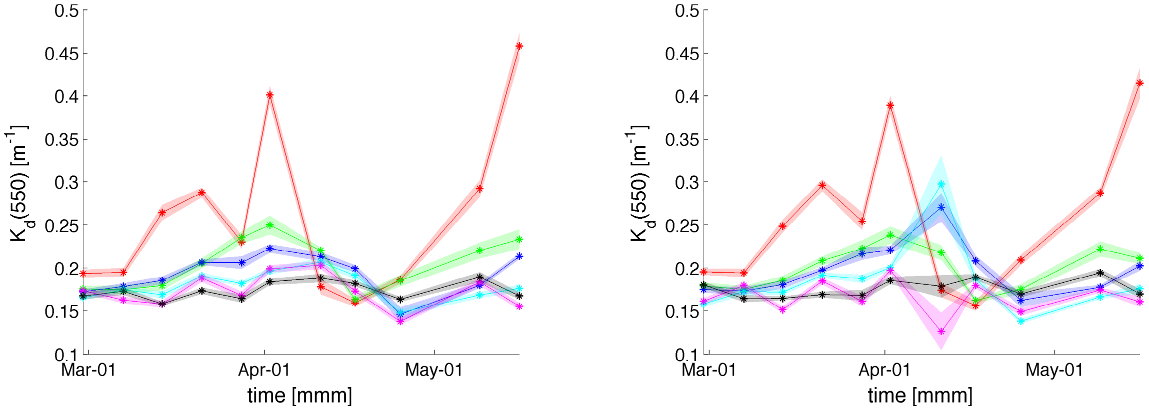
$Sensor$	variable	$Feb29$	$Mar07$	$Mar14$	$Mar21$	$Mar28$	$Apr02$	$Apr11$	$Apr17$	$Apr25$	$May09$	$May16$
$11A0091$	d	-3.05	-4.05	-3.35	-4.85	-5.00	-5.00	-5.00	-4.55	-4.50	-1.30	-3.15
	$K_d(550)$	0.20	0.20	0.26	0.26	0.25	0.39	0.16	0.15	0.23	0.28	0.48
	$\sigma_{K_d(550)}$	0.0034	0.0057	0.0077	0.0056	0.011	0.0082	0.011	0.0035	0.0073	0.0081	0.017
$11A0214$	d	-1.35	-2.55	-3.75	-4.45	-4.45	-4.50	-4.65	-3.75	-5.00	-4.05	-4.90
	$K_d(550)$	0.20	0.21	0.26	0.28	0.24	0.41	0.13	0.16	0.25	0.30	0.45
	$\sigma_{K_d(550)}$	0.0048	0.0073	0.0087	0.0086	0.013	0.014	0.012	0.0031	0.0081	0.0073	0.020
$11A0254$	d	-0.35	-1.95	-1.95	-2.00	-2.80	-2.50	-2.35	-1.15	-4.90	-1.8	-2.35
	$K_d(550)$	0.20	0.20	0.26	0.29	0.23	0.40	0.18	0.16	0.19	0.29	0.46
	$\sigma_{K_d(550)}$	0.0045	0.0052	0.0090	0.0057	0.079	0.011	0.0090	0.0045	0.0049	0.0098	0.016
$11A0256$	d	-0.7	-2.75	-2.75	-2.05	-3.35	-3.45	-2.75	-1.80	-5.00	-2.00	-3.50
	$K_d(550)$	0.20	0.19	0.25	0.30	0.25	0.39	0.17	0.16	0.21	0.29	0.41
	$\sigma_{K_d(550)}$	0.0038	0.0050	0.0076	0.0068	0.010	0.011	0.0081	0.0061	0.0062	0.0067	0.018
$11A0257$	d	1.80	-2.05	-1.10	-1.15	-2.15	-2.25	-3.00	-0.90	-4.75	-2.55	-3.85
	$K_d(550)$	0.21	0.20	0.28	0.32	0.23	0.40	0.19	0.15	0.20	0.29	0.40
	$\sigma_{K_d(550)}$	0.0053	0.0083	0.012	0.011	0.0078	0.0085	0.011	0.0036	0.0062	0.0077	0.024
hyper-pro	d	-0.65	-4.00	-4.15	-4.45	-2.65	-2.60	-2.95	-3.35	-3.15	-2.85	-3.60
	$K_d(550)$	0.22	0.20	0.24	0.32	0.30	0.32	0.26	0.17	0.24	0.28	0.36
	$\sigma_{K_d(550)}$	0.0016	0.0007	0.0036	0.0053	0.0077	0.0039	0.0009	0.0010	0.0028	0.0026	0.0079

Table 3.2: Root mean square error ($RMSE$) of $K_d(550)$ estimates from seal tags and the hyper-pro for different 10 m levels. Different colours of tagID refer to the corresponding time series of $K_d(550)$ in Figure 3.8. The values in gray measure the magnitude of RMSEs when compared to mean $K_d(550)$ estimates for each tag and are calculated using $\%ErrorFraction = 100 \times RMSE/\overline{K_d(550)}$.

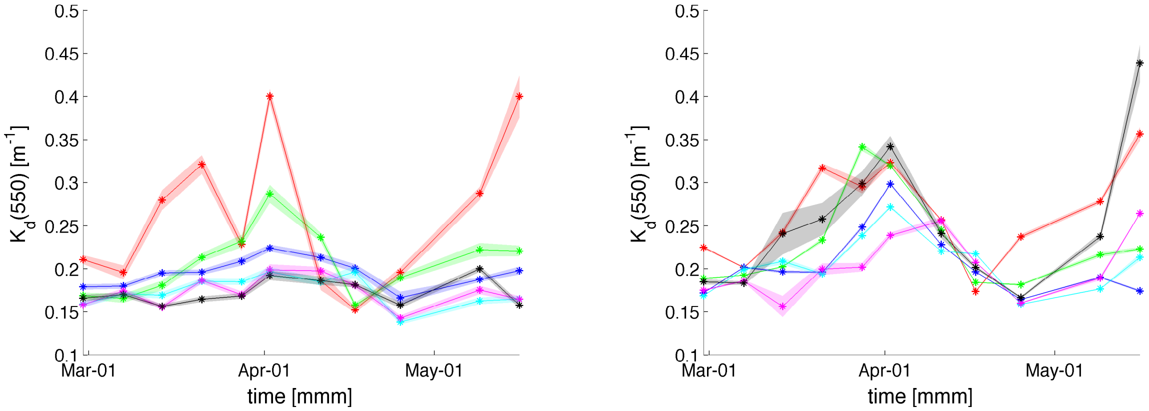
$TagID$	$0-10\ m$	$5-15\ m$	$10-20\ m$	$15-25\ m$	$20-30\ m$	$25-35\ m$
$11A0091$	0.0551 (21%)	0.0375 (17%)	0.0346 (18%)	0.0321 (18%)	0.0437 (26%)	0.1174 (70%)
$11A0214$	0.0599 (23%)	0.0367 (18%)	0.0332 (17%)	0.0346 (19%)	0.0407 (23%)	0.1162 (70%)
$11A0254$	0.0539 (21%)	0.0415 (20%)	0.0308 (16%)	0.0351 (20%)	0.0423 (24%)	0.1111 (64%)
$11A0256$	0.0412 (16%)	0.0460 (23%)	0.0315 (16%)	0.0412 (22%)	0.0542 (33%)	0.1100 (62%)
$11A0257$	0.0437 (17%)	0.0379 (18%)	0.0275 (14%)	0.0379 (22%)	0.0402 (23%)	0.1124 (65%)



(a) MK10 Serial # 11A0091 (b) MK10 Serial # 11A0214



(c) MK10 Serial # 11A0254 (d) MK10 Serial # 11A0256



(e) MK10 Serial # 11A0257 (f) the hyper-pro

Figure 3.9: Comparison of the estimated $K_d(550)$ per 10 m level (from 0-35 m) at Bedford Basin. Data points were obtained from February 29 - May 16, 2012. (a)-(e) $K_d(500)$ estimates from seal tags. Seal tags are determined by their serial numbers. (f) $K_d(550)$ estimates from the hyper-pro. The light shade for each line represents the standard deviation calculated from the slope ($K_d(550)$) of the regression.

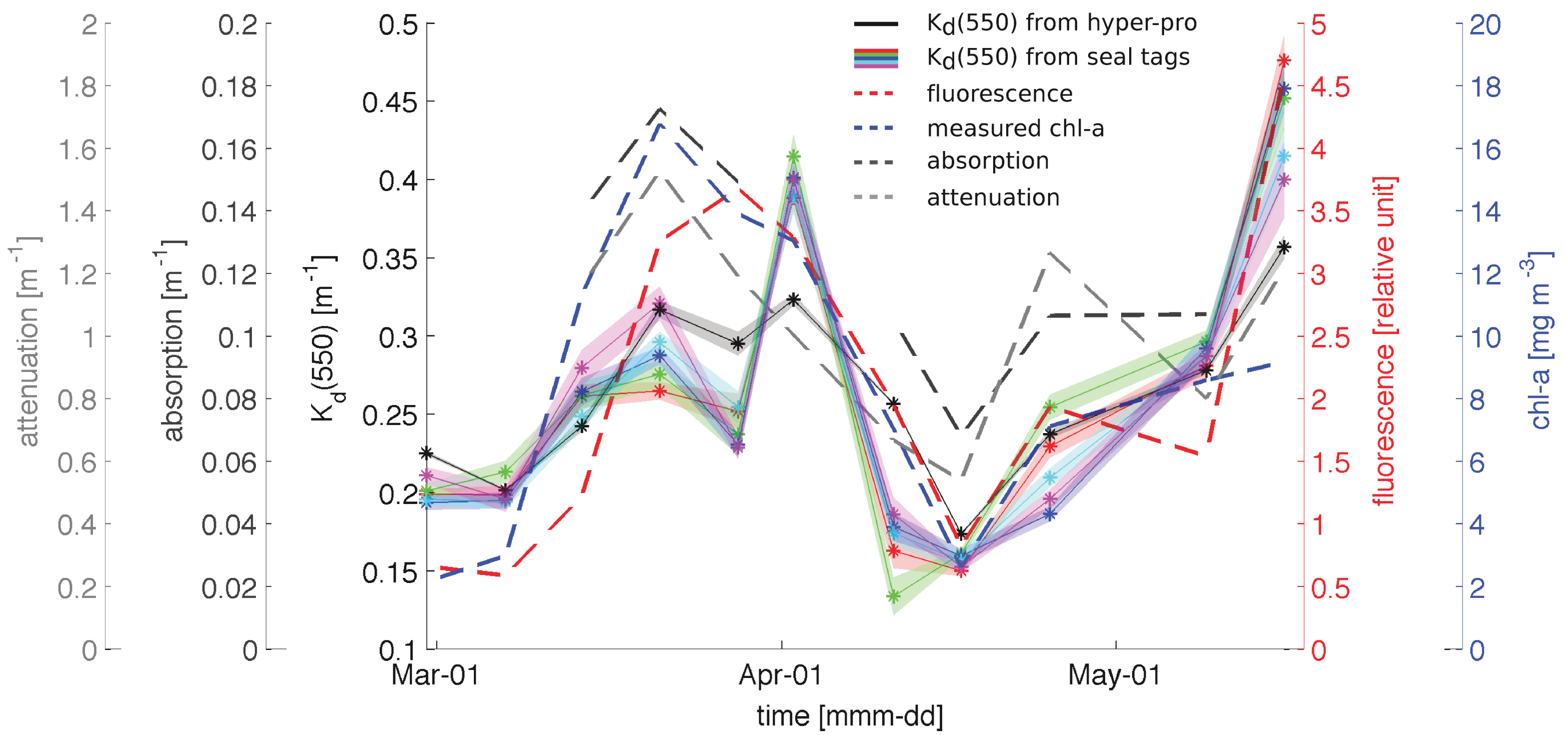


Figure 3.10: Time series of the estimated $K_d(550)$ from seal tags (solid coloured lines) and the hyper-pro (black line), absorption excluding contribution from seawater (broken dark gray line), beam attenuation (broken light gray line), fluorescence (broken red line) and mean measured chl a (broken blue line) for the upper 10 m. Fluorescence (F) was measured using the hyper-pro, beam attenuation ($\alpha(550)$) and absorption ($a(550)$) were measured from the bio-optical profiler attached to CTD and chl a were calculated as described in appendix B.1.

Table 3.3: Correlation matrix showing relationships among possible indicators of phytoplankton biomass in the upper 10 m.

		$K_d(550)$									
		$11A0091$	$11A0214$	$11A0254$	$11A0256$	$11A0257$	the hyper-pro	F	$a(550)$	$\alpha(550)$	chl
$K_d(550)$	$11A0091$	1	-	-	-	-	-	-	-	-	-
	$11A0214$	0.9579	1	-	-	-	-	-	-	-	-
	$11A0254$	0.9095	0.8423	1	-	-	-	-	-	-	-
	$11A0256$	0.9533	0.8919	0.9637	1	-	-	-	-	-	-
	$11A0257$	0.8779	0.8135	0.9756	0.9441	1	-	-	-	-	-
	hyper-pro	0.6655	0.5479	0.5309	0.6842	0.5164	1	-	-	-	-
	F	0.4449	0.3079	0.3770	0.5537	0.3877	0.8443	1	-	-	-
	$a(550)$	0.7147	0.5725	0.7136	0.7877	0.7837	0.8222	0.7800	1	-	-
	$\alpha(550)$	0.6999	0.6763	0.5646	0.6952	0.6373	0.7516	0.7213	0.8610	1	-
	chl	0.7204	0.5883	0.7343	0.8094	0.7580	0.8334	0.8446	0.9542	0.8604	1

Units: $K_d(550)$ [m^{-1}], F [μg L^{-1}], a(550) [m^{-1}], $\alpha(550)$ [m^{-1}] and chl-a [mg m^{-3}]

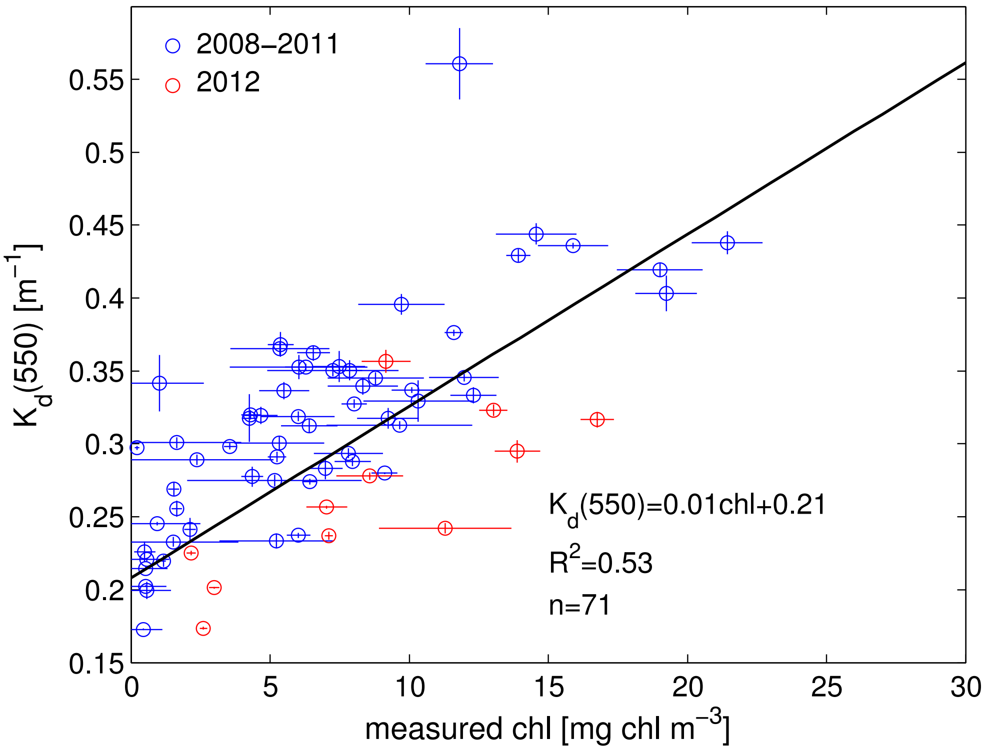


Figure 3.11: The relationship between estimated $K_d(550)$ from the hyper-pro and the mean measured *chl-a* concentration from bottle samples in the upper 10 m at Bedford Basin for the years 2008-2012. The error bars associated with $K_d(550)$ are from regression errors. The error bars associated with measured chl a concentration are mean standard deviation of chl a measurements in the upper 10 m. The regression line ($R^2 = 0.53$) is generated using type-II regression (*York et al.*, 2004) for 88 data points.

Table 3.3 presents the correlation matrix showing relationships among possible indicators of phytoplankton biomass. There is high correlation ($\rho > 0.8$) among estimates from electronic tags, this is not surprising since high precision is observed in the upper 10 m of $K_d(550)$ estimates (Figure 3.8a). Correlation of the electronic tags to estimates from the hyper-pro is lower ($0.51 < \rho < 0.67$) suggesting significant differences in measured light attenuation. Correlation between $K_d(550)$ estimates from electronic tags and other potential indicators of phytoplankton biomass like fluorescence, absorption, beam attenuation, *chl-a* are calculated as well. High correlations ($0.55 < \rho < 0.94$) are calculated with the exception of fluorescence ($0.3 < \rho < 0.55$). Among the biomass indicators, $K_d(550)$ from electronic tags correlates very well with measured *chl-a* ($0.59 < \rho < 0.81$) suggesting that temporal changes in phytoplankton biomass in Bedford Basin is captured in the $K_d(550)$ estimates.

Figure 3.11 shows a bio-optical relationship relating $K_d(550)$ to chl-a measurements from 2008-2012 at Bedford Basin for the upper 10 m. The linear relationship $K_d(550)$ = 0.01 chl+0.21 is found to explain 53% of the variability in the dataset. The presence of large and varying quantities of absorbing materials can affect the $K_d(\lambda)$ estimates. In an enclosed bay such as Bedford Basin, high concentration of CDOM and suspended particles are expected to be high because of contribution from the surrounding human activities. Estimates of $K_d(\lambda)$ are therefore expected to be higher than in more dilute environments (e.g. the Scotian Shelf). The use of light sensors sensitive to 550 nm is also another point worth mentioning. Phytoplankton absorption is dominated by blue and near-infrared spectra (*Brewin et al.*, 2011) while CDOM absorption is dominated by the blue spectra (*Retamal et al.*, 2007). Little information about phytoplankton and CDOM is therefore contained in the collected irradiance data which can cause incorrect chl estimates when using the derived bio-optical model.

3.4 Estimation of phytoplankton biomass on the Scotian Shelf

3.4.1 Estimation of attenuation coefficient

Seasonal changes in temperature and salinity affects density which creates stratification that breaks a water column into distinct strata. This results in non-uniform mixing, often a well-mixed layer on top of a poorly mixed deep layer (*Wuest et al.*, 2000). Phytoplankton may be found in abundance at the surface, right above the boundary, or directly below the mixed layer; wherever adequate supply of nutrients and sufficient sunlight intersects (*Mellard et al.*, 2010). Furthermore, turbulent motions in water caused by storm events can also influence the vertical distribution of phytoplankton (*Lewis et al.*, 1984) (e.g. advection of nutrient-rich deep waters to the surface, triggering primary production).

Numerous studies have shown that in the Scotian Shelf waters, phytoplankton is found in abundance within the mixed layer (*Dauchez et al.*, 1996; *Mousseau et al.*, 1996; *Greenan et al.*, 2004; *DFO*, 2005). Light attenuation in the upper mixed layer therefore contains enough information to infer phytoplankton biomass surrounding the shelf. The Scotian Shelf is categorized as case II however and is thus optically complex. Parameterizations of bio-optical models need to be locally tuned as algal and non-algal optical signals can vary regionally (*Craig et al.*, 2006; *Darecki and Stramski*, 2004).

The estimation of $K_d(550)$ in this section follows a similar procedure as in section 3.4.1, with the exception that regression of light attenuation profiles is done only for the mixed layer. Figure 3.3 shows that the slope of the light attenuation profile is steeper in the mixed layer than compared to the profile below it. This suggests that stronger light attenuation is associated in the upper mixed layer possibly because of the presence of phytoplankton and other light attenuants such as CDOM and suspended particles.

The mixed layer depth (MLD) was estimated from the temperature profiles collected by electronic tags. The temperature profiles were smoothed first using *loess* smoother in Matlab to remove the effect of temperature fluctuations and noise near the surface. A first difference criterion (difference of surface temperature to temperature values at each depth) was then used to determine the mixed layer depth. A threshold of $\Delta T = 0.4°C$ was chosen since it corresponds to the proper prediction of MLDs for most profiles. The criterion usually fails when seals are in shallow regions around Sable Island. In certain cases, there is strong fluctuation in temperature that exceeds $\Delta T = 0.4°C$, however if checked visually the reported MLD is incorrectly estimated. Hence, to ensure reliable estimation all calculated MLDs were visually inspected and were corrected to agree with visual MLD. Even visual inspection can fail at times however in cases where the seals didn't dive deep enough to reach the thermocline. In these cases, no MLD is reported. The use of criteria based on temperature works in most cases since daily and seasonal cycles in surface forcing have large heat fluxes, and weak salt fluxes (*Brainerd and Gregg*, 1995). However, in the event of intense rainfall strongly stratified pools of fresh water might necessitate the accounting for salinity in calculating MLD.

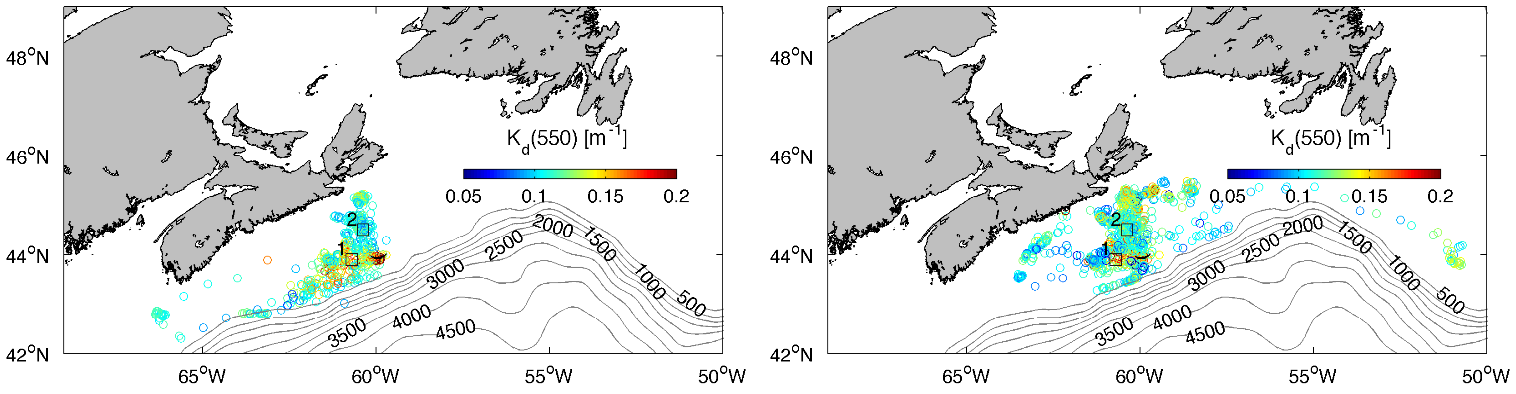
The spatial distribution of the estimated $K_d(550)$ values is shown in Figure 3.12 for 2009-2012. There seems to be a consistent spatial pattern that is independent of time. A closer look at Figures 3.12a-3.12c reveals that there is a higher probability of finding $K_d(550)$ estimates higher than 0.15 m^{-1} at the west of Sable Island (see box 1 in Figure 3.12) than at any region in the Scotian Shelf for all years. Similarly, values of $K_d(550)$ over the Middle Banks (north of Sable Island) remains below 0.15 m^{-1} for most estimates for all years (see box 2 in Figure 3.12). Seals however frequent the Middle Banks more often than West Bank despite low $K_d(550)$ estimates in the region (Figure 3.12d). This possibly suggests biological activities are minimally responsible for the high $K_d(550)$ near West

Bank.

Figures 3.13, 3.14 and 3.15 show the corresponding time series for seals tracks in Figures 3.12a-3.12c, respectively (see Figures B.2-B.22 in the appendix for all seals). These Figures further support the assertion made in Figure 3.12 regarding the time-independence of the observed $K_d(550)$ on the Scotian Shelf from June to February. Figure 3.13 (see also B.2-B.7) shows that $K_d(550)$ estimates from October 2009-January 2010 fluctuate between to 0.1-0.15 m^{-1}, except when seals cross the West Bank and in regions immediately adjacent to Sable Island. A similar trend is observed in Figures 3.14 and 3.15 (see also B.8-B.22) from September 2010-February 2011 and June 2011-January 2012, respectively. The estimated $K_d(550)$ values seem to remain independent of the temperature and the estimated MLD in the region.

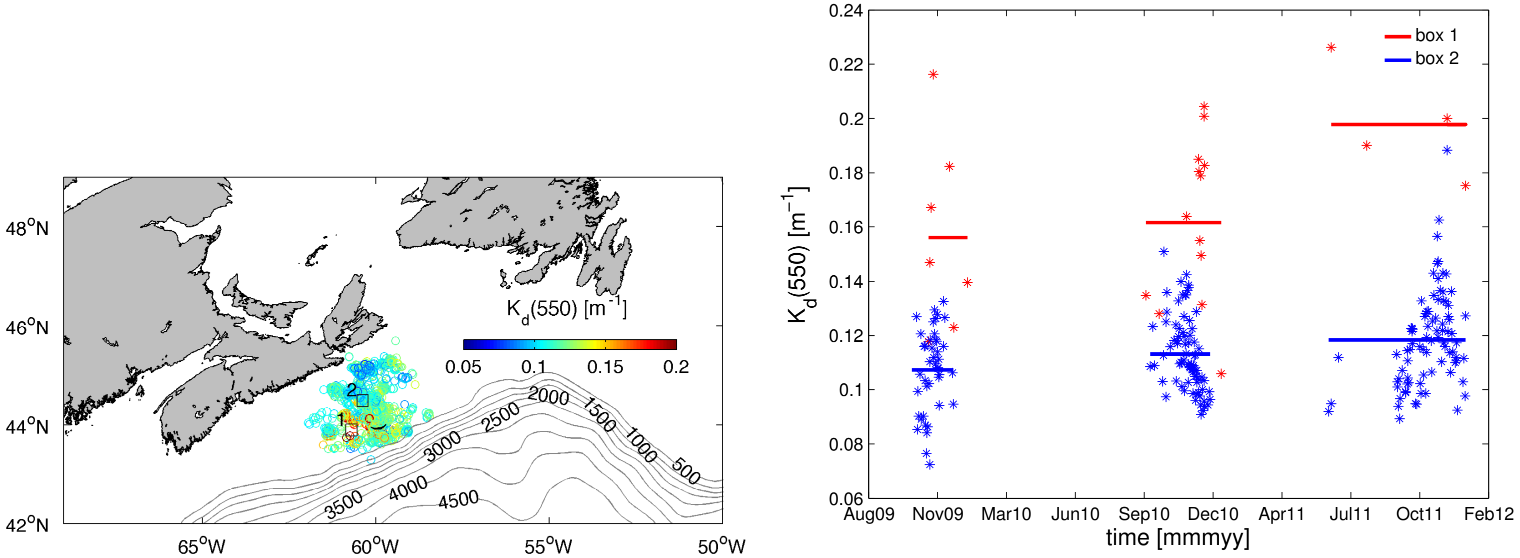
The relatively flat temporal trend in $K_d(550)$ is however not surprising. chl concentration in the Scotian Shelf remains fairly low from summer to fall. In the summer, chl concentration is generally less than 1 mg m^{-3} due to stratification and weak winds that limits the diffusion of the nutrients to the euphotic zone from nutrient-rich waters below. Increased wind mixing in the fall permits infusion of nutrients in the upper mixed layer which triggers a fall bloom with peak concentrations of ∽2 mg chl m^{-3} (*Greenan et al.*, 2004). Estimated $K_d(550)$ might change however in the spring as primary production on the Scotian Shelf has a peak chl concentration of ∽8 mg chl m^{-3}.

The consistency of measurements from electronic tags can also be seen in the data. Figures 3.16-3.18 shows time series of the regions (∼676 km^2) visited by at least 5 grey seals for each deployment year. The time series show estimates from different electronic tags in the same region are comparable except for data points near the coast and west of Sable Island. Regions that are visited more frequently can be found over the Middle Bank as compared to any other area in the Scotian Shelf. Table 3.4 shows the break down of the number of unique seal visits in each box and the number of $K_d(550)$ estimates that the trips generated. Varying seal dive frequency can be observed from month to month (varying number of $K_d(550)$ estimates). It can be seen as well that seals are likely to stay around longer in one region consistent with findings of *Harvey et al.* (2012) that seals concentrate on foraging zones to conserve energy. Hence, biologically active regions are sampled better.



(a) October 2009-January 2010

(b) September 2010-February 2011



(c) June 2011-January 2012

(d) June 2011-January 2012

Figure 3.12: Spatial distribution of the estimated of $K_d(550)$ from October 2009-January 2012. Location of $K_d(550)$ are mean daily positions calculated from 10:00-14:00. The bathymetry is shown in grey contours. (a) $K_d(550)$ estimates for 13 grey seals from October 2009-January 2010, (b) for 19 grey seals from September 2010-February 2011, and (c) for 13 grey seals from June 2011-January 2012. (d) Time series of the $K_d(550)$ estimates (asterisk) in boxes 1 and 2. The solid lines are the mean values of the estimates. The boxes are at the same location for all years and covers an area of 676 km^2.

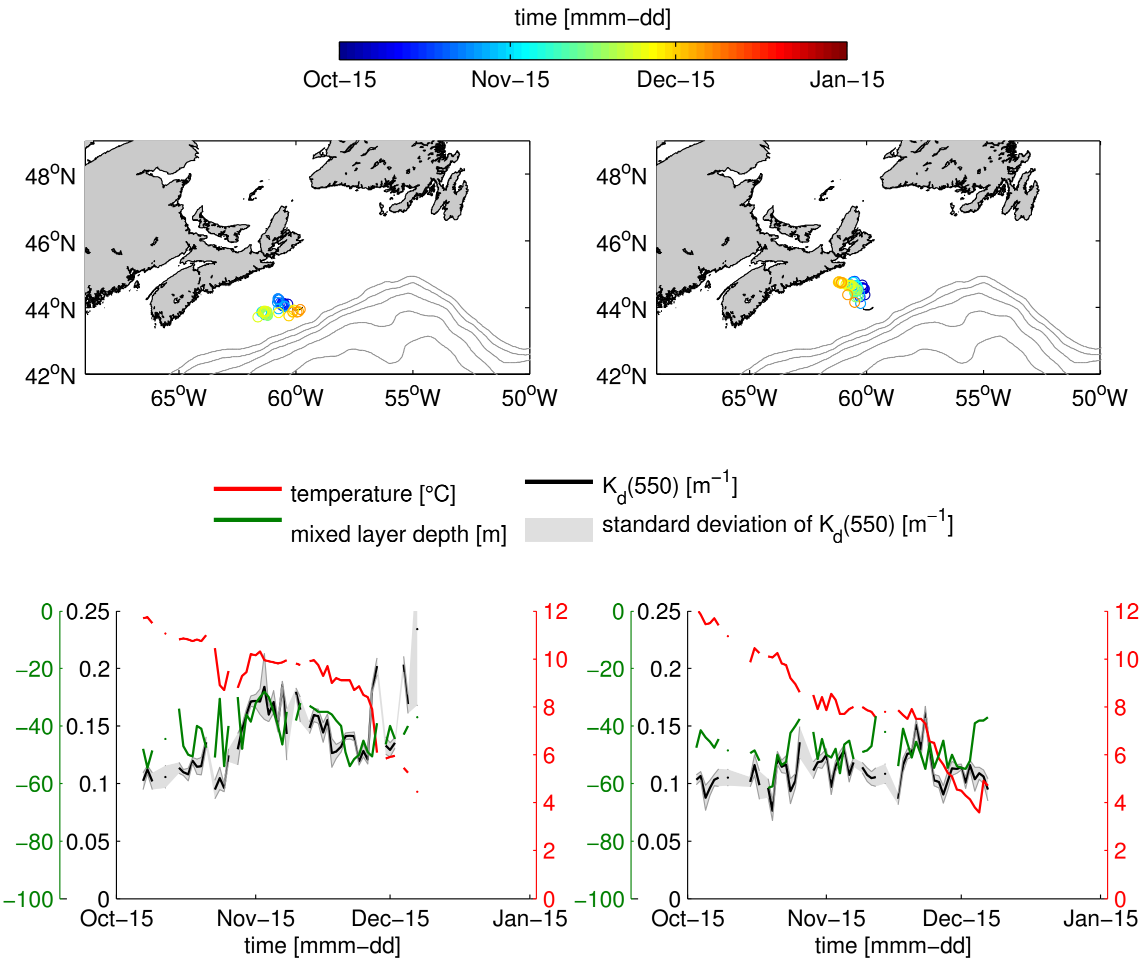


Figure 3.13: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9935 (left) and 78 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below. See Figures B.2 to B.7 for plots for other seals in the region.

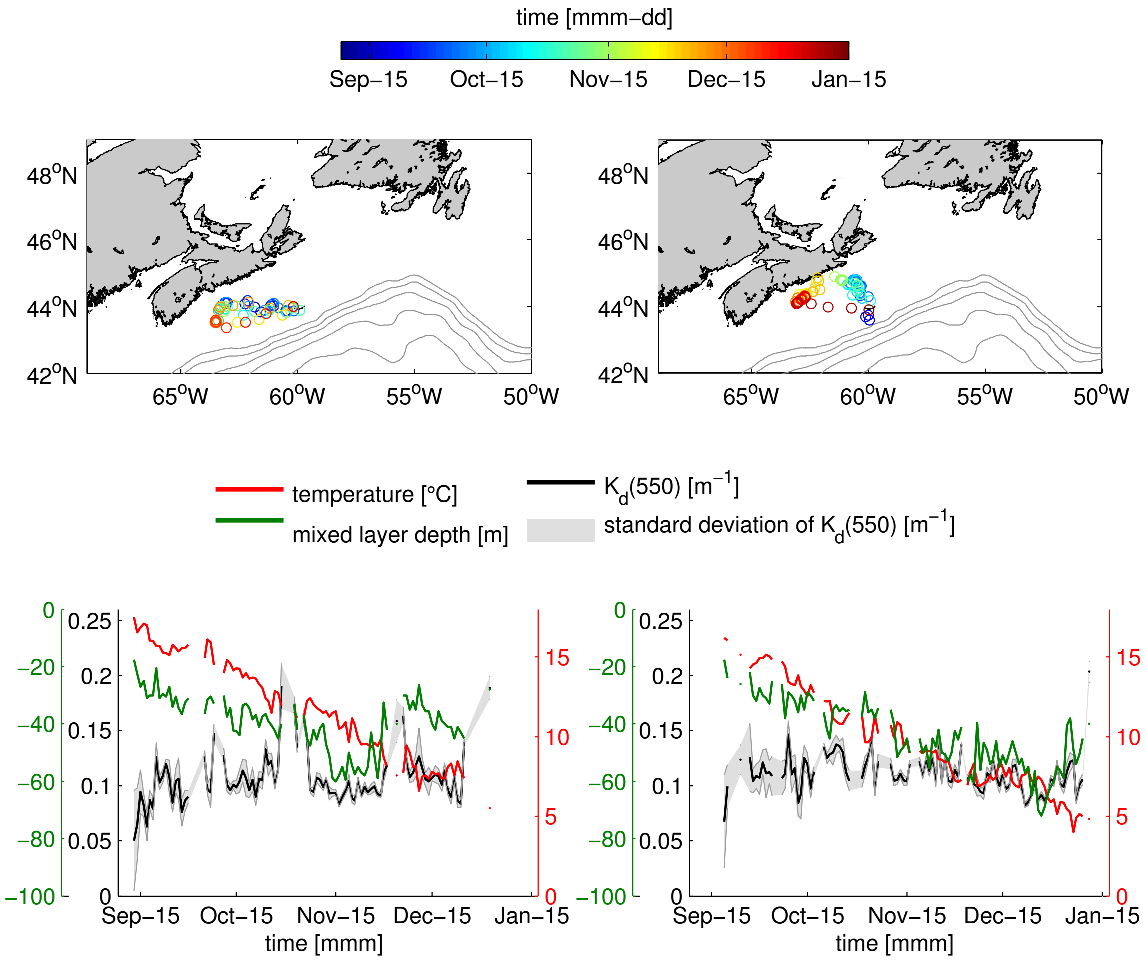


Figure 3.14: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 125 (left) and 9411 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below. See Figures B.8 to B.16 for plots for other seals in the region.

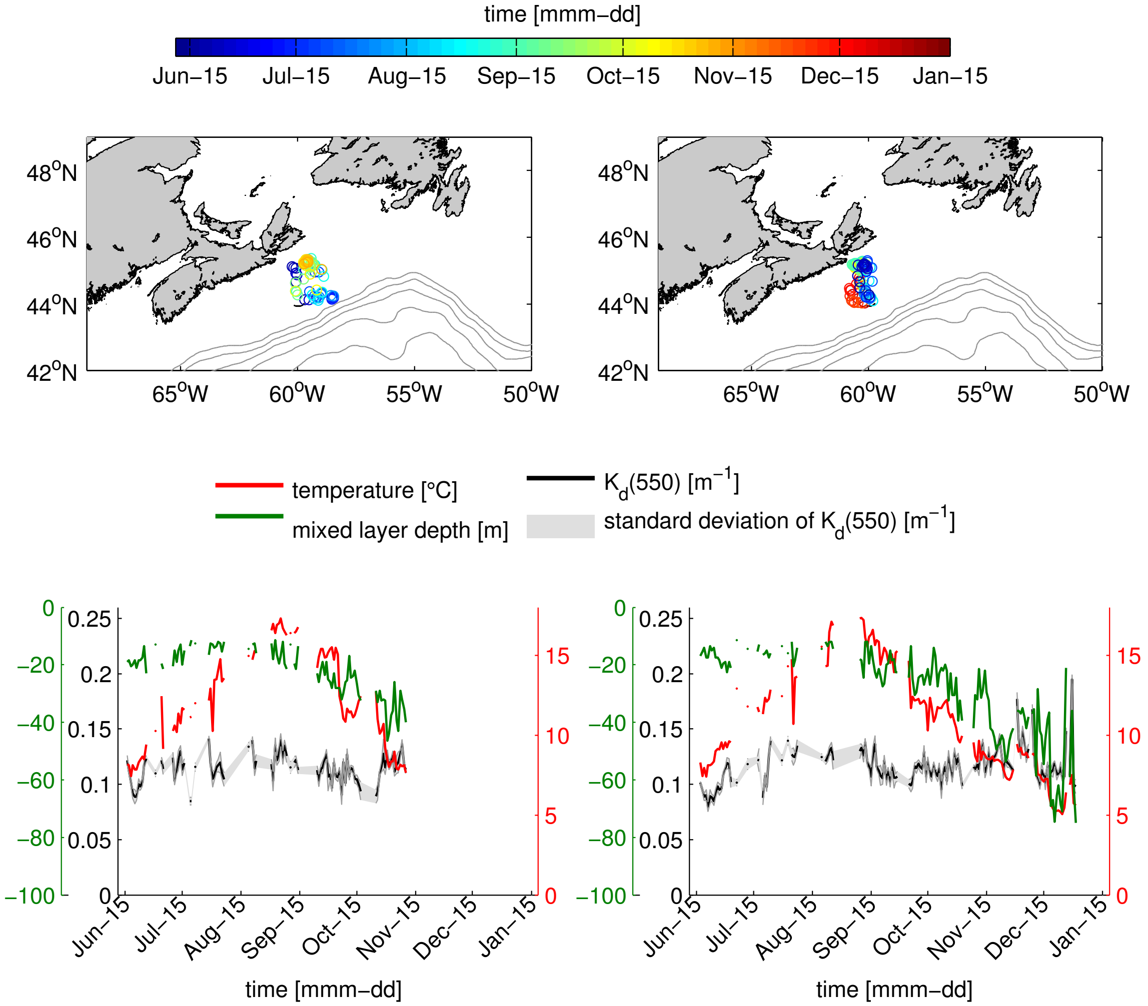
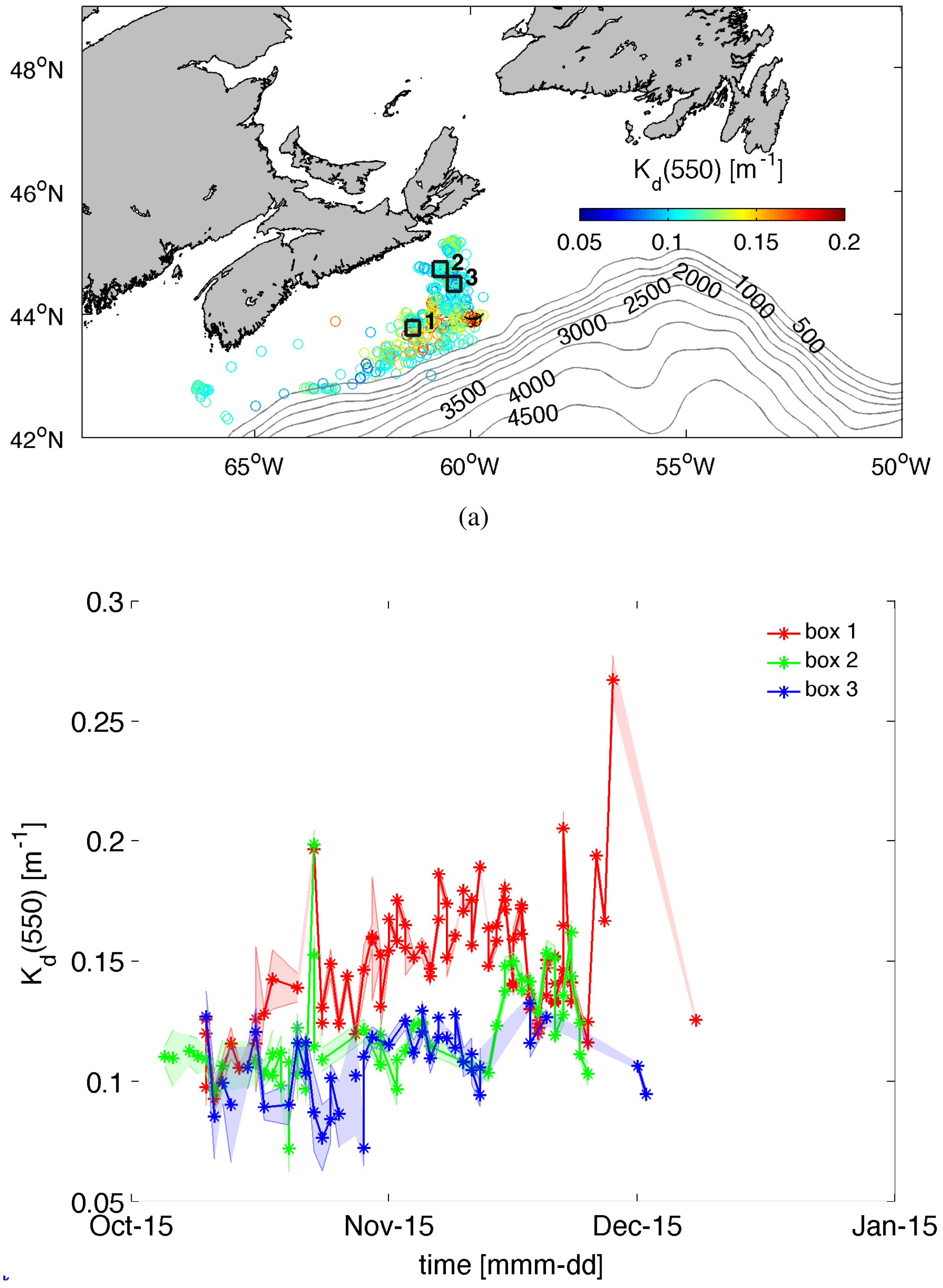


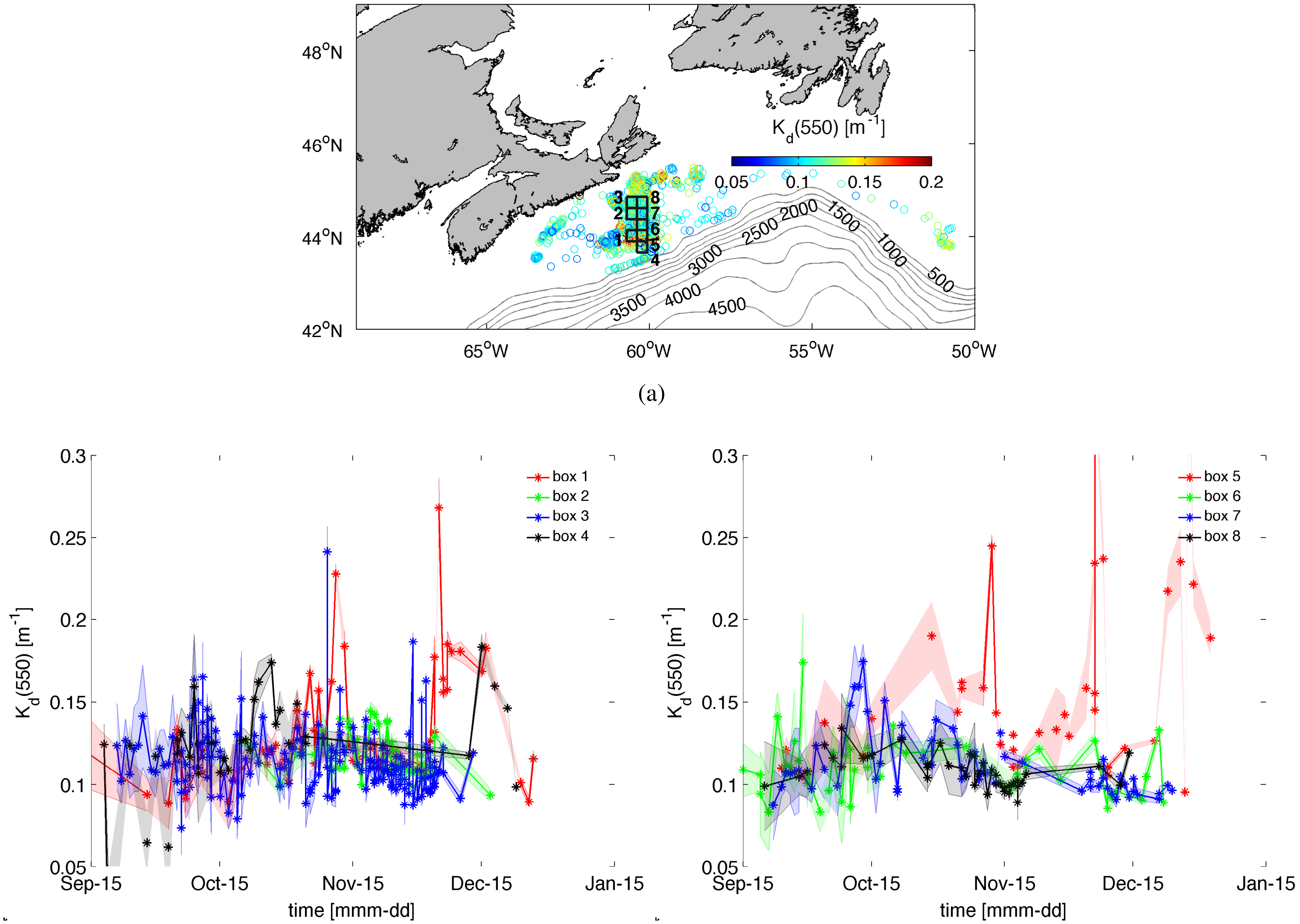
Figure 3.15: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 6630 (left) and 2718 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below. See Figures B.17 to B.22 for plots for other seals in the region.



(a)

(b)

Figure 3.16: Time series of the estimated $K_d(550)$ from selected regions along the Scotian Shelf. (a) Similar to Figure 3.12a with added emphasis on region of interest (boxes). Each box represents an area of approximately ∽676 km^2. (b) The corresponding time series of $K_d(550)$ estimates inside each box in (a). The location of the boxes were chosen such that there are at least 30 day $K_d(550)$ estimates and there are interaction for at least 5 grey seals for the duration of the experiment.

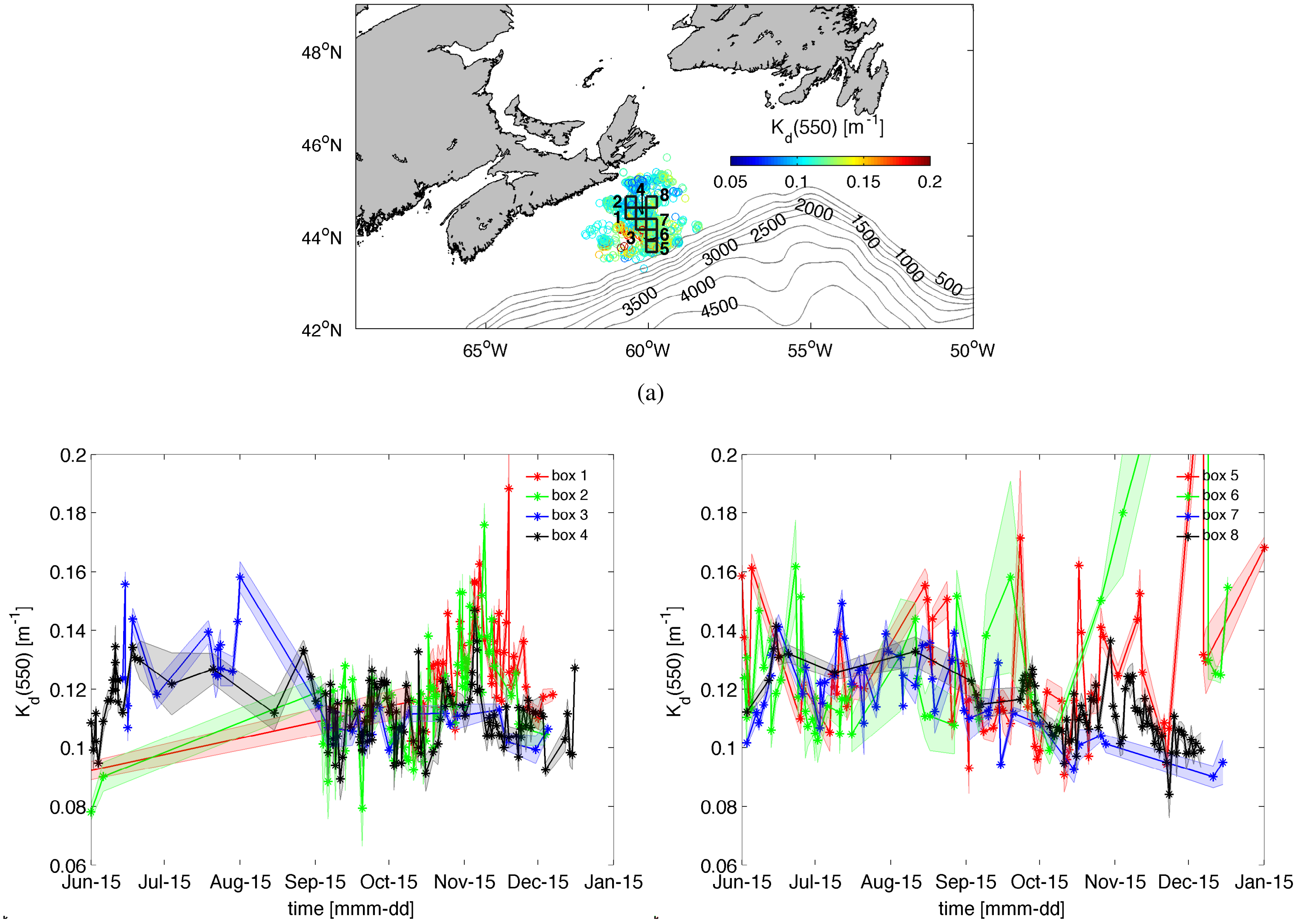


(a)

(b)

(c)

Figure 3.17: Time series of the estimated $K_d(550)$ from selected regions along the Scotian Shelf. (a) Similar to Figure 3.12b with added emphasis on region of interest (boxes). Each box represents an area of approximately 676 km^2. (b) The corresponding time series of $K_d(550)$ estimates inside each box in (a). The location of the boxes were chosen such that there are at least 30 day $K_d(550)$ estimates and there are interaction for at least 5 grey seals for the duration of the experiment.



(a)

(b)

(c)

Figure 3.18: Time series of the estimated $K_d(550)$ from selected regions along the Scotian Shelf. (a) Similar to Figure 3.12c with added emphasis on region of interest (boxes). Each box represents an area of approximately $676\ km^2$. (b) The corresponding time series of $K_d(550)$ estimates inside each box in (a). The location of the boxes were chosen such that there are at least 30 day $K_d(550)$ estimates and there are interaction for at least 5 grey seals for the duration of the experiment.

3.4.2 Bio-optical relationship in Scotian Shelf

In this section, chl observations from MODIS satellite were used to derive a bio-optical relationship on the Scotian Shelf. The chl data have temporal resolution of 1 day and were defined on a $\sim$ 4 km spatial grid. The chl data were matched in time and in space to the $K_d(550)$ estimates. Matching in time is straightforward since chl data and $K_d(550)$ estimates both have a similar temporal resolution (daily). To match the data points spatially however, $K_d(550)$ values were mapped first onto a spatial grid that matches the resolution of the chl data. $K_d(550)$ estimates were then assigned to the coordinates of the closest grid point if chl data is present, otherwise $K_d(550)$ estimates are discarded. Due to this nature of matching, it is possible then that multiple $K_d(550)$ estimates are matched with a single chl value. Figure 3.19 show the spatial distribution of satellite-based chl concentration from 2009-2012 that are nearest to location of $K_d(550)$ estimates.

Figure 3.20 shows the relationship between estimated $K_d(550)$ and satellite measured chl at Scotian Shelf. The Figure shows no noticeable pattern between satellite-measured chl and $K_d(550)$, and a $R^2 = 0.006$ was obtained when linearly regressed. The loss of important bio-optical information by using a sensor sensitive only to green light coupled with the large errors associated with satellite-measured chl in case II waters might have caused the insignificant relationship observed between $K_d(550)$ and satellite-measured chl along the Scotian Shelf.

3.5 Summary and Conclusions

Determining the appropriate bio-optical relationship is complicated in coastal waters such as Bedford Basin and on the Scotian Shelf. The large and varying quantities of absorbing and scattering materials can be introduced to surface waters from sediments and terrestrial sources (*K.Carder et al.*, 1989; *E.D'Sa and Miller*, 2003) in a short span of time which affect the bio-optical condition significantly. In spite of these, phytoplankton are more abundant and variable near the coast (*Y.Kang et al.*, 2006; *Moore et al.*, 2007; *O'Boyle and Silke*, 2010; *Song et al.*, 2011), and most of the times their optical properties are distinct from other absorbing and scattering materials (*Yentsch and Phinney*, 1985; *Hunter et al.*, 2008; *Moisan et al.*, 2012). Thus, with the right approach it is possible to use optical methods for detecting biological variability in coastal waters.

In this thesis, a simple approach was taken to derive a bio-optical relationship between

Table 3.4: Number of grey seal (N_{seals}) and $K_d(550)$ estimates (N_{K_d}, in grey) per month in specified regions (boxes) in Figures 3.16-3.18. $N_{totseals}$ is the total number of seals that visited each box and N_{totK_d} (in grey) is the total number of $K_d(550)$ estimates in each box. The hyphen in the table means that no tags are attached to the seals for a particular month.

		N_{seals} N_{K_d}									$N_{totseals}$ N_{totK_d}
Year range	location	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	
2009-2010	box 1	-	-	-	-	5	5	4	0	0	7
		-	-	-	-	9	43	28	0	0	80
	box 2	-	-	-	-	2	5	2	0	0	6
		-	-	-	-	10	28	12	0	0	50
	box 3	-	-	-	-	3	6	3	0	0	6
		-	-	-	-	7	30	5	0	0	42
2010-2011	box 1	-	-	-	1	5	4	3	0	-	9
		-	-	-	2	17	17	16	0	-	52
	box 2	-	-	-	0	2	7	3	0	-	10
		-	-	-	0	2	43	6	0	-	51
	box 3	-	-	-	2	5	11	6	0	-	12
		-	-	-	9	61	88	22	0	-	180
	box 4	-	-	-	2	5	1	1	0	-	8
		-	-	-	5	20	3	5	0	-	33
	box 5	-	-	-	3	4	7	9	1	-	19
		-	-	-	3	4	15	16	1	-	39
	box 6	-	-	-	7	7	4	7	0	-	14
		-	-	-	13	16	7	9	0	-	45
	box 7	-	-	-	4	7	6	8	0	-	11
		-	-	-	8	20	12	20	0	-	60
	box 8	-	-	-	3	9	4	3	0	-	10
		-	-	-	3	12	23	3	0	-	41
2011-2012	box 1	1	0	0	0	2	4	2	0	-	6
		1	0	0	0	5	31	13	0	-	50
	box 2	2	0	0	2	4	5	2	0	-	7
		2	0	0	16	27	35	5	0	-	85
	box 3	2	1	2	2	2	1	1	0	-	8
		5	2	9	3	8	4	3	0	-	34
	box 4	7	3	2	5	7	5	4	0	-	11
		19	4	2	19	42	26	21	0	-	133
	box 5	4	3	2	2	2	2	2	2	-	6
		5	9	4	13	16	10	7	2	-	66
	box 6	10	10	4	2	3	1	2	0	-	13
		10	18	6	4	3	2	6	0	-	49
	box 7	4	3	3	3	2	1	2	0	-	8
		7	11	12	9	4	2	2	0	-	47
	box 8	3	1	1	3	2	1	1	0	-	6
		5	2	1	4	25	24	17	0	-	78

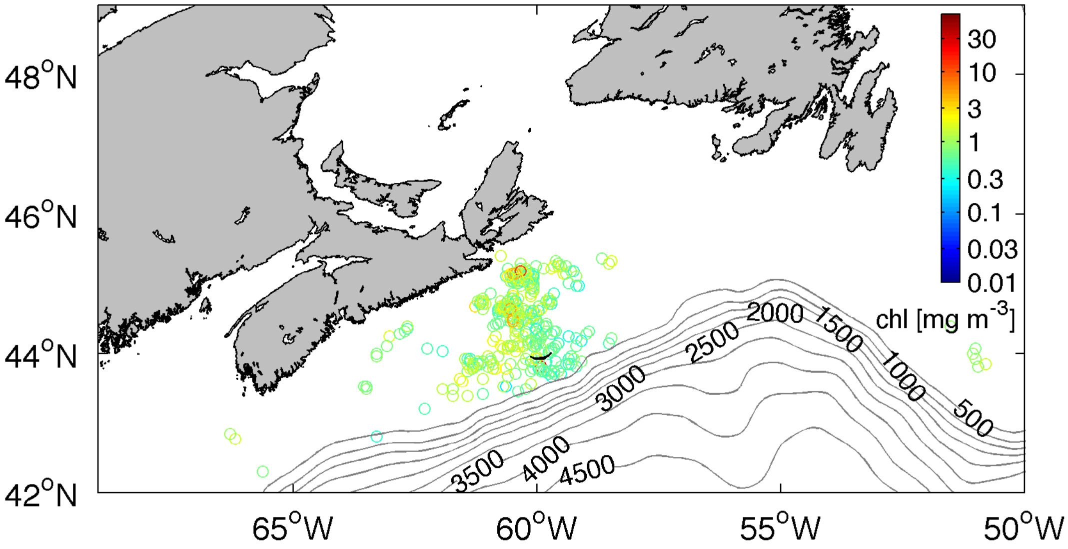


Figure 3.19: Spatial distribution of satellite-measured chl concentration matched on to locations of $K_d(550)$ estimates from 2009-2012. A mean distance of 1.57±0.32 km is calculated between chl grid and location of estimated $K_d(550)$. A total number of 407 data points were matched from 2009-2012.

$K_d(550)$ estimates to chl concentration at Bedford Basin and Scotian Shelf. $K_d(550)$ were estimated from light attenuation profiles measured using sensors sensitive only to 550 nm. The $K_d(550)$ estimates were able to uncover temporal and spatial patterns. In Bedford Basin, the temporal evolution of the $K_d(550)$ estimates from electronic tags (up until the end of spring bloom) match that of the measured chl, including other indicators of biomass such as fluorescence, absorption and attenuation. It can fail however to predict presence (or absence) of biomass at times when the contribution from other attenuating agents like CDOM and suspended particles dominates the basin. In Scotian Shelf, it was discovered through $K_d(550)$ estimates that bio-optical conditions in the region do not change much from June to February. It was also discovered spatial patterns in abundance of light attenuants, with the West Bank having higher concentration than in other parts of the Scotian Shelf.

The use of a sensor sensitive only to 550 nm poses a difficult challenge to overcome since a lot of important bio-optical information is lost because absorption of light in phytoplankton is dominated by blue and near infrared wavelengths. In addition, satellite-measured chl may itself not be particularly reliable in case II waters like on the Scotian

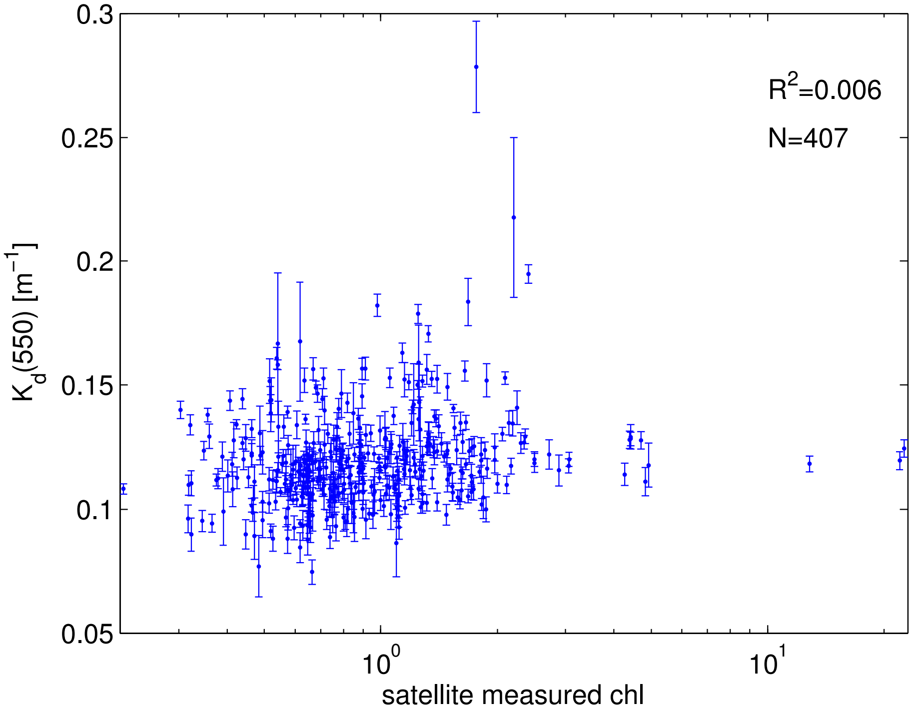


Figure 3.20: Relationship between estimated $K_d(550)$ and chl concentration from satellite along the Scotian Shelf. The error bars were calculated from estimation of $K_d(550)$ using type-I regression. $K_d(550)$ is plotted as a function of log $x - axis$ to expand points at small values of chl concentration.

Shelf. The approach of comparing actual values of chl, and thus estimation of phytoplankton biomass, might not be appropriate yet at this point as seen from the results from Scotian Shelf. $K_d(550)$ estimates can however reliably predict instances where there is overwhelming presence of phytoplankton like in the spring bloom and spatial distributions of light attenuants and may thus be considered for qualitative analysis.

CHAPTER 4

SUMMARY AND DISCUSSION

Two important topics to the the marine biogeochemical community have been explored: i) improving models for biogeochemical state estimation and ii) assessing possible new biogeochemical data sources. There is urgency in developing better biogeochemical models to study how marine food webs and marine biogeochemical cycle will be affected by the changes that the Earth is experiencing including global warming, ocean acidification and increased occurrence of eutrophication. This is accompanied by a greater need for biogeochemical data, both for model validation and data assimilation.

The first part of this thesis (chapter 2) described a simple method for suppressing bias and drift in biogeochemical models. The basic idea behind this scheme is to nudge the model toward climatology in prescribed frequency bands (hence called frequency-dependent nudging) while allowing high-frequency variability (sub-seasonal variations) to evolve freely. Frequency-dependent nudging was applied to a 1D biogeochemical model with 7 variables. The model's physical dynamics is described by a simple vertical diffusion equation with different diffusivity coefficients (η_1 and η_2) within and below the prescribed mixed layer. This is coupled to the biogeochemical model of *Fennel et al.* (2006). Three sets of experiment were done to test the performance of frequency-dependent nudging under varying biogeochemical conditions. Values from 3 different stations from a full 3D model of Bianucci et al. (*in prep*) were used to initialize the experiments. The same values were used as well as a benchmark of improvement gained through nudging (referred in the thesis as synthetic observations).

Improvements gained from frequency-dependent nudging were compared to conventional nudging. In general, both nudging methods gave considerable improvement when compared to results from the simple model. A closer look at the results however reveals that frequency-dependent nudging gave better results than conventional nudging. The model benefited from the extra frequencies allowed (greater than 1 cycle per year), which resulted in better state estimate for the frequency-dependent nudging case. Results of the experiments also reveal that no optimum nudging parameter, γ, exist for all variables. A wide range of γ values will give frequency-dependent nudging a considerable advantage over conventional nudging.

The second part of this thesis (chapter 3) assesses the feasibility of using data from electronic tags for ocean observations. In particular, the focus is on estimating biomass from light attenuation data. This is also consistent with the general goal of chapter 2 to improve biogeochemical estimates since observations are important component of data assimilation. There is value in using data from electronic tags because the data collected are of high temporal resolution (up to 1 data point per second) and can cover wide spatial domains when aggregated. Biologically active regions are naturally sampled as well because the frequent need for nutrition of the agents of data collections (marines animals).

A calibration experiment was carried out in Bedford Basin from February 29-May 16, 2012. Calibration is necessary since electronic tags were not originally intended for quantitative measurement of irradiance. Measured temperatures were calibrated as well against measurements using CTD. The experiment revealed that there are differences in depth measured by the tags, CTD and hyper-pro. *Beck et al.* (2003) (who used similar tags) pointed out that this might be due to the differences of the capability of the pressure sensor. Correction of the depth difference is necessary since it can affect the magnitude of the estimated $K_d(550)$. After adjusting for depth differences (calculated depth differences were added to seal tags measurements, see chapter 3 for complete details), the resulting temperature and irradiance profiles were found to be comparable to measurements from CTD and hyper-pro. There is good agreement as well in the temporal trend of the $K_d(550)$ estimates among tags and the hyper-pro, and when compared to other possible indicators of phytoplankton biomass (e.g. absorption, attenuation, fluorescence).

Results from analysis of the Scotian Shelf data reveal that electronic tags can be used to deduce spatial patterns in $K_d(550)$. Consistent spatial pattern of values of $K_d(550)$ over 3 years were observed from the Scotian Shelf i.e. high values of $K_d(550)$ in the west coast of Sable Island and low $K_d(550)$ in the Middle Bank. The possible reasons for this can't be provided however as, to my knowledge, there are no available biogeochemical data in these regions at a similar time frame.

The goal of the second part of this thesis is to estimate phytoplankton biomass from light attenuation coefficients. The results suggest that this is not yet possible; attempts at generating estimates of phytoplankton biomass did not yield encouraging results. Phytoplankton dominates absorption in the blue and near infrared region and in using a sensor sensitive only to 550 nm (green light), a lot of important information about the phytoplankton is lost. The presence of other attenuating agents in the coastal environment (Bedford Basin and Scotian Shelf) make biomass estimation more challenging as the right proportion of attenuating agents can't be established using only absorption at 550 nm. Despite this, results suggest that $K_d(550)$ estimates deduced the temporal and spatial patterns that may be associated with phytoplankton biomass.

4.1 Future work

In the first part of the thesis, it was shown that frequency-based nudging can be used to improve 1D biogeochemical model estimates. As an extension, frequency-dependent nudging should be considered for 3-D biogeochemical models since more variability is involved due to the presence of horizontal physical dynamics. This is particularly true if strong nudging ($<$ 30 days) is being considered since, as previously seen, higher frequencies are being suppressed and artificial phase lags are possibly being introduced which can then affect the timing of biogeochemical cycles. Another possible application of frequency-dependent nudging is the estimation of model bias (e.g. *Nerger and Gregg*, 2008) which can lead to better understanding of biogeochemical processes.

Results from the second part of the thesis suggest that physical variables (e.g. temperature and depth measurements) of the electronic tag data can already be used for ocean observations. More work is still needed however to be able to estimate phytoplankton biomass. Future work should focus on inclusion of another sensor that is sensitive to a different wavelength (blue or red). Several studies on satellite algorithms (*Gitelson et al.*, 2009; *Moses et al.*, 2009; *Gurlin et al.*, 2011) show that at least two wavelengths are needed to be able to retrieve chl-a concentration in case-II waters with confidence, and discriminate contributions from CDOM.

APPENDIX A

ADDITIONAL FIGURES FOR CHAPTER 2

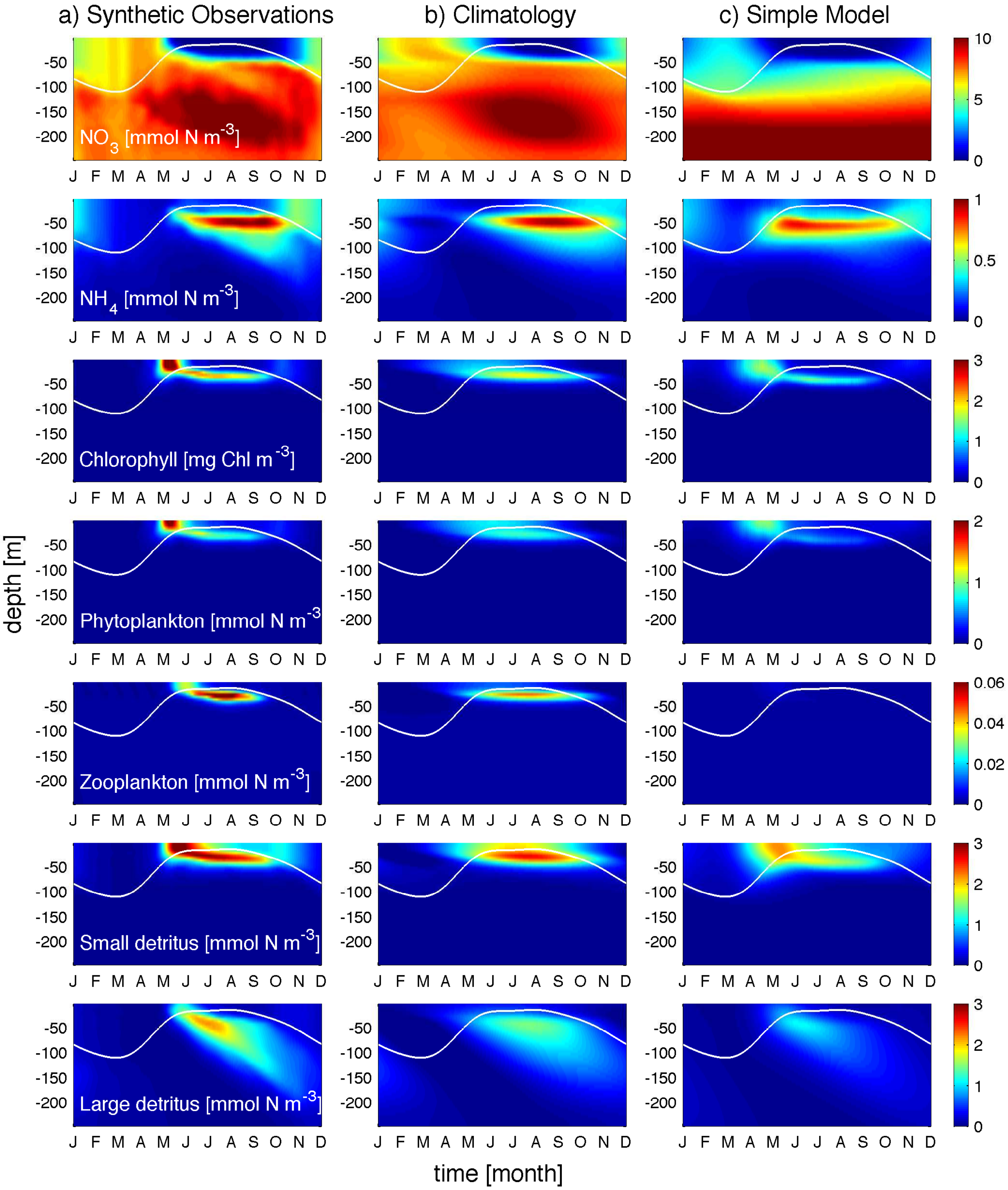


Figure A.1: Same as Figure 2.6 but for station 2.

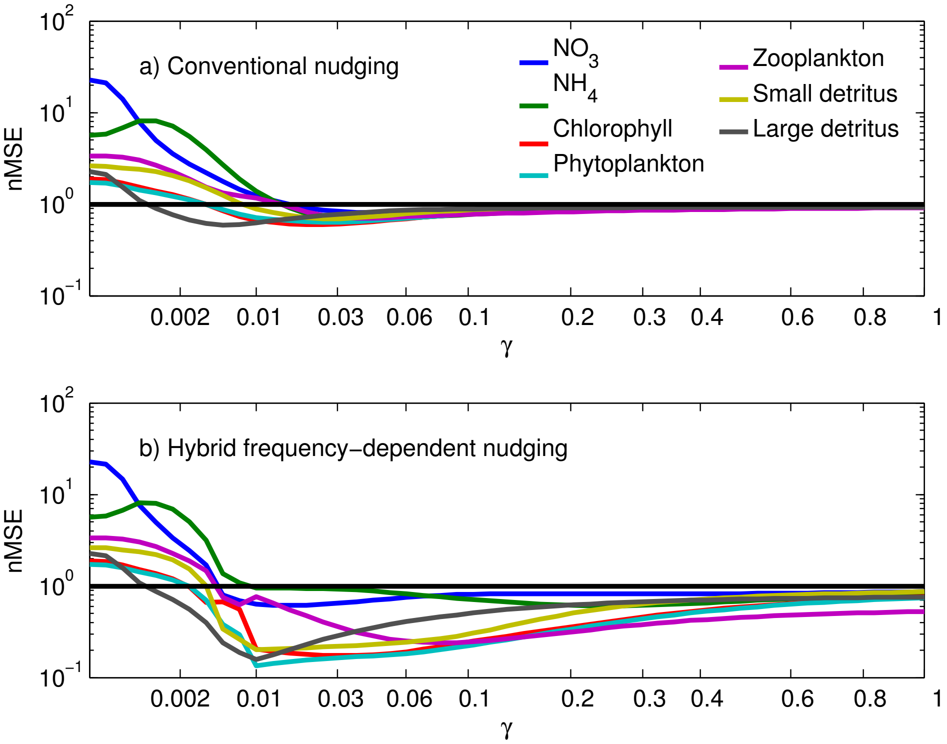


Figure A.2: Same as Figure 2.7 but for station 2.

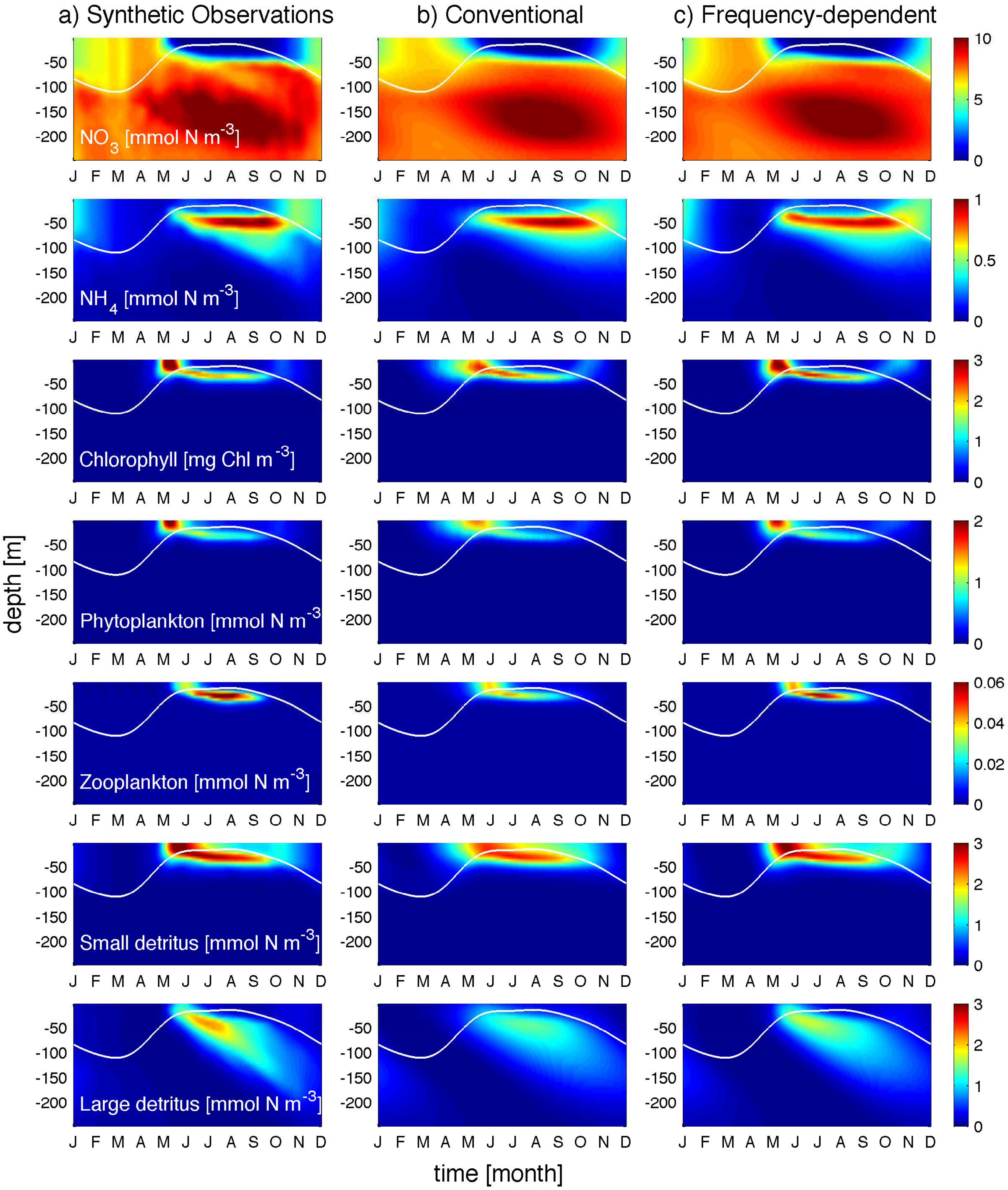


Figure A.3: Same as Figure 2.8 but for station 2.

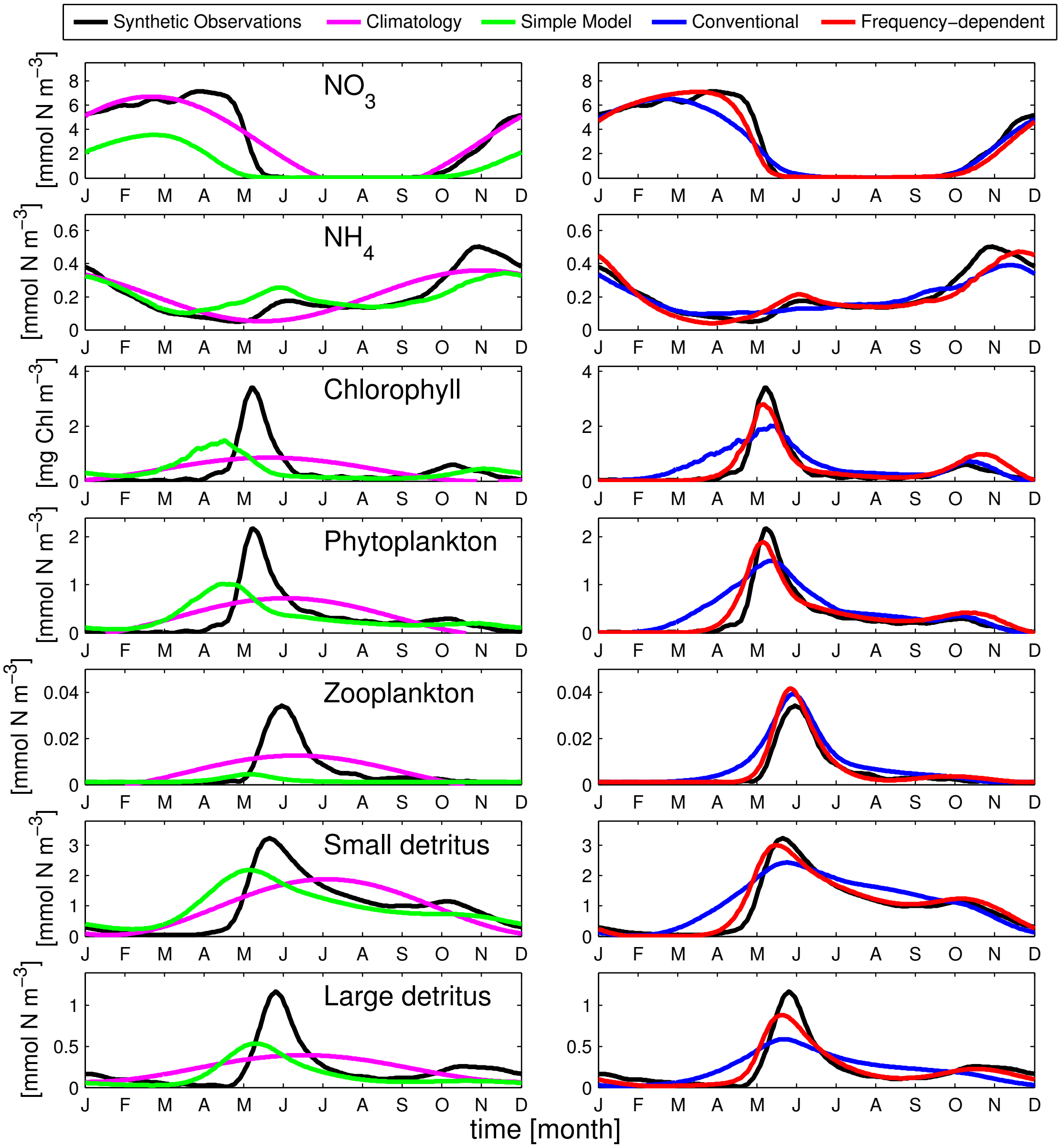


Figure A.4: Same as Figure 2.9 but for station 2.

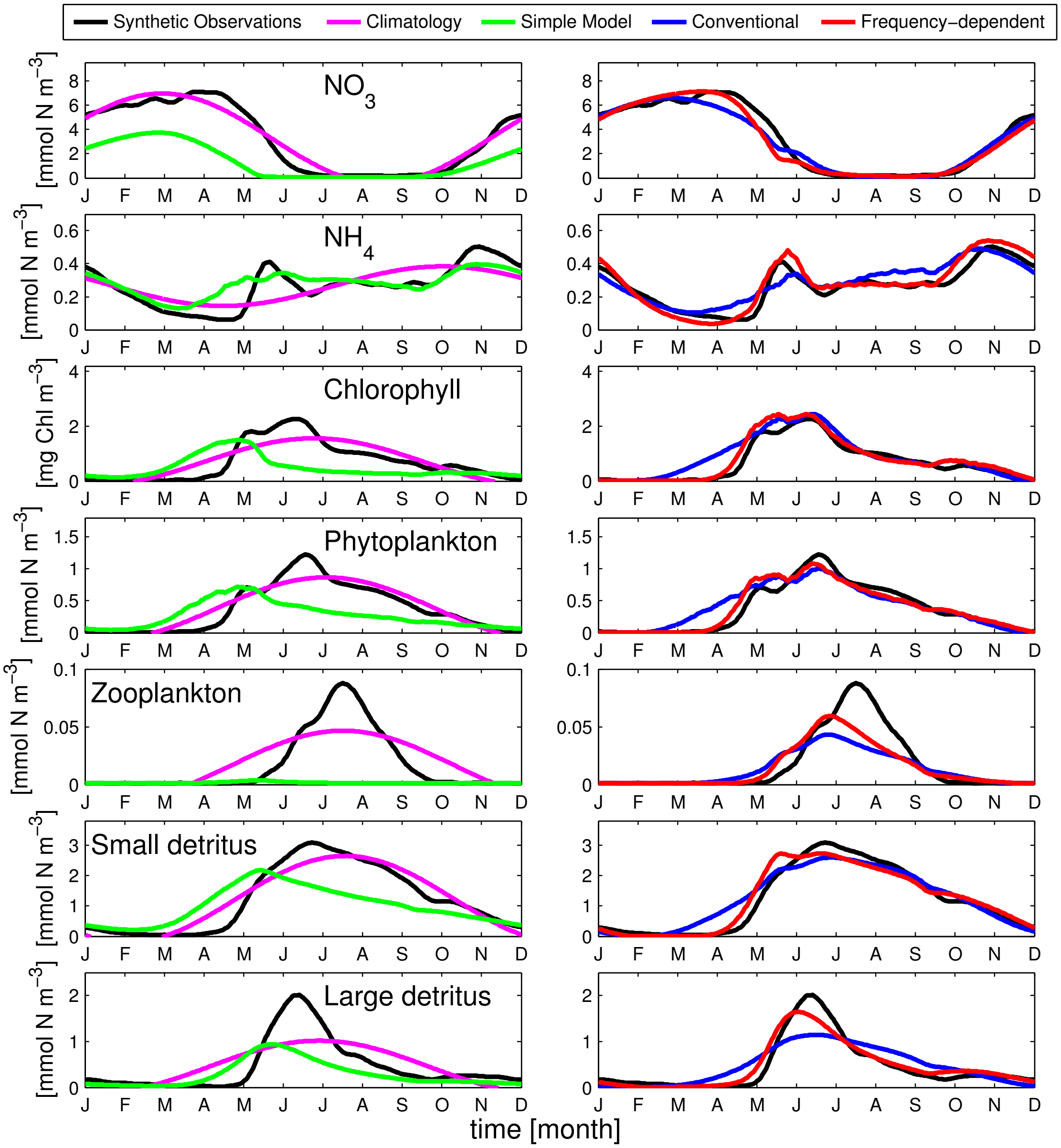


Figure A.5: Same as Figure 2.10 but for station 2.

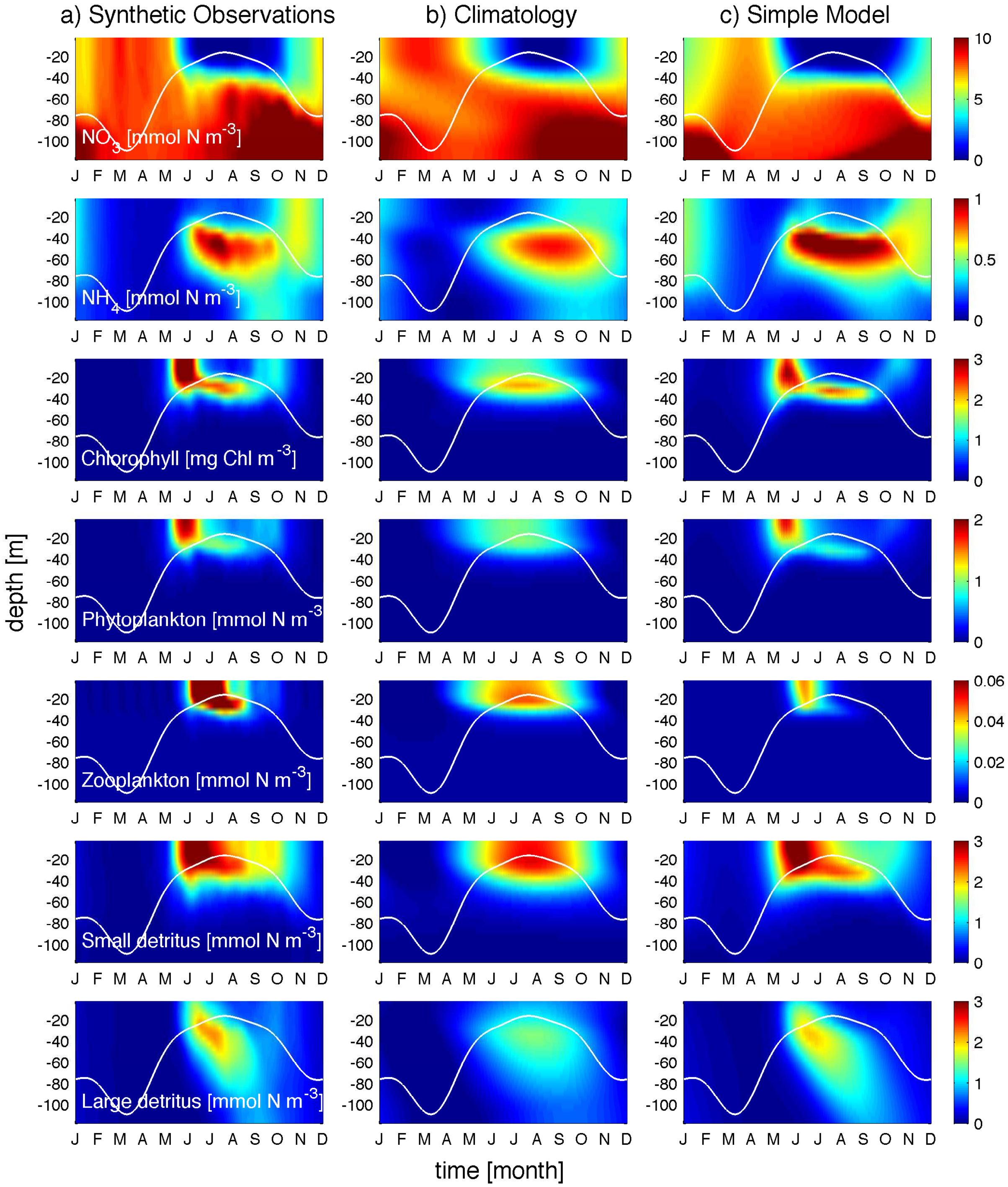


Figure A.6: Same as Figure 2.6 but for station 3.

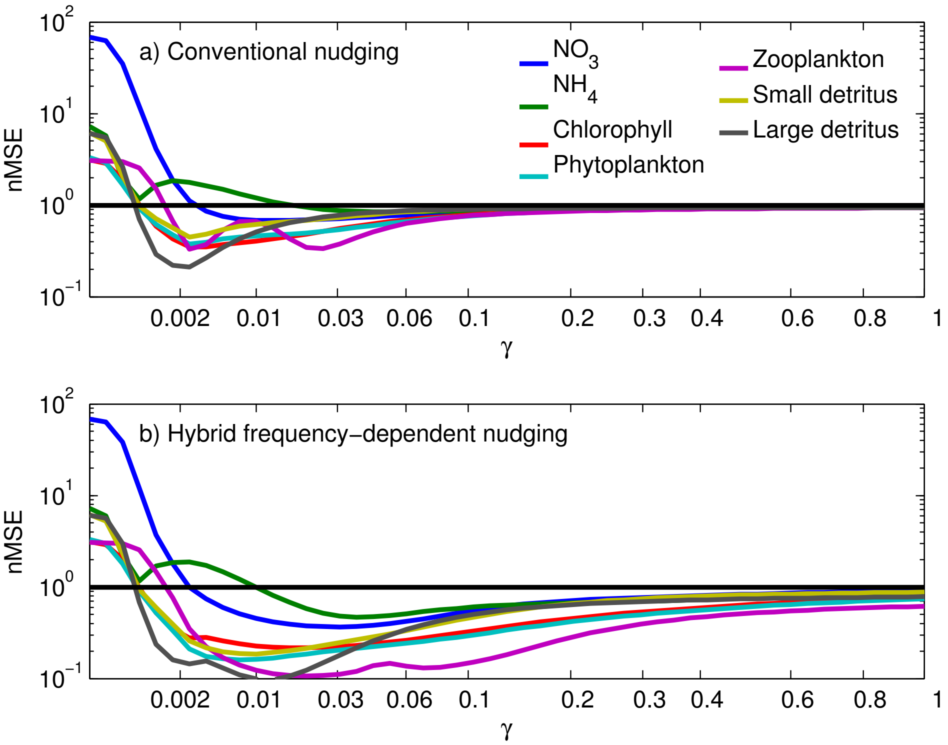


Figure A.7: Same as Figure 2.7 but for station 3.

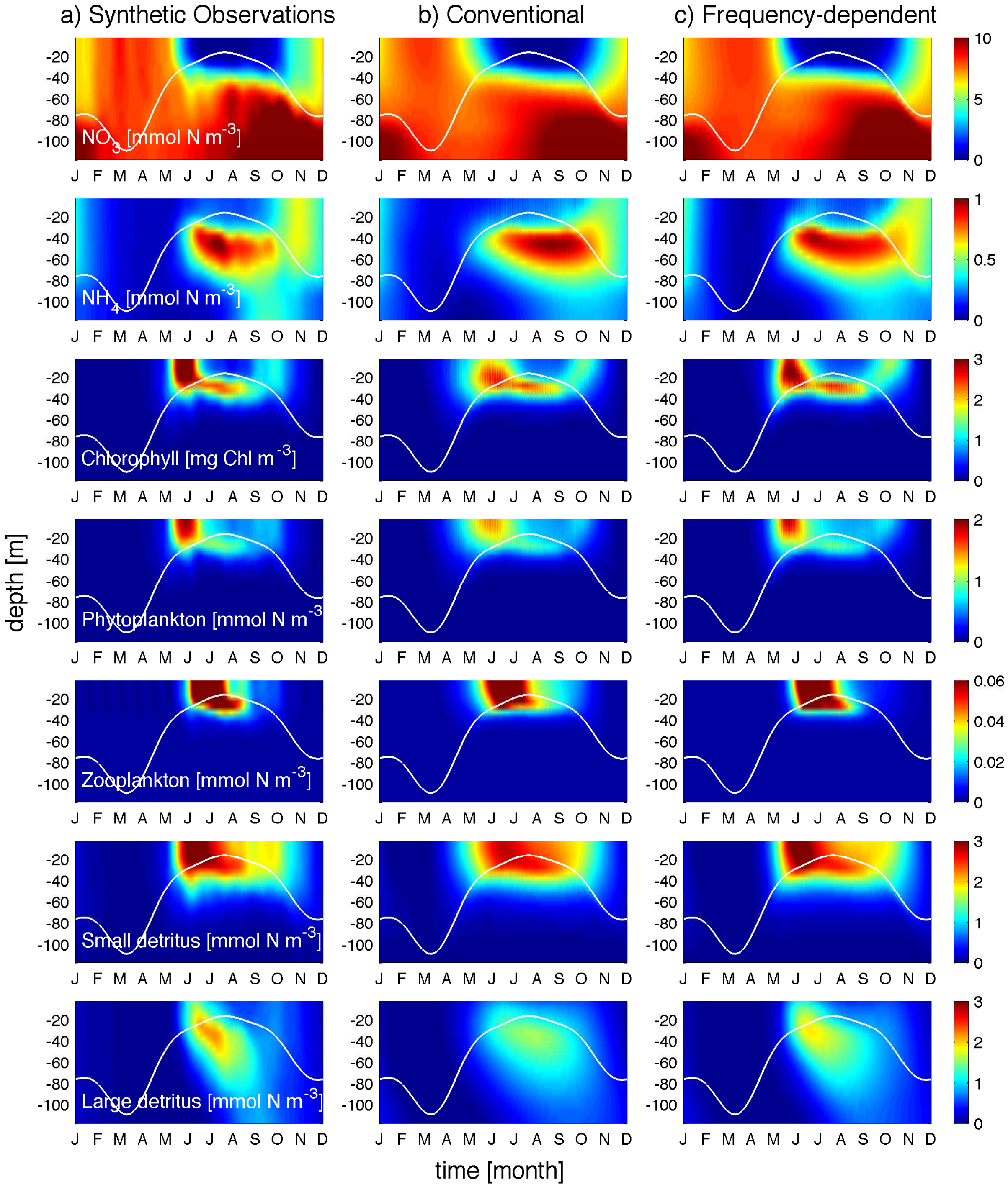


Figure A.8: Same as Figure 2.8 but for station 3.

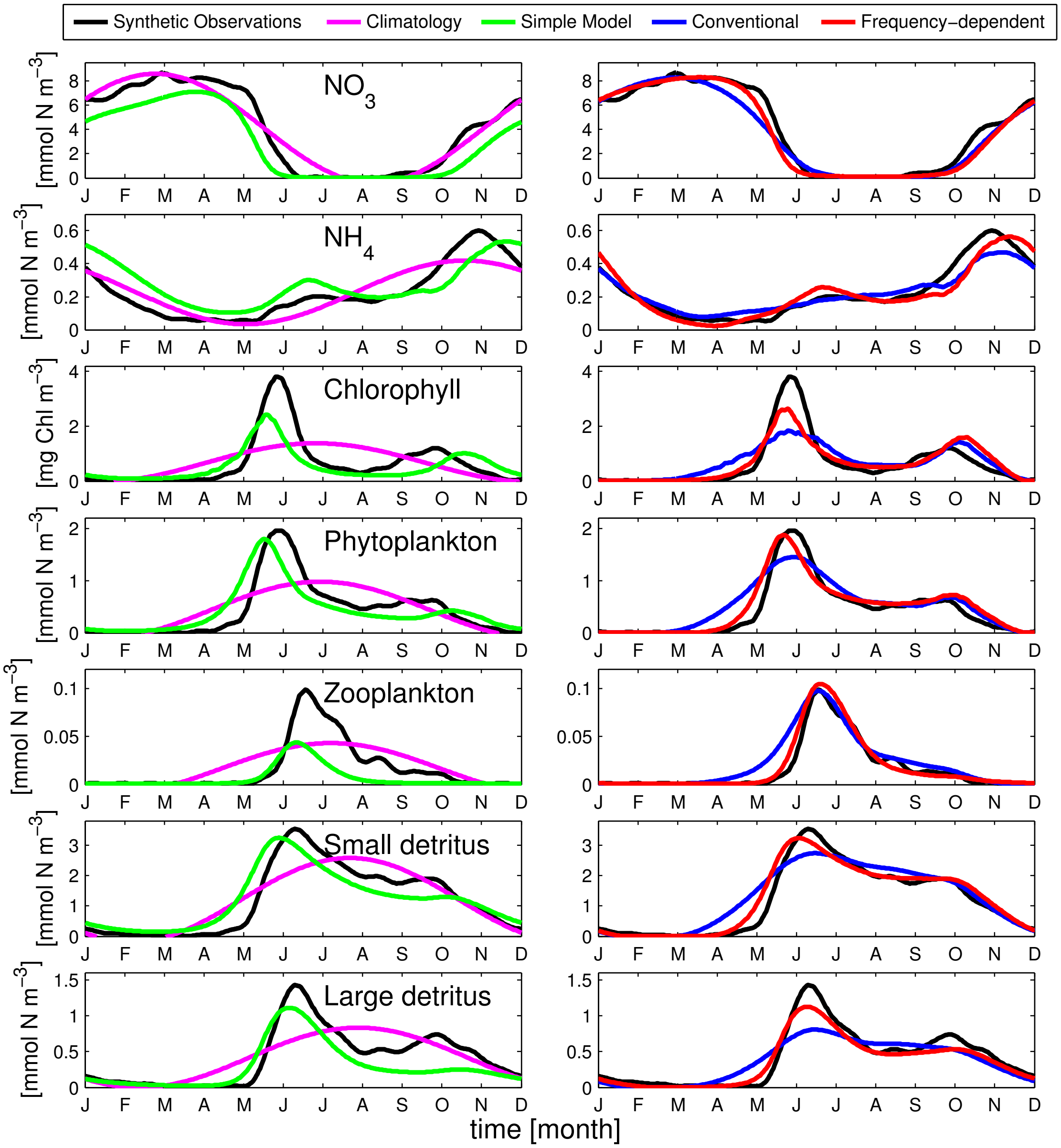


Figure A.9: Same as Figure 2.9 but for station 3.

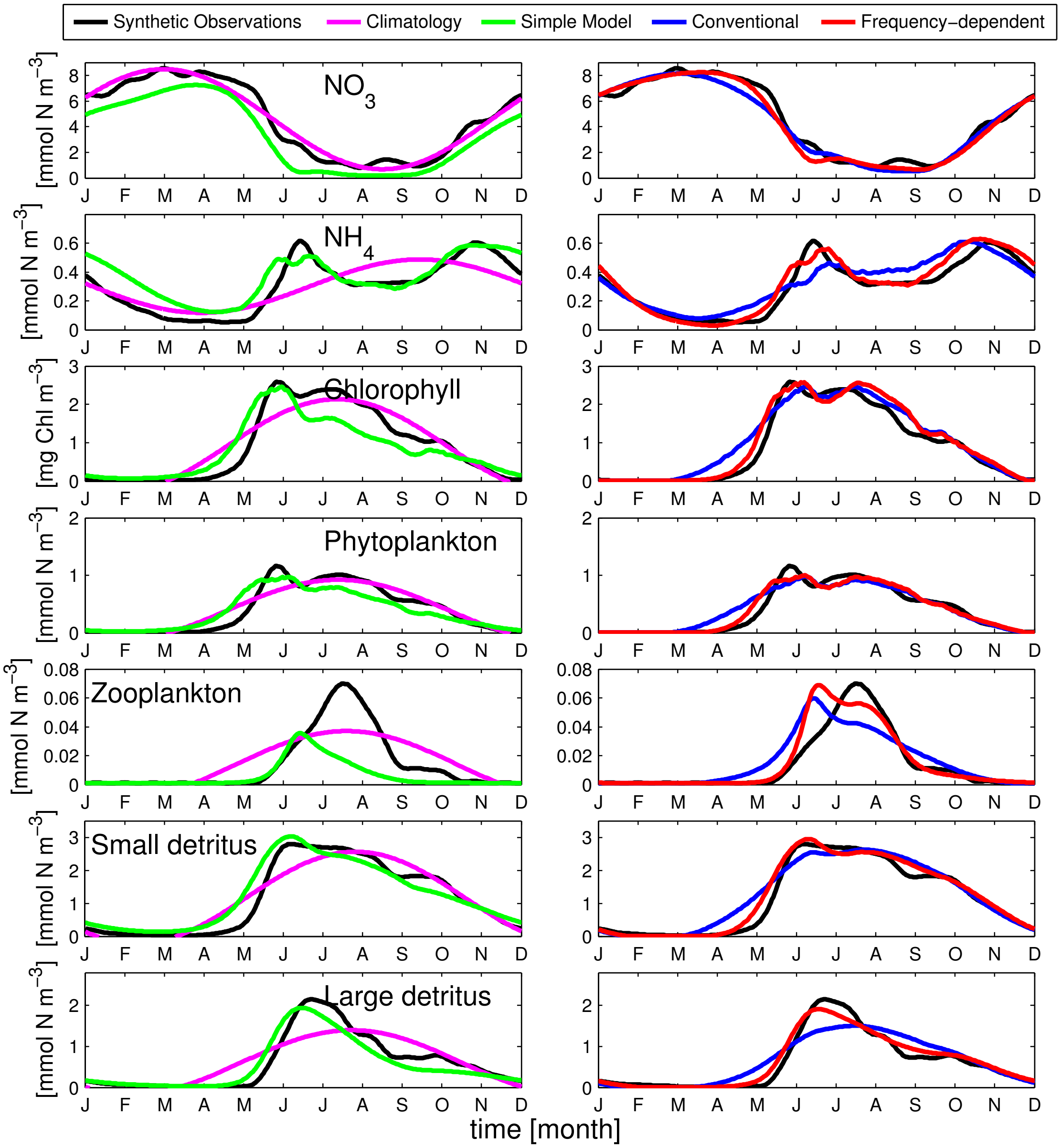


Figure A.10: Same as Figure 2.10 but for station 3.

APPENDIX B

ADDITIONAL TABLES AND FIGURES FOR CHAPTER 3

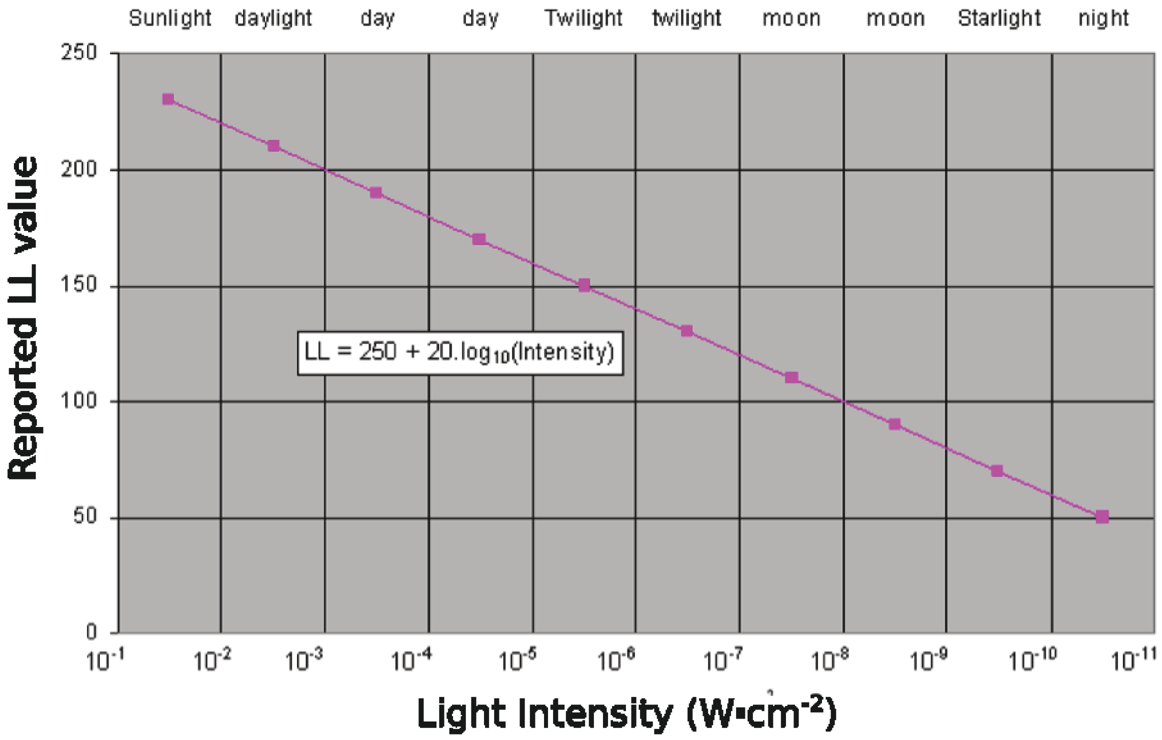


Figure B.1: Logarithmic conversion of the light intensity measured by MK10 loggers to reported light level (LL) values. The relationship is provided by Wildlife Computer, USA. The light sensor is not intended for quantitative measurement of irradiance but rather to simply differentiate between dusk and dawn.

Table B.1: Seal tags deployment information for October 2009-January 2010. Tags, identified by their seal IDs, were attached to seals at Sable Islands in October 2009 and recovered in January 2010. Seals without recovery date means that tags are lost. MK10Serial is the serial number of the electronic tags and MK10Ptt is the corresponding archive filename given to the data collected using the software provided by Wildlife Computers, USA.

SealID	Sex	Deployment Date	MK10Serial	MK10Ptt	Recovery Date
	M	15-Oct-09	9A0794	98432	
	M	13-Oct-09	9A0782	98426	
4057	M	21-Oct-09	9A0804	98437	06-Jan-10
5685	M	22-Oct-09	9A0796	98433	05-Jan-10
6900	M	21-Oct-09	9A0806	98438	04-Jan-10
9022	M	20-Oct-09	9A0807	98439	06-Jan-10
9413	M	13-Oct-09	9A0791	98440	05-Jan-10
104	F	26-Oct-09	9A0792	98431	05-Jan-10
118	F	27-Oct-09	9A0786	98429	02-Jan-10
125	F	28-Oct-09	9A0799	98434	02-Jan-10
4267	F	20-Oct-09	9A0784	98428	31-Dec-09
5682	F	18-Oct-09	9A0783	98427	04-Jan-10
6165	F	30-Oct-09	9A0787	98430	07-Jan-10
9410	F	17-Oct-09	9A0802	98436	04-Jan-10
9411	F	16-Oct-09	9A0800	98435	04-Jan-10

Table B.2: Seal tags deployment information for September 2010-February 2011

SealID	Sex	Deployment Date	MK10Serial	MK10Ptt	Recovery Date
9929	F	10-Sep-10	10A549	66487	1-Jan-11
9930	F	10-Sep-10	10A565	66476	1-Jan-11
9937	F	17-Sep-10	10A585	66556	2-Jan-11
9938	F	18-Sep-10	10A582	66541	31-Dec-10
4515	M	11-Sep-10	10A579	66505	15-Dec-10
6122	F	11-Sep-10	10A071	66555	18-Dec-10
9933	F	11-Sep-10	10A577	66501	2-Jan-11
9931	F	10-Sep-10	10A575	66498	1-Jan-11
3648	M	15-Sep-10	10A572	66489	2-Jan-11
9932	F	10-Sep-10	10A578	66504	18-Dec-10
9934	F	11-Sep-10	10A567	66477	1-Jan-11
3271	F	08-Sep-10	10A573	66490	1-Jan-11
5954	M	16-Sep-10	10A559	66494	17-Dec-10
9928	F	09-Sep-10	10A580	66506	2-Jan-11
6195	M	13-Sep-10	10A571	66488	2-Jan-11
9936	M	13-Sep-10	10A576	66499	1-Jan-11
9935	M	12-Sep-10	10A568	66479	3-Jan-11
78	F	18-Sep-10	10A569	66486	17-Jan-11
9414	F	09-Sep-10	10A581	66508	1-Jan-11
9939	F	18-Sep-10	10A583	66548	1-Jan-11

Table B.3: Seal tags deployment information for June 2011-January 2012

SealID	Sex	Deployment Date	MK10Serial	MK10Ptt	Recovery Date
10334	F	15-Jun-11	11A090	106705	19-Jan-12
10322	F	11-Jun-11	11A091	106706	20-Dec-11
10332	F	12-Jun-11	11A151	106707	
6120	F	15-Jun-11	11A213	106708	23-Dec-11
10330	F	14-Jun-11	11A214	106709	30-Dec-11
3736	F	15-Jun-11	11A215	106710	28-Dec-11
10323	F	11-Jun-11	11A216	106711	
10327	F	12-Jun-11	11A217	106712	03-Jan-12
10333	F	15-Jun-11	11A245	106713	01-Jan-12
5846	F	14-Jun-11	11A251	106714	01-Jan-12
137	F	15-Jun-11	11A252	106715	31-Dec-11
10324	F	11-Jun-11	11A253	106716	16-Jan-12
10331	F	14-Jun-11	11A254	106717	27-Dec-11
10328	F	12-Jun-11	11A256	106718	07-Jan-12
2999	F	12-Jun-11	11A257	106719	31-Dec-11
10329	F	13-Jun-11	11A258	106720	
6630	F	14-Jun-11	11A260	106721	03-Jan-12
2574	F	12-Jun-11	11A261	106722	
10325	F	12-Jun-11	11A262	106723	30-Dec-11
2718	F	14-Jun-11	11A267	106724	10-Jan-12

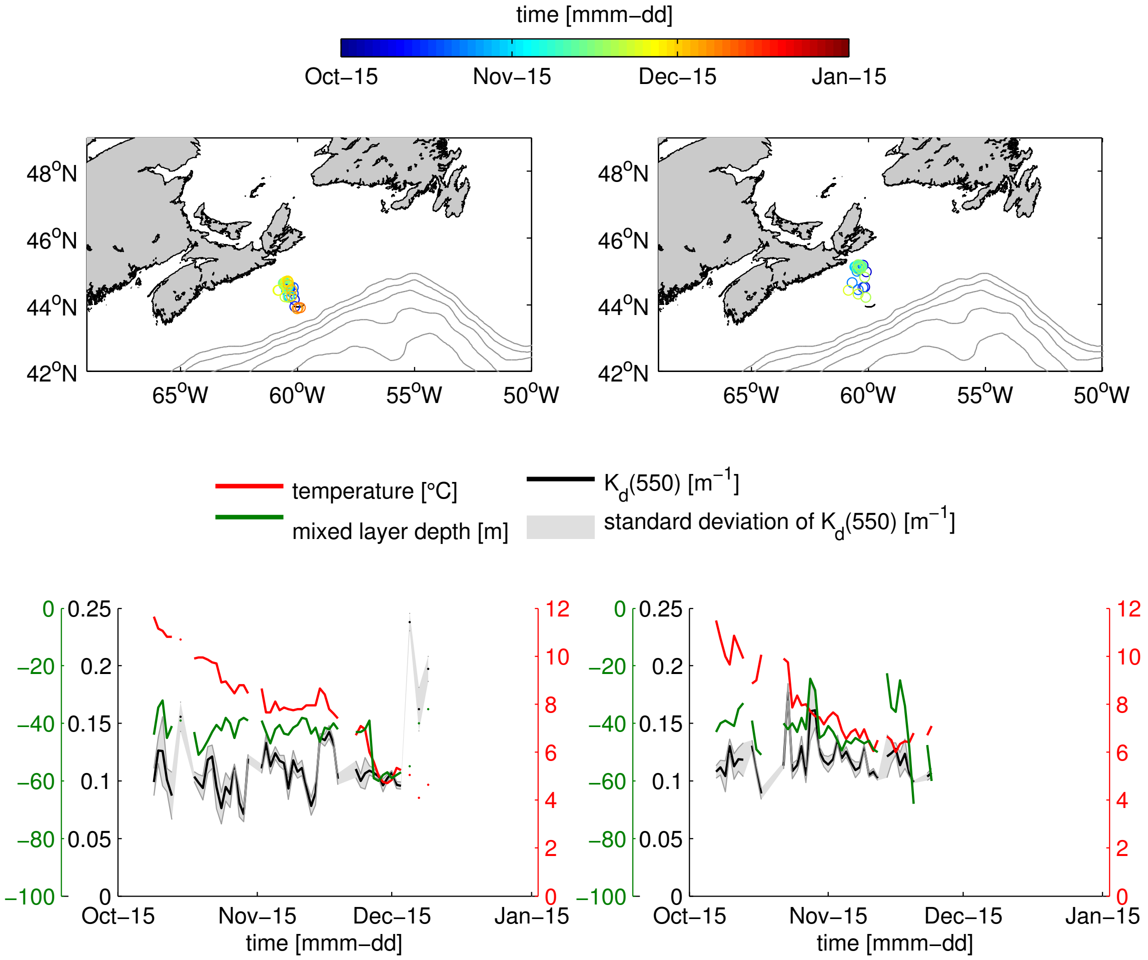


Figure B.2: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 5682 (left) and 4267 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

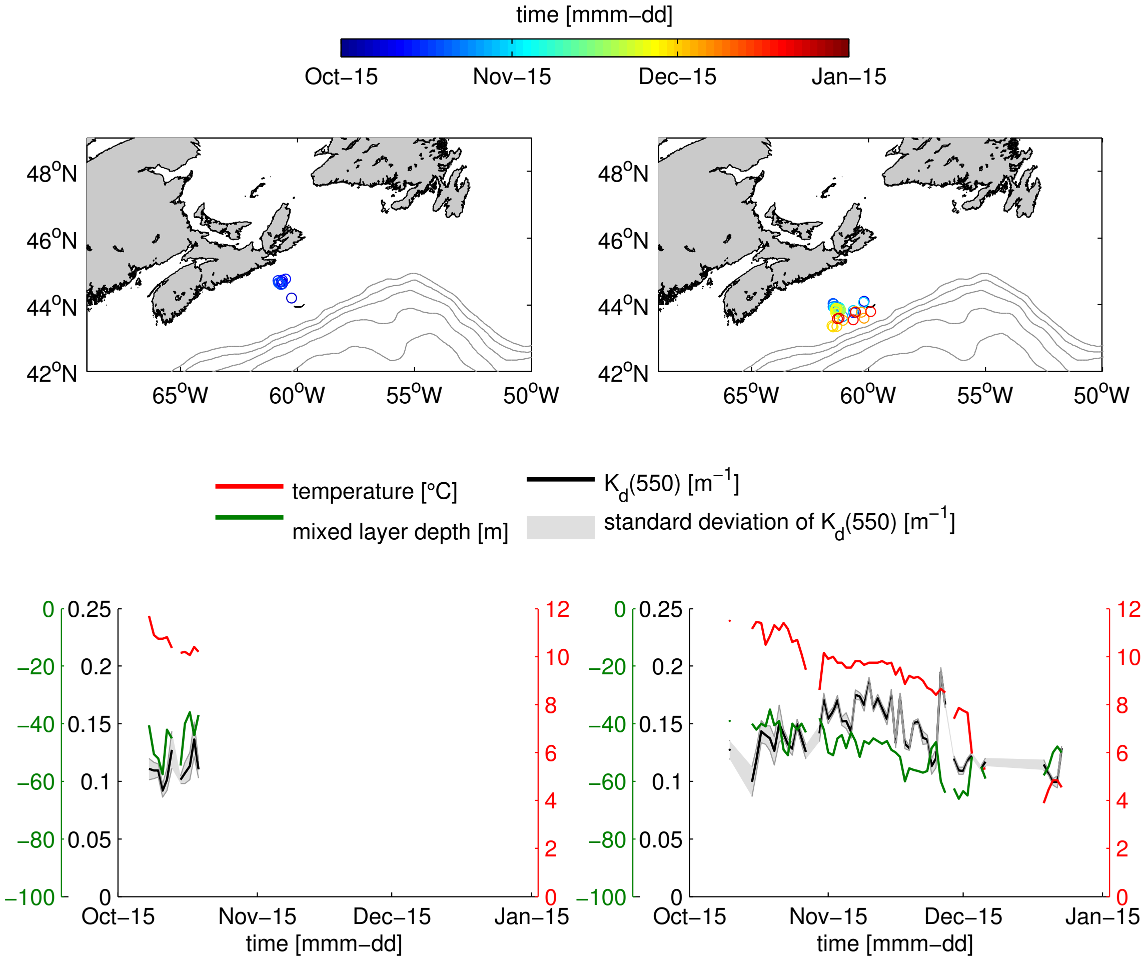


Figure B.3: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 118 (left) and 6165 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

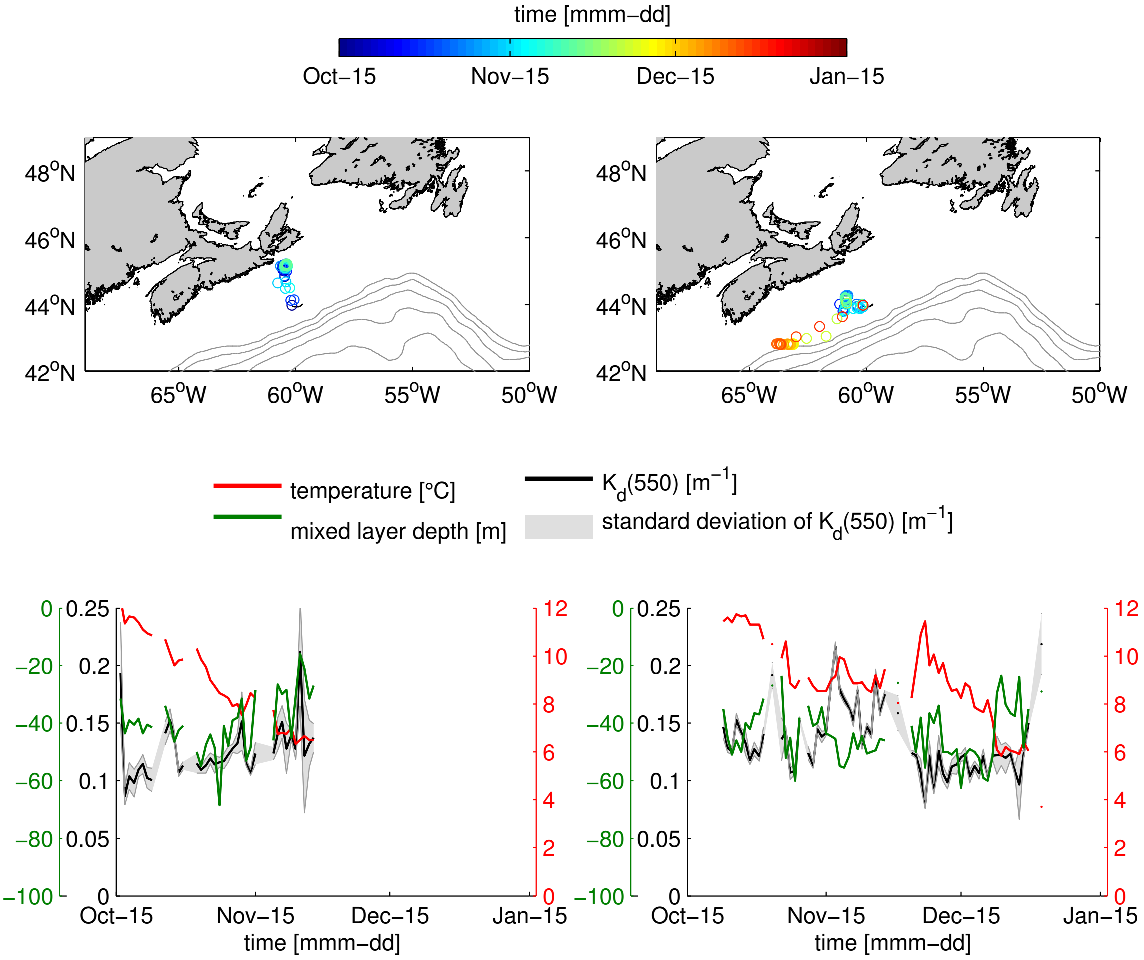


Figure B.4: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 104 (left) and 5685 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

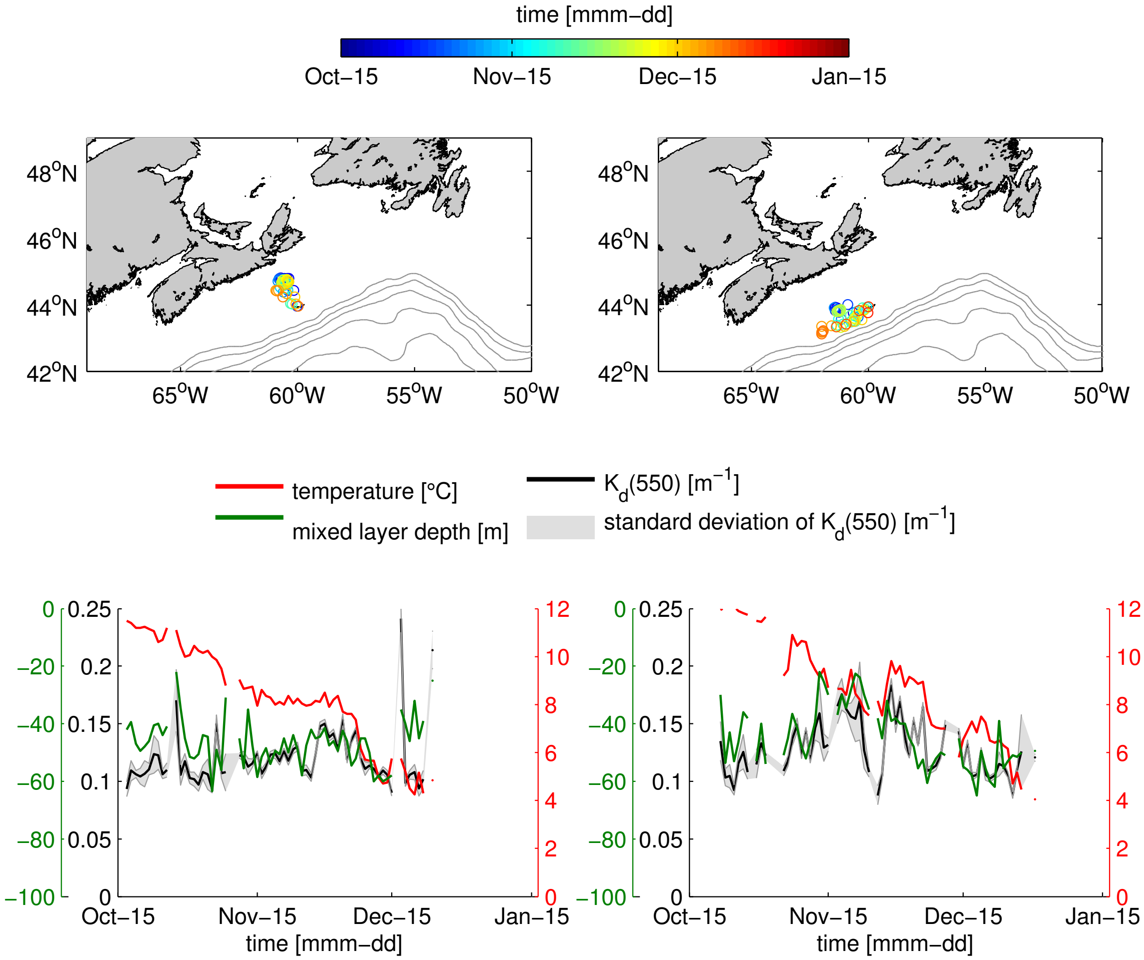


Figure B.5: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9410 (left) and 4057 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

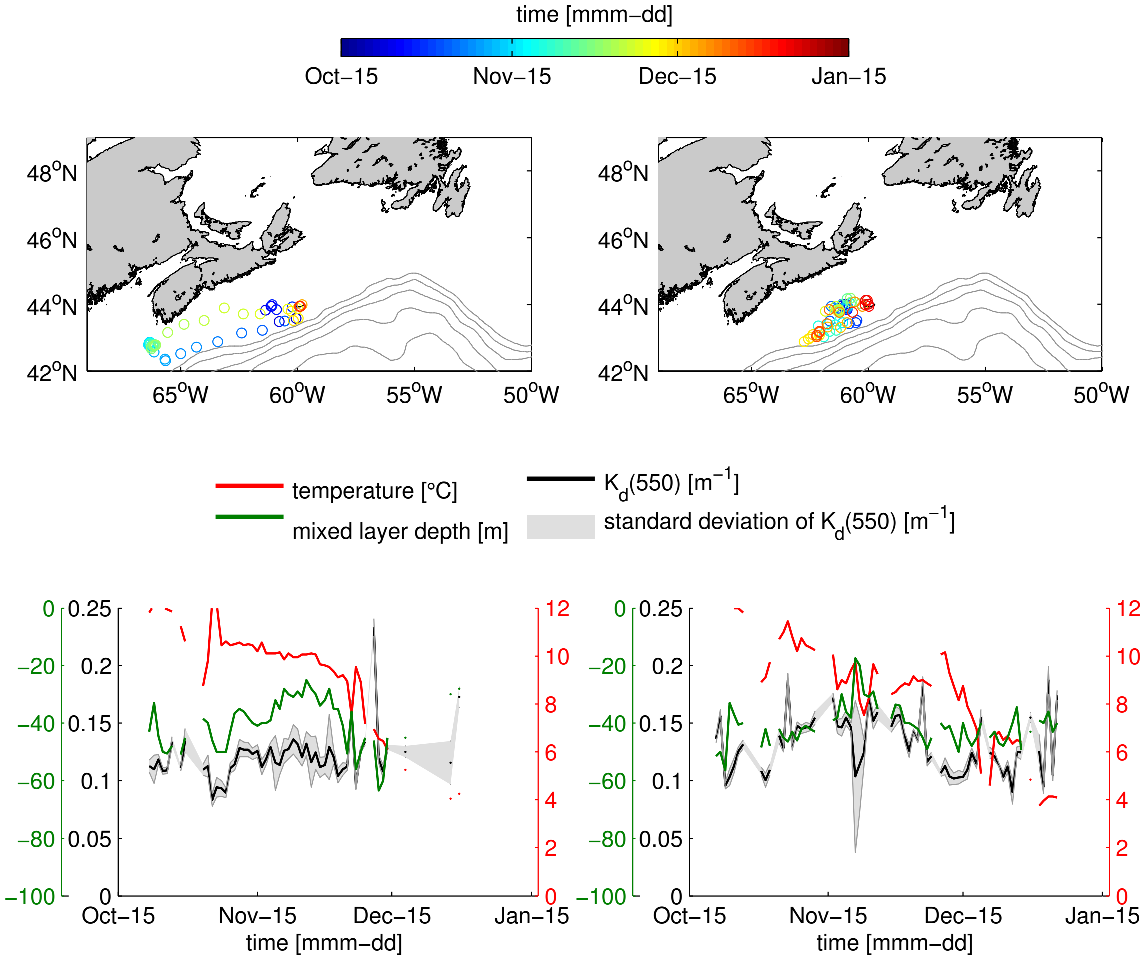


Figure B.6: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9022 (left) and 6911 (right)] from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

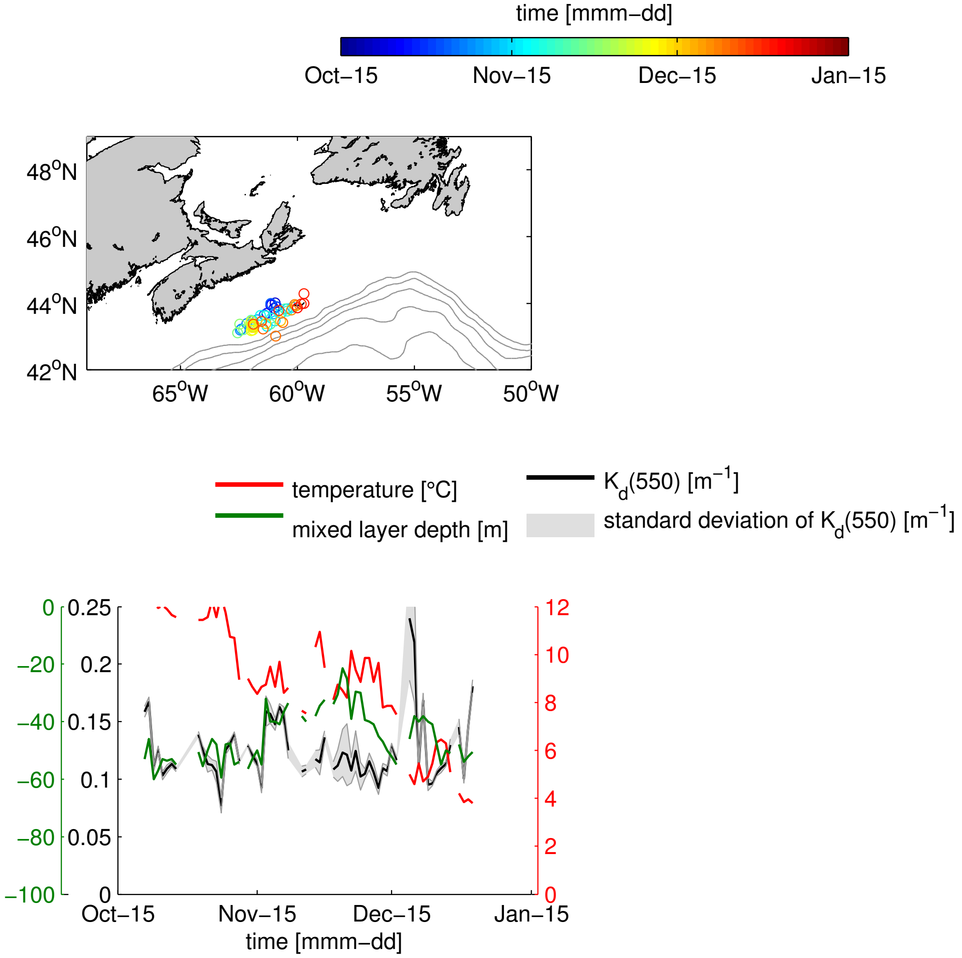


Figure B.7: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for one seal (sealID 9413) from Figure 3.12a. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

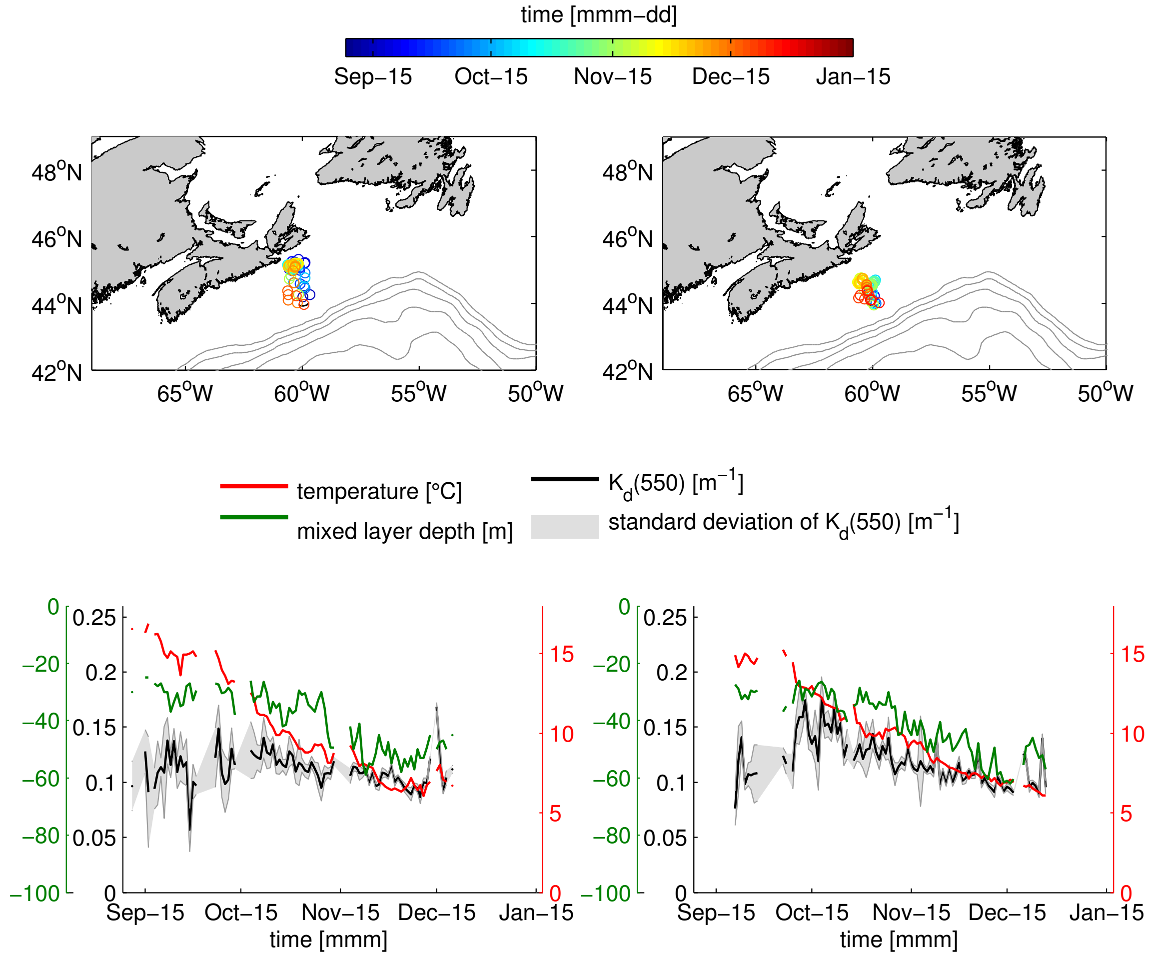


Figure B.8: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9930 (left) and 9934 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

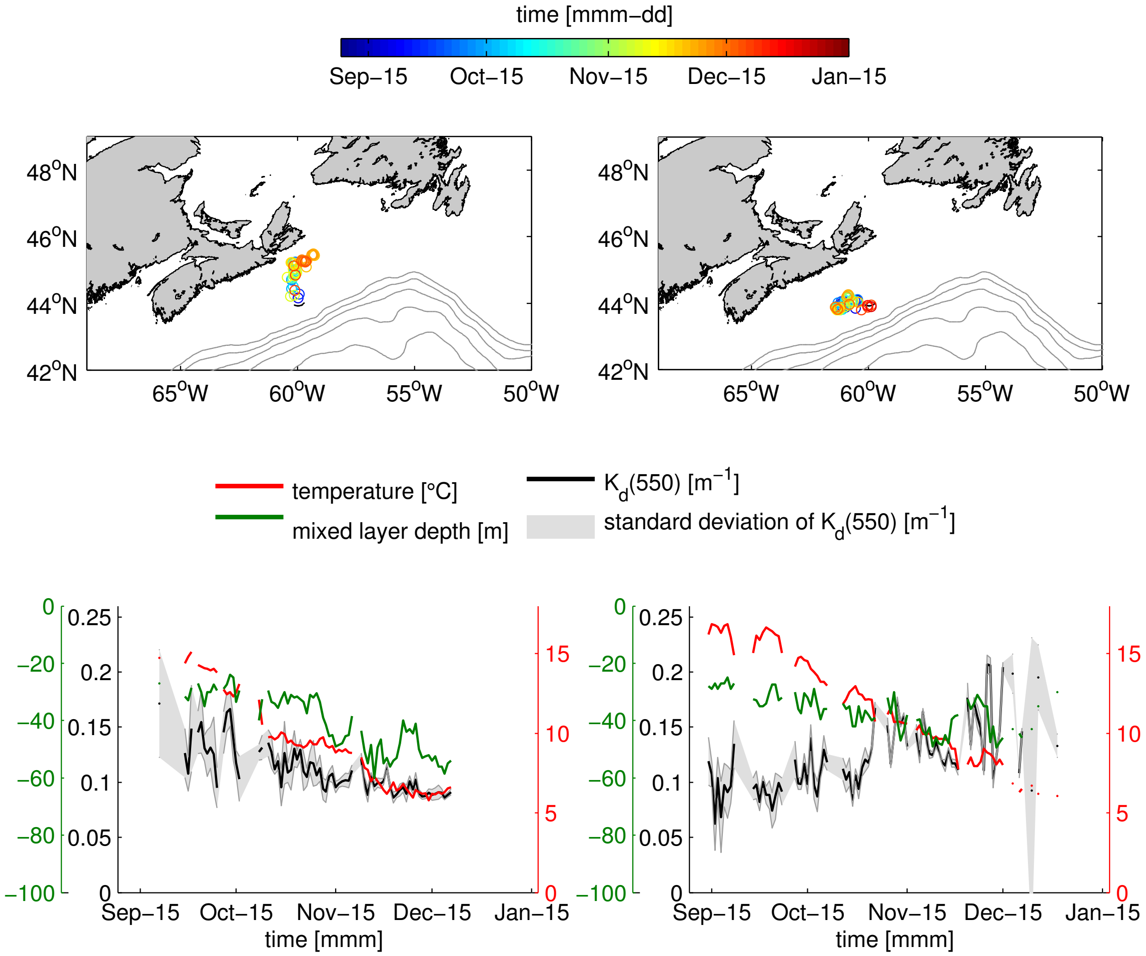


Figure B.9: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9929 (left) and 6195 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

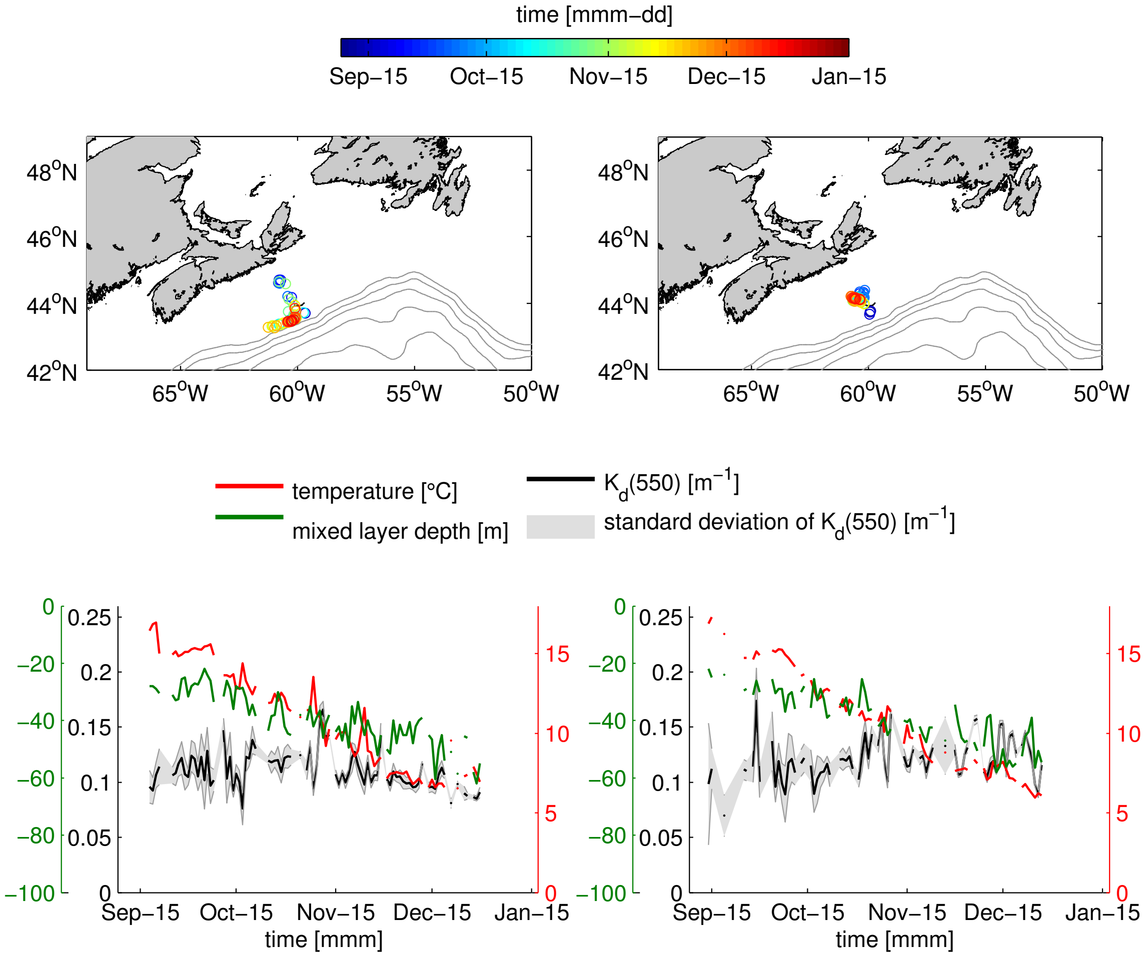


Figure B.10: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 3648 (left) and 3271 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

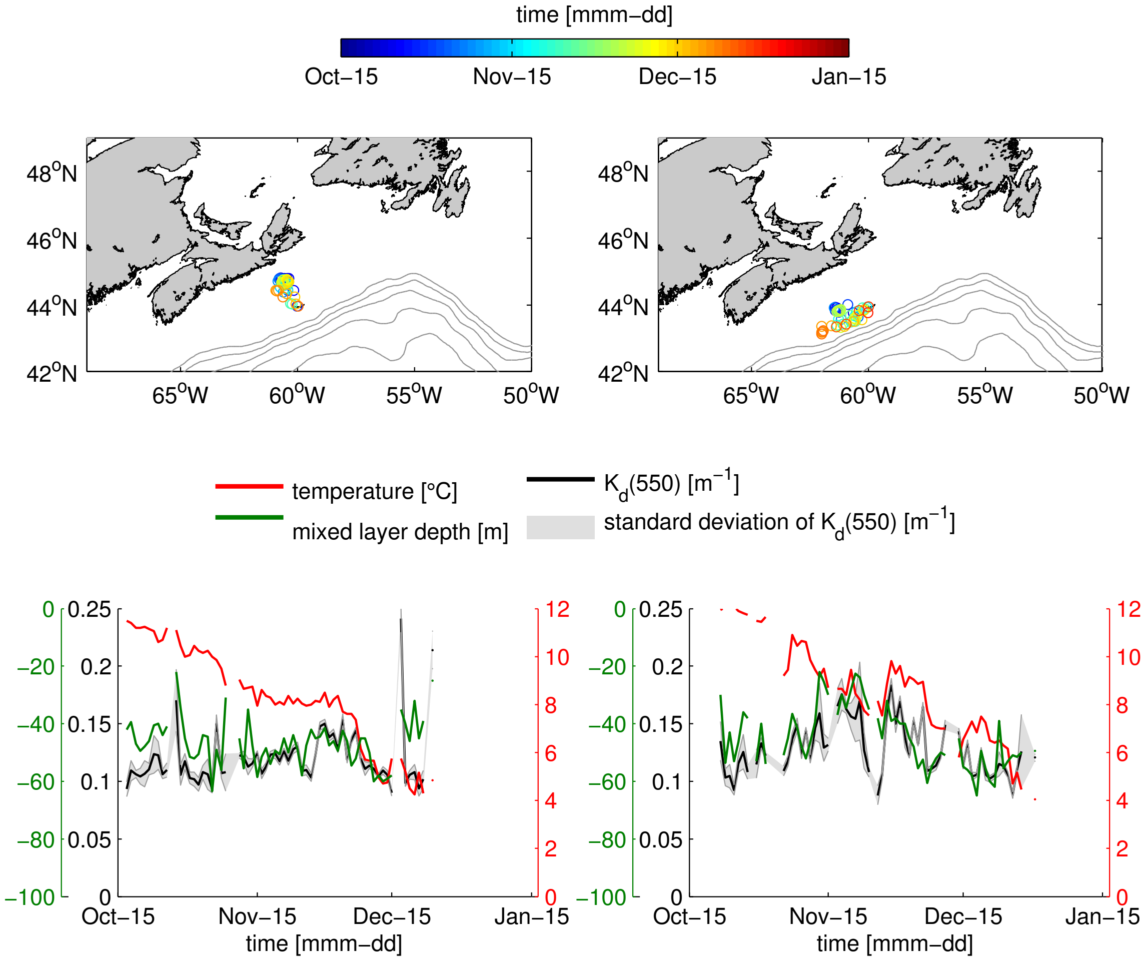


Figure B.11: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9931 (left) and 9936 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

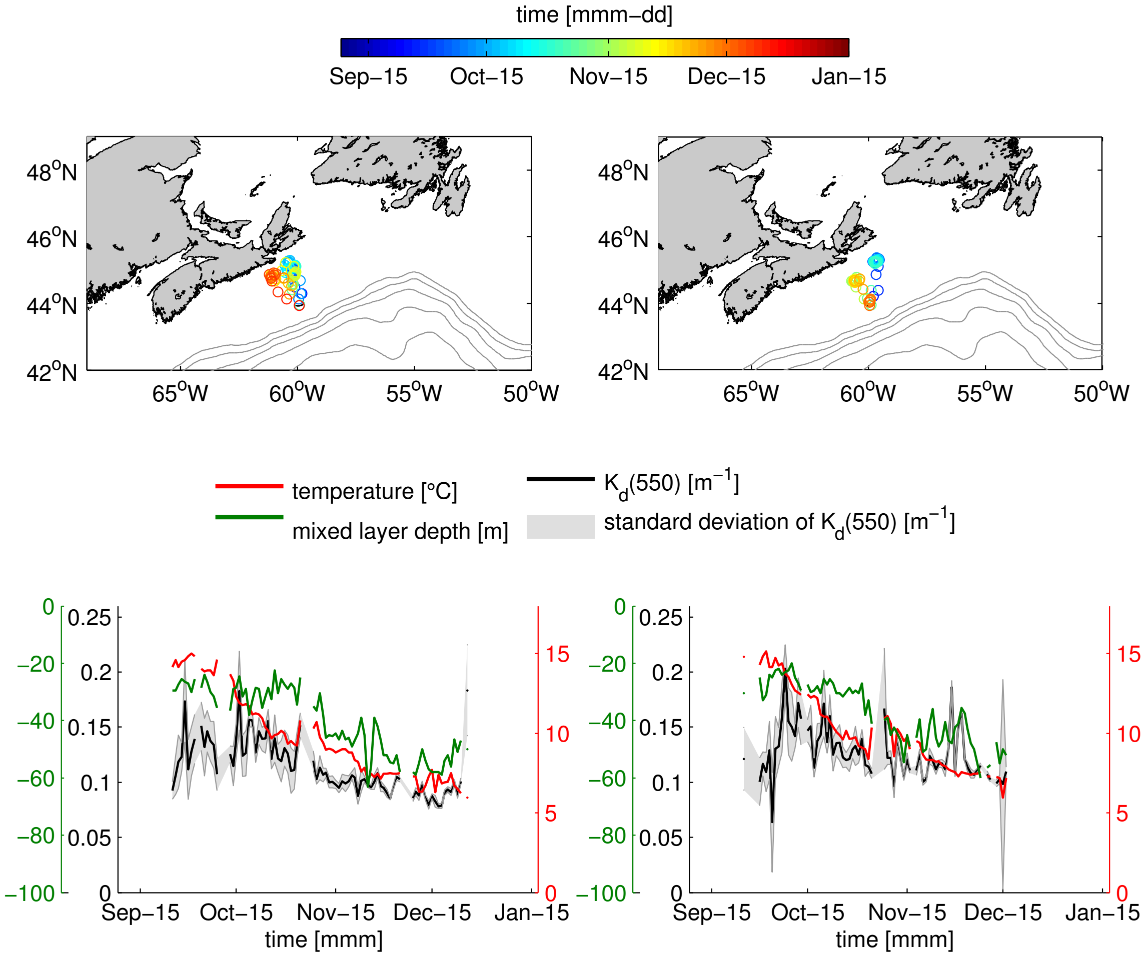


Figure B.12: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9933 (left) and 9932 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

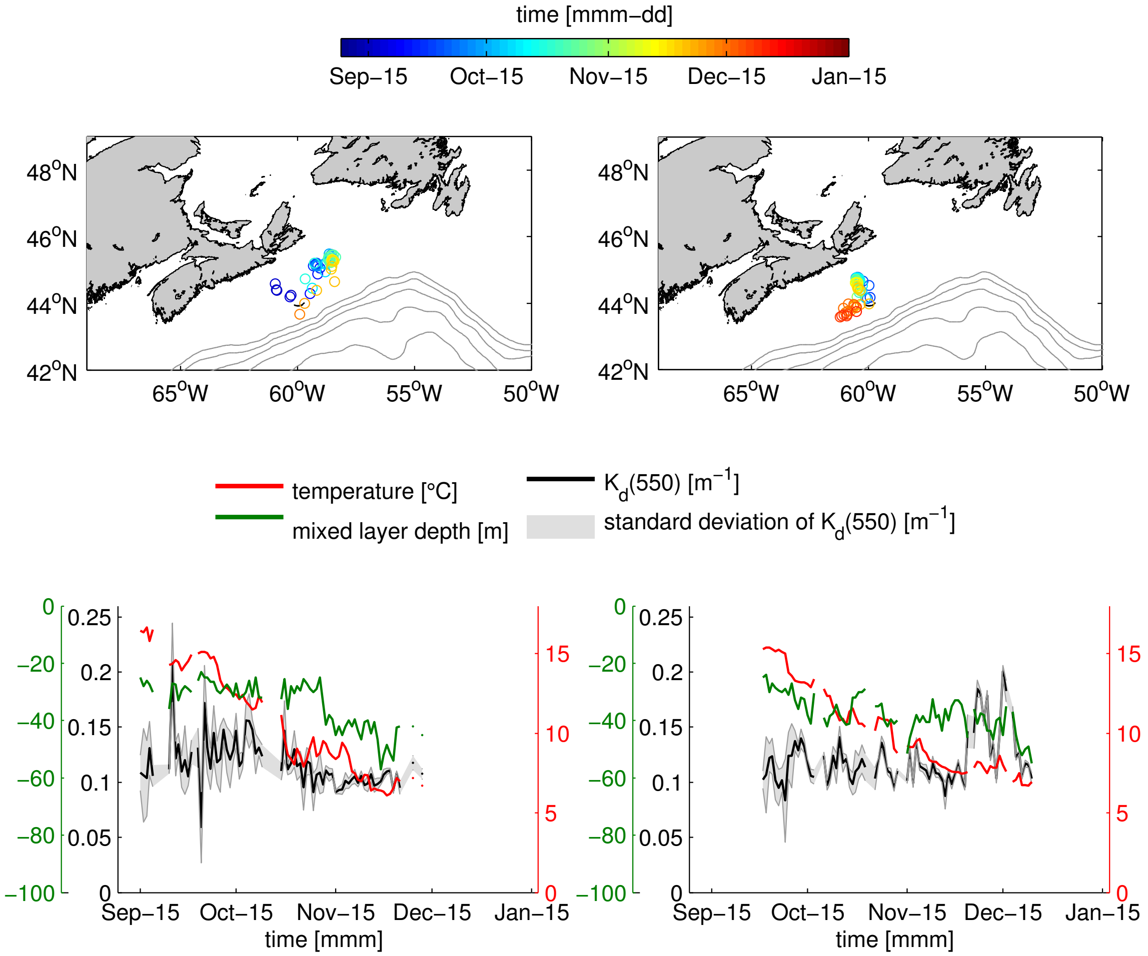


Figure B.13: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 4515 (left) and 9928 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

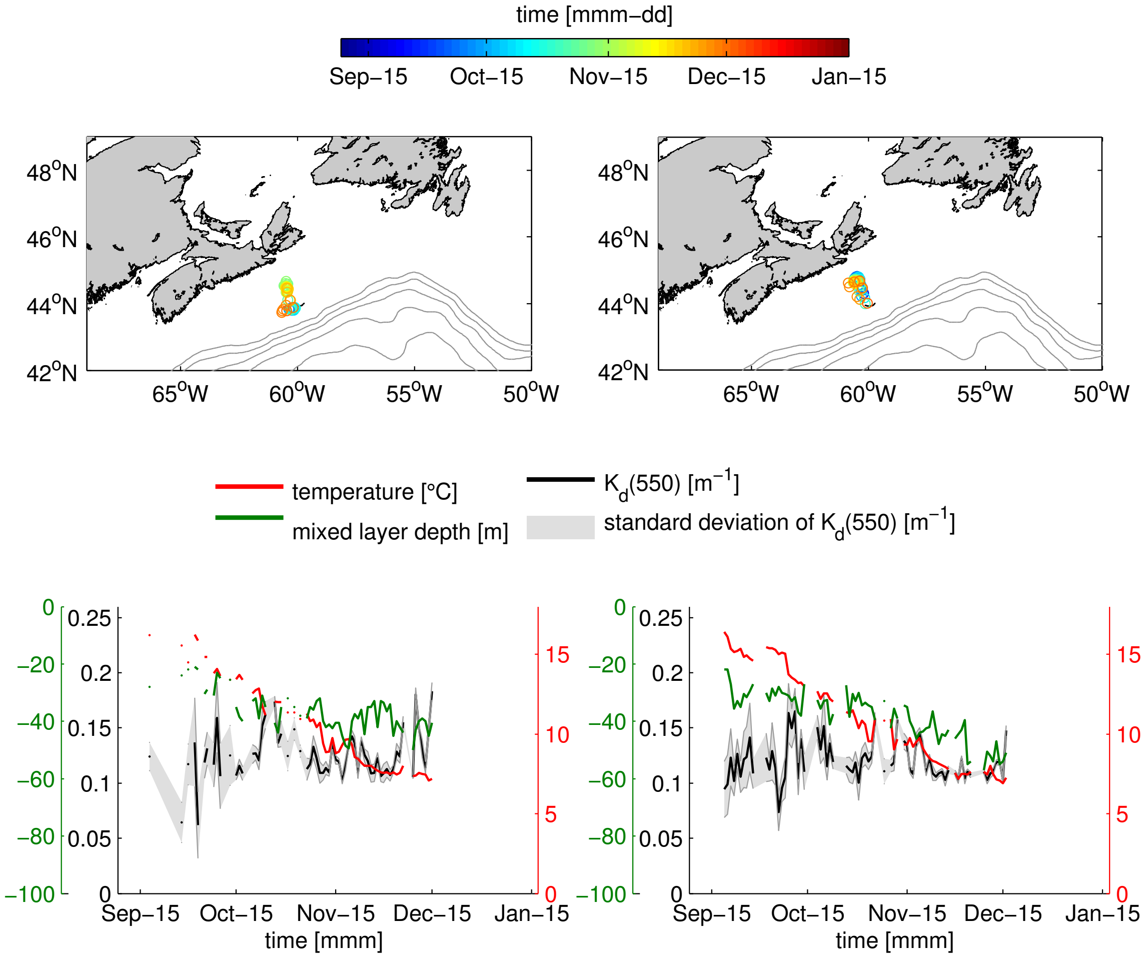


Figure B.14: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9414 (left) and 9938 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

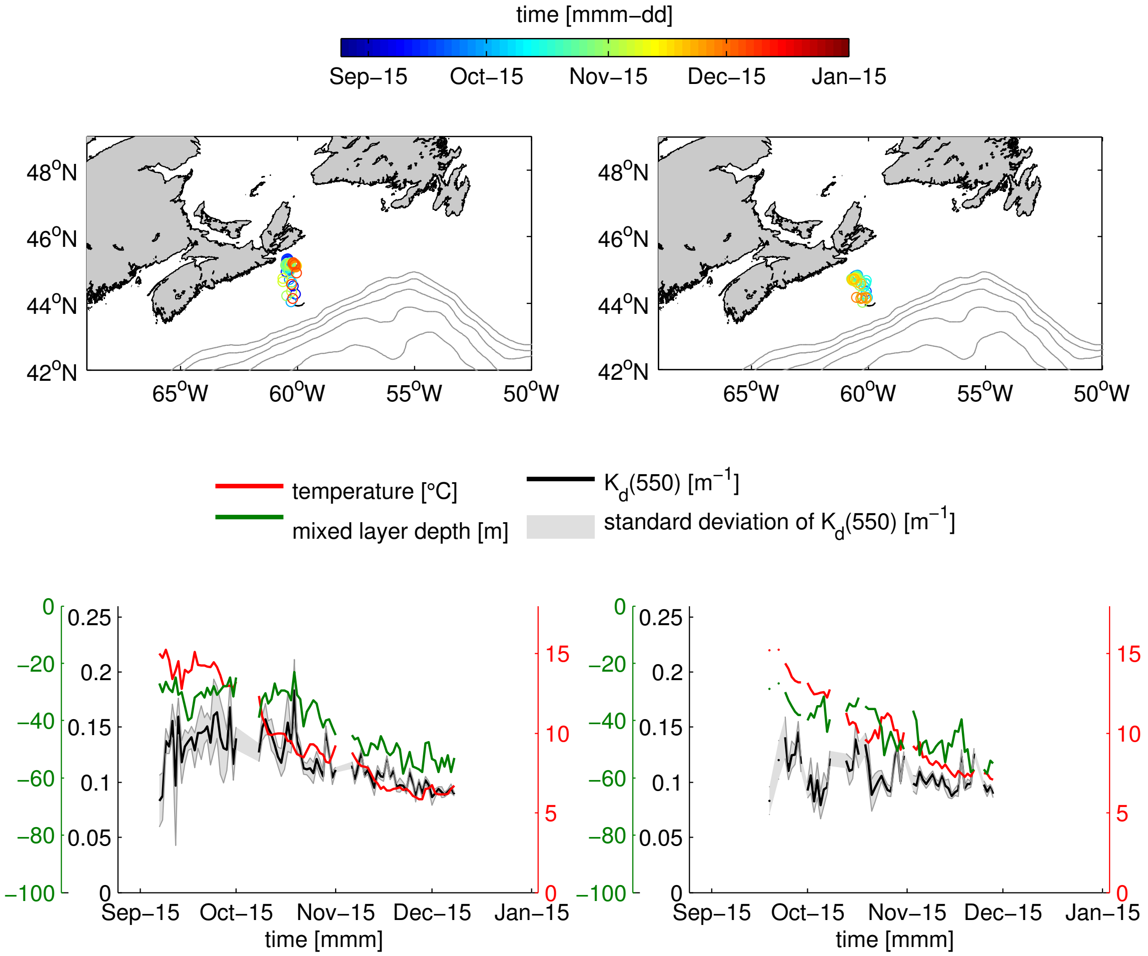


Figure B.15: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 9939 (left) and 6122 (right)] from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

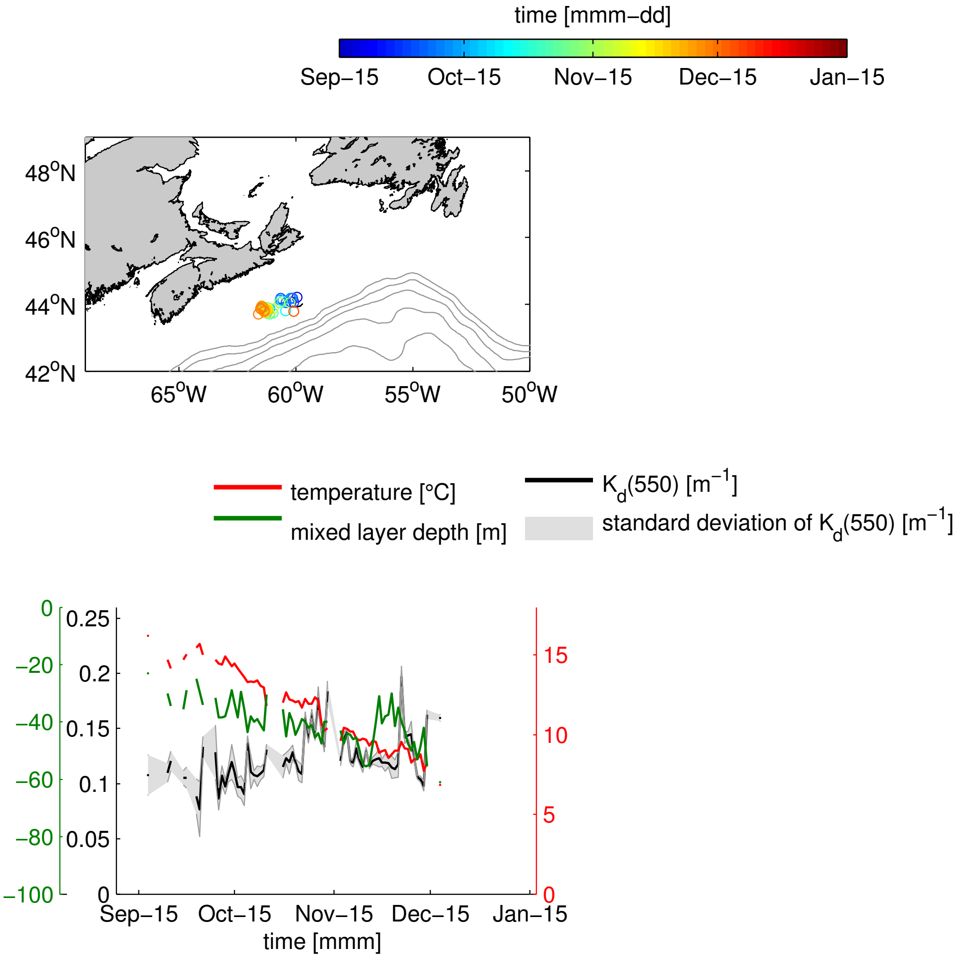


Figure B.16: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for one seal (sealID 9937) from Figure 3.12b. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

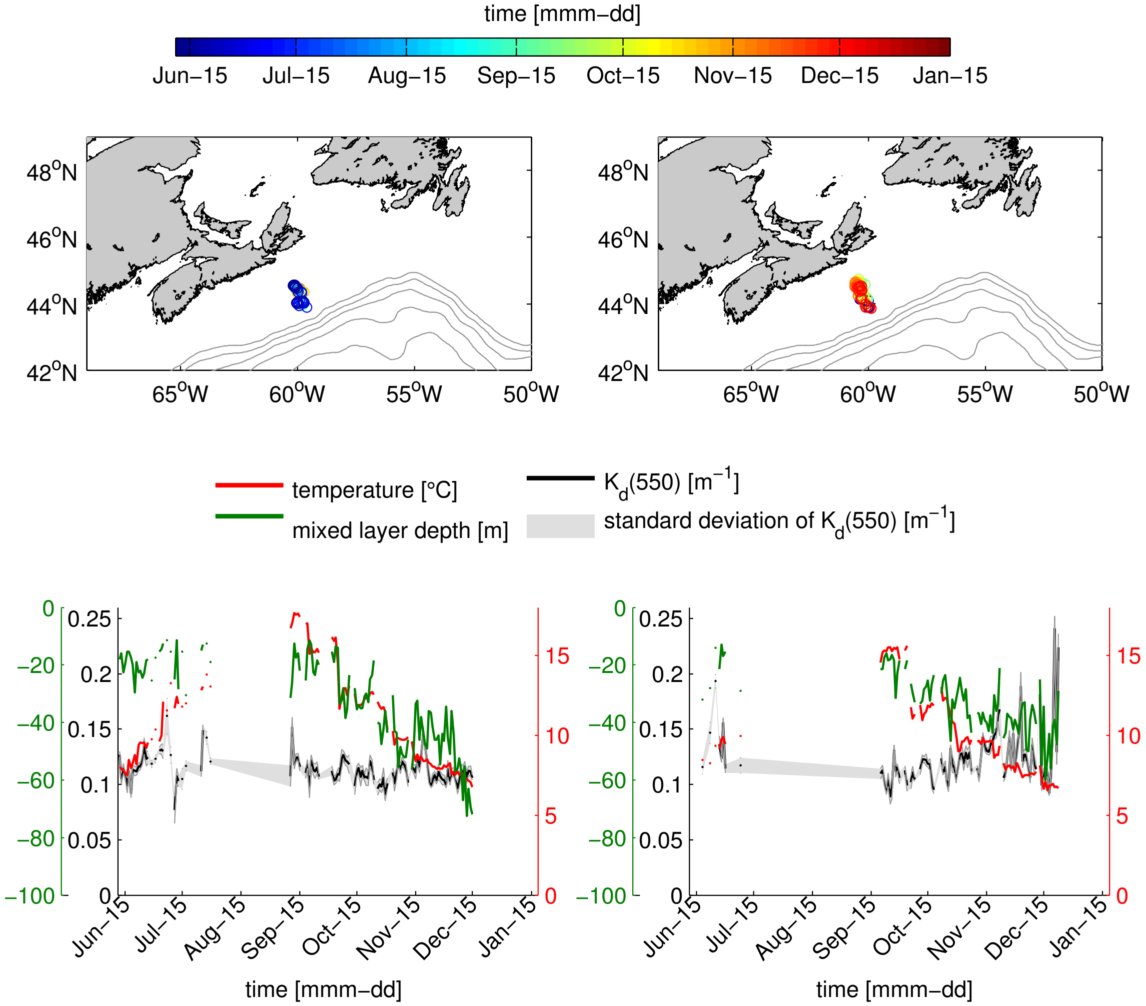


Figure B.17: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 10322 (left) and 10331 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

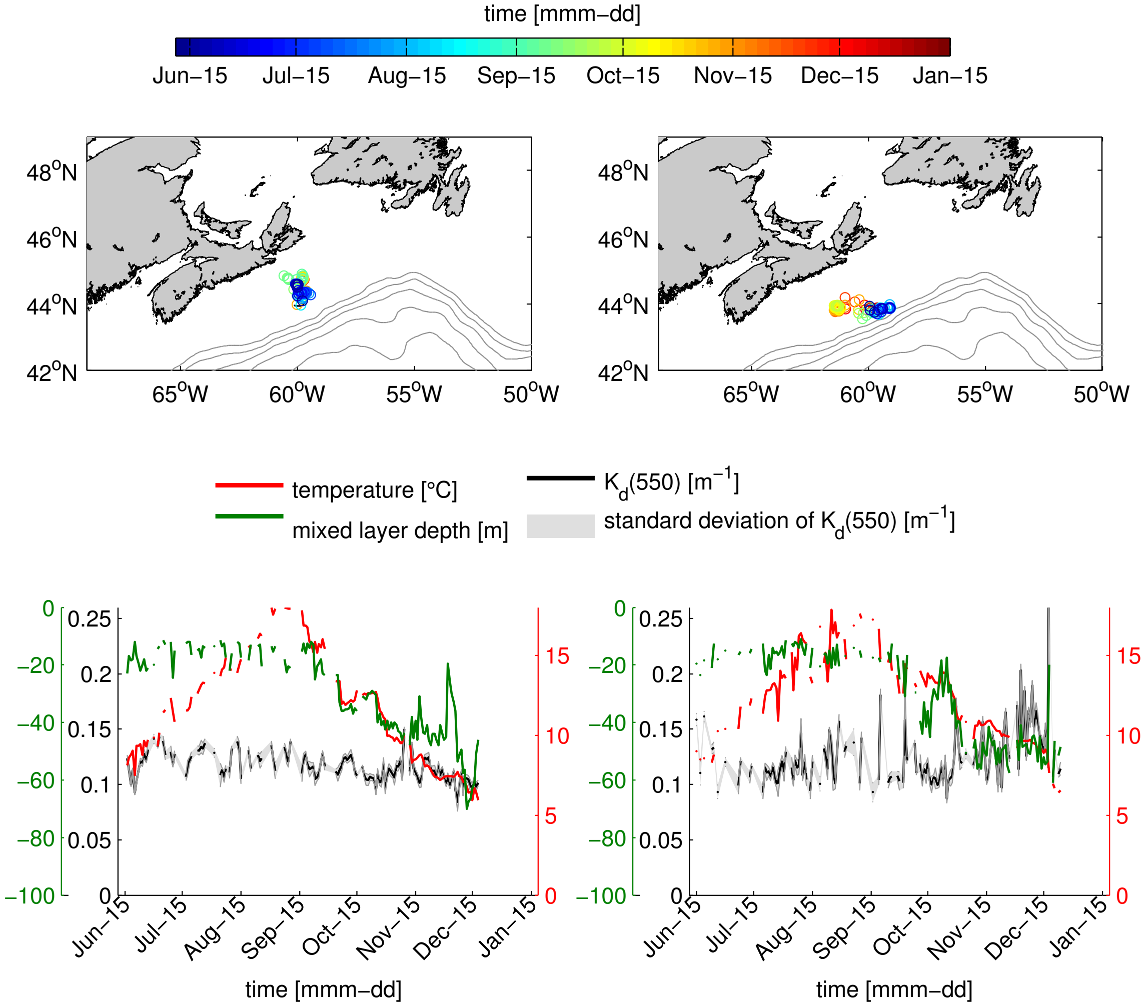


Figure B.18: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 10327 (left) and 5846 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

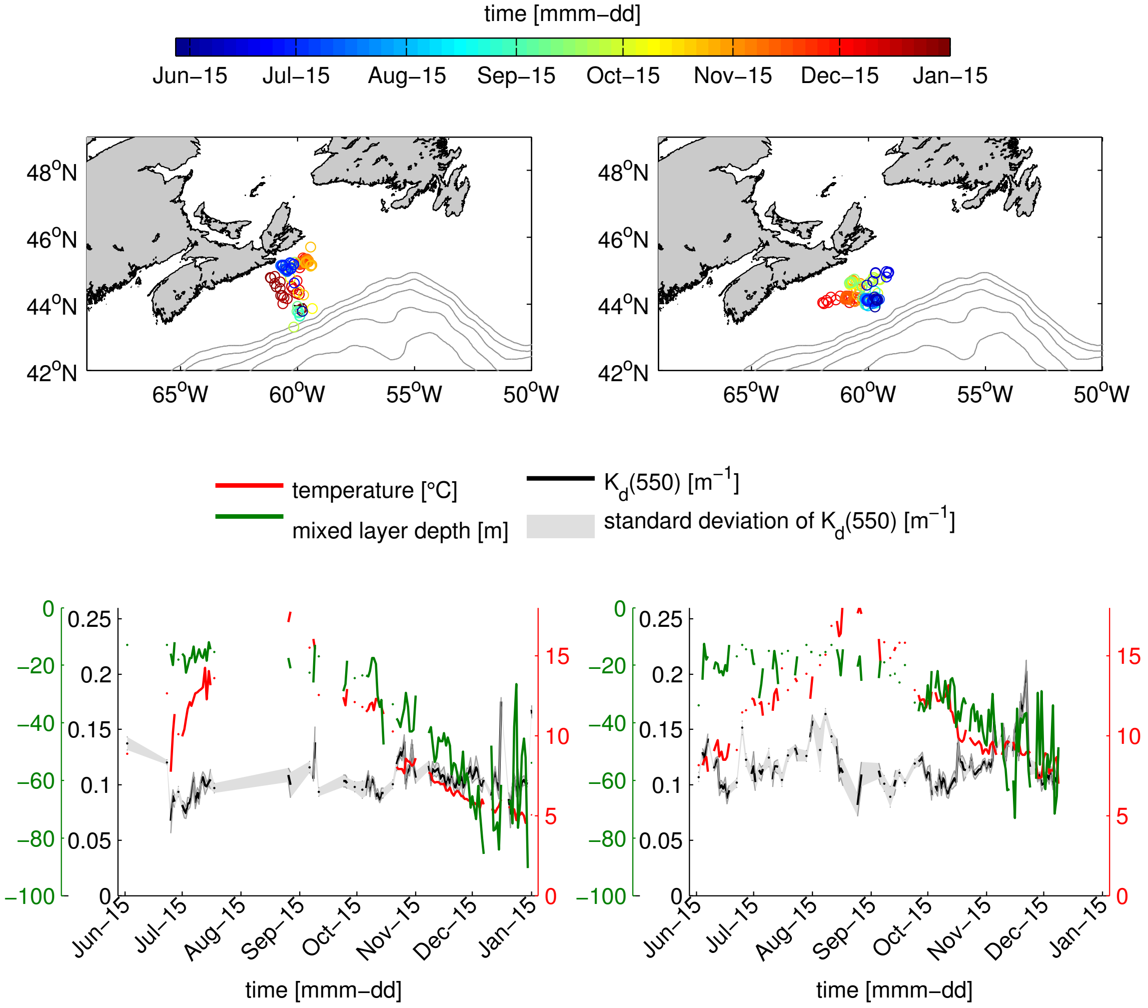


Figure B.19: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 10334 (left) and 10330 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

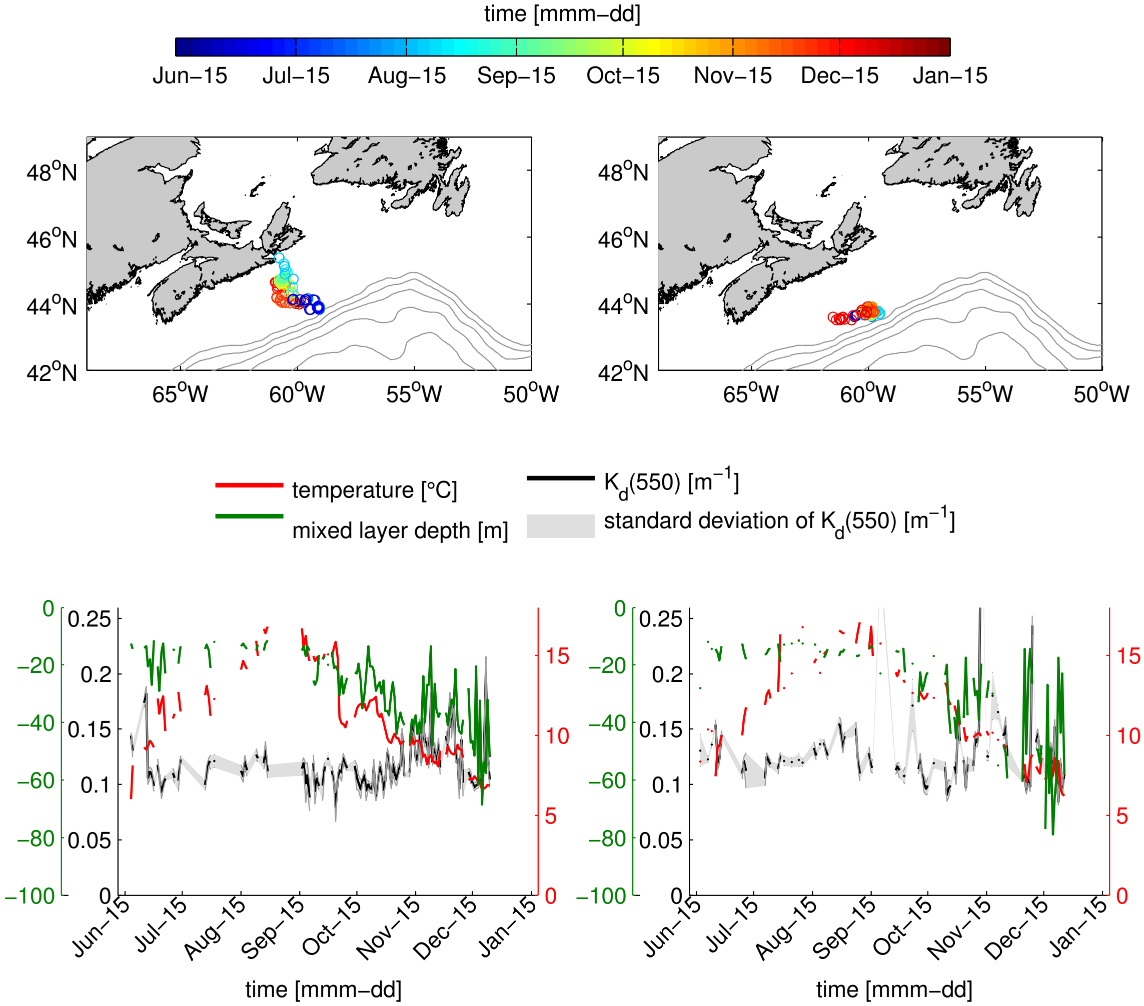


Figure B.20: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 137 (left) and 10331 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

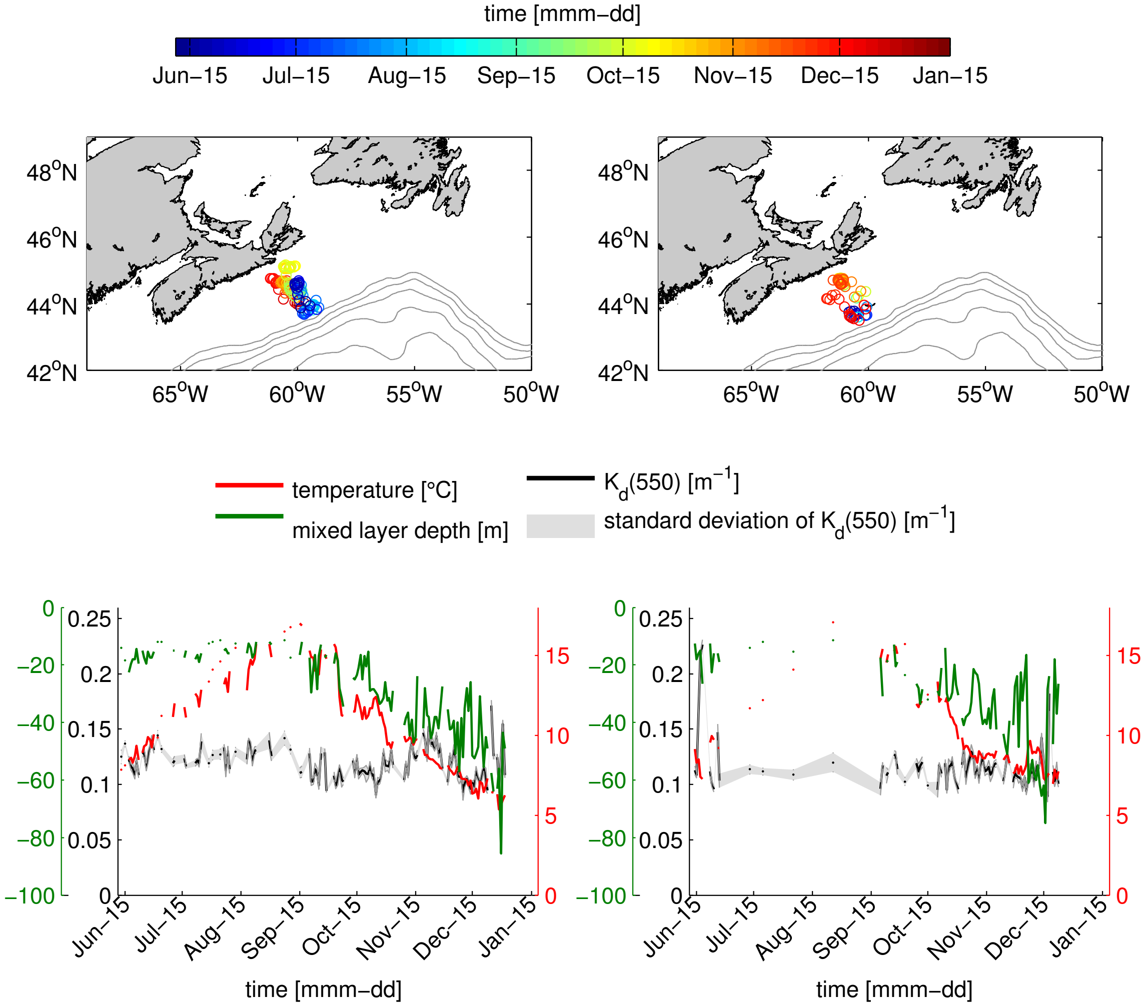


Figure B.21: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for two seals [sealID 10328 (left) and 2999 (right)] from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

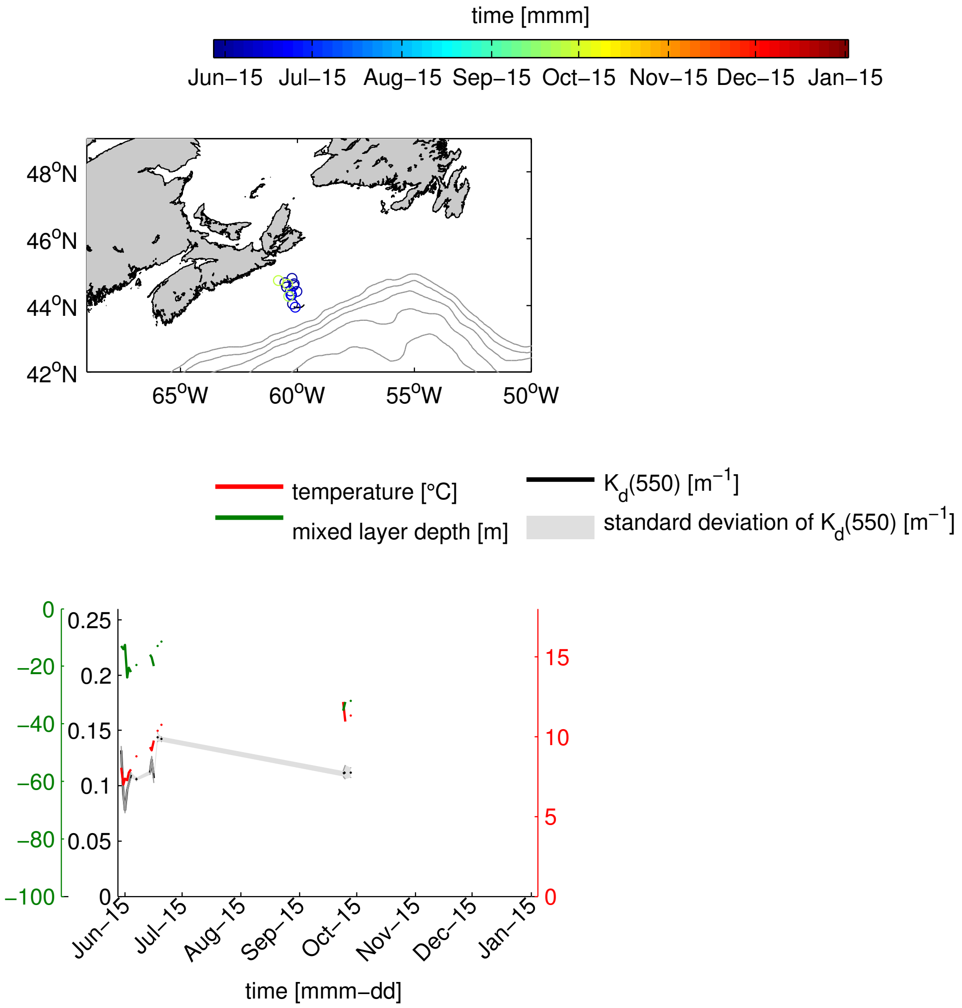


Figure B.22: Spatio-temporal distribution of the estimated $K_d(550)$ (top row) for one seal (sealID 10325) from Figure 3.12c. The corresponding time series (black line), calculated mixed layer depth (mld, green line) and mean temperature (calculated up to mld, red line) are shown below.

BIBLIOGRAPHY

Allen, J., M. Eknes, and G. Evensen, An ensemble Kalman Filter with a complex marine ecosystem model: hind casting phytoplankton in the Cretan Sea, *Ann. Geophys.*, *21*, 399–411, 2003.

Andersen, J., L. Schluter, and G. Aertebjerg, Coastal eutrophication: recent developments in definitions and implications for monitoring strategies, *Journal of Plankton Research*, *28*, 621–628, 2006.

Antoine, D., A. Morel, A. Gordon, V. Banzon, and R. Evans, Bridging ocean color observations of the 1980's and 2000's in search of long-term trends, *J. Geophys. Res.*, *110*, 2005.

Balch, W., D. Drapeau, B. Bowler, E. Booth, J. Goes, A. Ashe, and J. Frye, A multi-year record of hydrographic and bio-optical properties in the Gulf of Maine: I. Spatial and temporal variability, *Prog. Ocea.*, *63*, 57–98, 2004.

Barlow, R., J. Aiken, P. Holligan, D. Cummings, S. Maritorena, and S. Hooker, Phytoplankton pigment and absorption characteristics along meridional transects in the Atlantic Ocean, *Deep-Sea Research I*, *47*, 637–660, 2002.

Beck, C., W. Bowen, J. Mcmillan, and S. Iversion, Sex differences in the diving behaviour of a size-dimorphic capital breeder: the grey seal, *Animal Behaviour*, *66*, 777–789, 2003.

Behrenfeld, M., R. O'Malley, and Coauthors, Climate-driven trends in contemporary ocean productivity, *Nature*, *444*, 752–755, 2006.

Blanco, J. U., and J. Sheng, Numerical investigation of inter annual variability of circulation and hydrography over the eastern Canadian Shelf, *Atmosphere-Ocean*, *50*, 277–300, 2012.

Boehlert, G., D. Costa, D. Crocker, P. Green, T. O'Brien, S. Levitus, and B. L. Boeuf, Autonomous pinniped environmental samplers: Using instrumented animals as oceanographic data collectors, *J. Atmos. Oceanic Technol.*, *18*, 1882–1893, 2001.

Bowen, W., Role of marine mammals in aquatic ecosystems, *Mar. Ecol. Prog. Ser.*, *158*, 267–274, 1997.

Bowen, W., O. Oftedal, and D. Boness, Mass and energy transfer during lactation in a small phocid, the harbor seal (*Phoca vitulina*), *Physiol. Zool.*, *65*, 844–866, 1992.

Bowen, W., C. Beck, and S. Iverson, Bioelectrical impedance analysis as a means of estimating total body water in grey seals, *Canadian Journal of Zoology*, *77*, 1–5, 1999.

Boyce, D., M. Lewis, and B. Worm, Global phytoplankton decline over the past century, *Nature*, *446*, 591–596, 2010.

Brainerd, K., and M. Gregg, Surface mixed and mixing layer depths, *Deep-Sea Research I*, *42*, 1521–1543, 1995.

Breed, G., I. Jonsen, R. Myers, W. Bowen, and M. Leonard, Sex-specific, seasonal foraging tactics of adult grey seals (*Halichoerus grypus*) revealed by state-space analysis, *Ecology*, *90*, 3209–3221, 2009.

Brewin, R., E. Devred, S. Sathyendranath, S. Lavender, and N. Hardman-Mountfors, Model of phytoplankton absorption based on three size classes, *Applied Optics*, *50*, 4535–4549, 2011.

Ciotti, A., M. Lewis, and J. Cullen, Assessment of the relationships between dominant size in natural phytoplankton communities and the spectral shape of the absorption coefficient, *Limnol. Oceanogr.*, *47*, 404–417, 2002.

Craig, S., S. Lohrenz, Z. Lee, K. Mahoney, G. Kirkpatrick, and O. Schofield, Use of hyper spectral remote sensing reflectance for detection and assessment of the harmful alga, *Karenia brevis*, *Applied Optics*, *45*, 5414–5425, 2006.

Crowley, T., Causes of climate change over the past 1000 years, *Science*, *289*, 270–277, 2000.

Darecki, M., and D. Stramski, An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea, *Remote Sensing of Environment*, *89*, 326–350, 2004.

Dauchez, S., L. Legendre, L. Fortier, and M. Levasseur, Nitrate uptake by size-fractioned phytoplankton on the Scotian Shelf (Northwest Atlantic): spatial and temporal variability, *Journal of Plankton Research*, *18*, 577–595, 1996.

Daufresne, M., K. Lengfellnera, and U. Sommer, Global warming benefit the small in aquatic ecosystem, *Proc. Natl. Acad. Sci*, 2009.

Davies-Colley, R., W. Vant, and D. Smith, *Colour and clarity of natural waters*, Ellis Horwood, New York, 1993.

DFO, 2004 State of the ocean: Chemical and biological oceanographic conditions in the Gulf of Main-Bay of Fundy, Scotian Shelf and the Southern Gulf of St. Lawrence, *Can. Sci. Advis. Sec. Sci. Advis. Rep.*, *046*, 2005.

Doney, S., V. Fabry, R. Feely, and J. Kleypas, Ocean acidification: The other CO_2 problem, *Annu. Rev. Mar. Sci.*, *1*, 169–192, 2009.

E.D'Sa, and R. Miller, Bio-optical properties in waters influenced by Mississippi River during low flow conditions, *Rem. Sens. Env.*, *84*, 538–549, 2003.

Edwards, M., and A. Richardson, Impact of climate change on marine pelagic phenology and trophic mismatch, *Nature*, *430*, 881–884, 2004.

Eknes, M., and G. Evensen, An ensemble Kalman Filter with a 1-D marine ecosystem model, *J. Mar. Syst.*, *36*, 75–100, 2002.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*, 1143–1162, 1994.

Falkowski, P., R. Barber, and V. Smetacek, Biogeochemical controls and feedbacks on ocean primary production, *Science*, *281*, 200–206, 1998.

Fennel, K., J. Wilkin, J. Levin, J. Moisan, and J. O'Reilly, Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget, *Global Biogeochemical Cycles*, *20*, 2006.

Fennel, K., J. Wilkin, M. Previdi, and R. Najjar, Denitrification effects on air-sea CO_2 flux in the coastal ocean: Simulations for the Northwest North Atlantic, *Geophys. Res. Let.*, *35*, 2008.

Field, C., M. Behrenfeld, J. Randerson, and P. Falkowski, Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, *281*, 237–240, 1998.

Gallegos, C., Calculating optical water quality targets to restore and protect submersed aquatic vegetation: Overcoming problems in partitioning the diffuse attenuation coefficient for photosynthetically active radiation, *Eestuaries and Coasts*, *24*, 381–397, 2001.

Gitelson, A., D. Gurlin, W. Moses, and T. Barrow, A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters, *Environ. Res. Let.*, *4*, 2009.

Gnanadesikan, A., J. Dunne, and J. John, What ocean biogeochemical models can tell us about bottom-up control of ecosystem variability, *ICES Journal of Marine Science*, *68*, 1030–1044, 2011.

Green, R., H. Sosik, and R. Olson, Contributions of phytoplankton and other particles to inherent optical properties in New England continental shelf waters, *Limnol. Oceanogr.*, *48*, 2377–2391, 2003.

Greenan, B., B. Petrie, W. Harrison, and N. Oakey, Are the spring and fall blooms on the Scotian Shelf related to short-term physical events?, *Continental Shelf Research*, *24*, 603–625, 2004.

Gregg, W., Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model, *J. Mar. Sys.*, *69*, 205–225, 2007.

Grewal, M., and A. Andrews, *Kalman Filtering: Theory and Practice*, Prentice Hall, New Jersey, 1993.

Guinotte, J., and V. Fabry, Ocean acidification and its potential effects on marine ecosystem, *Annals of the New York Academy of Sciences*, *1134*, 320–342, 2010.

Gurlin, D., A. Gitelson, and W. Moses, Remote estimation of chlorophyll-a concentration in turbid productive waters-Return to a simple two-band NIR-red model?, *Rem. Sen. Env.*, *115*, 3479–3490, 2011.

Haidvogel, D., H. Arango, and Coauthors, Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the Regional Ocean Modeling System, *J. Comput. Phys.*, *227*, 3595–3524, 2008.

Hallegraeff, G., Ocean climate change, phytoplankton community responses, and harmful algal bloom: A formidable predictive challenge, *Journal of Phycology*, *46*, 220–235, 2010.

Harvey, V., M. Hammill, D. Swain, G. Breed, C. Lydersen, and K. Kovacs, Winter foraging by a top predator, the grey seal, in relation to the distribution of prey, *Mar. Eco. Prog. Ser.*, *462*, 273–286, 2012.

Hu, J., K. Fennel, J. Mattern, and J. Wilkin, Data Assimilation with a local ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic Bight, *J. Mar. Syst.*, *94*, 145–156, 2012.

Hunter, P., A. Tyler, M. Presing, A. Kovacs, and T. Preston, Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution, *Remote Sensing of Environment*, *112*, 1527–1544, 2008.

Irwin, A., and M. Oliver, Are ocean deserts getting larger?, *Geophys. Res. Let.*, *36, L18609, doi:10.1029/2009GL039883*, 2009.

Jerlov, N., *Marine Optics*, vol. 14, Elsevier, Amsterdam, 1976.

Justic, D., N. Rabalais, and E. Turner, Modeling the impacts of decadal changes in riverine nutrient fluxes on coastal eutrophication near the Mississipi River Delta, *Ecological Modelling*, *152*, 33–46, 2002.

K.Carder, R. Steward, G. Harvey, and P. Ortner, Marine humid and fulvic acids: Their effects on remote sensing of ocean chlorophyll, *Limnol. Oceanogr.*, *34*, 68–81, 1989.

Kirk, J., *Light and photosynthesis in aquatic ecosystems*, Cambridge University Press, 1994.

Kirkpatrick, G., D. Millie, M. Moline, and O. Schofield, Optical discrimination of a phytoplankton species in natural mixed populations, *Limnol. Oceanogr.*, *45*, 467–471, 2000.

Large, W., and S. Yeager, Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies, *NCAR Technical Note*, *460*, 2004.

Laurent, A., K. Fennel, J. Hu, and R. Hetland, Simulating the effects of phosphorus limitation in the Mississippi and Atchafalaya River plumes, *Biogeosciences Discuss*, *9*, 5625–5657, 2012.

Lehmann, M., K. Fennel, and R. He, Statistical validation of a 3-D biophysical model of western North Atlantic, *Biogeosciences*, *6*, 1961–1974, 2009.

Lewis, M., Biological Oceanography, University Lecture, 2011.

Lewis, M., E. Horne, J. Cullen, N. Oakey, and T. Platt, Turbulent motions may control phytoplankton photosynthesis in the upper ocean, *Nature*, *311*, 49–50, 1984.

Lorenc, A., Analysis methods for numerical weather prediction, *Q.J.R.Meteorol. Soc.*, *112*, 1177–1194, 1986.

Mattern, P., K. Fennel, and M. Dowd, Estimating time-dependent parameters for a biological ocean model using an emulator approach, *Journal of Marine Systems*, *96*, 32–47, 2012.

Meinke, I., B. Geyer, F. Feser, and H. von Storch, The impact of spectral nudging on cloud simulation with a regional atmospheric model, *American Meteorological Society*, *23*, 815–824, 2006.

Mellard, J., K. Yoshiyama, E. Litchman, and C. Kalusmeier, The vertical distribution of phytoplankton in stratified water columns, *Journal of Theoretical Biology*, *269*, 16–30, 2010.

Mobley, C., *Ocean optics web book*, http://www.oceanopticsbook.info/view/introduction/overview, 2010.

Moisan, T., S. Sathyendranath, and H. Bouman, *Remote sensing of biomass-Principles and Applications*, Dr. Lola Fatoyinbo, 2012.

Moore, T., R. Matear, J. Marra, and L. Clementson, Phytoplankton variability off the Western Australian Coast: Mesoscale eddies and their role in cross-shelf exchange, *Deep Sea Research Part II: Topical studies in Oceanography*, *54*, 943–960, 2007.

Moran, X., A. Lopez-Urrutia, A. Calvo-Diaz, and W. Li, Increasing importance of small phytoplankton in a warmer ocean, *Glob. Change Biol.*, *16*, 1137–1144, 2010.

Morel, A., Optical modeling of the upper ocean in relation to its biogenous matter content (case 1 water), *J. Geophys. Res.*, *93*, 749–768, 1988.

Morel, A., and S. Maritorena, Bio-optical properties of oceanic waters: A reappraisal, *J. Geophys. Res.*, *106*, 7163–7180, 2001.

Morel, A., and L. Prieur, Analysis of variations in ocean colour, *Limnol. Oceanogr.*, *22*, 709–722, 1977.

Moses, M., A. Gitelson, S. Berdnikov, and V. Povazhnyy, Estimation of chlorophyll-a concentration in case II waters using MOIS and MERIS data-successes and challenges, *Environ. Res. Let.*, *4*, 2009.

Mousseau, L., L. Legendre, and L. Fortier, Dynamics of size-fractionated phytoplankton and trophic pathways on the Scotian Shelf and at the shelf break, Northwest Atlantic, *Aquatic Microbial Ecology*, *10*, 149–163, 1996.

Natvik, L., and G. Evensen, Assimilation of ocean colour data into a biochemical model of the North Atlantic: part 1. Data assimilation experiments, *J. Mar. Syst.*, *40*, 127–153, 2003.

Nerger, L., and W. Gregg, *Estimation of model bias by the assimilation of satellite ocean chlorophyll data into a global model*, Presented as the Ocean Sciences Meeting, Orlando, FL, USA, March, 2008.

O'Boyle, S., and J. Silke, A review of phytoplankton ecology in estuarine and coastal waters around Ireland, *J. Plankton. Res.*, *32*, 99–118, 2010.

Perry, R., and S. Smith, Identifying habitat associations of marine shes using survey data: an application to the northwest Atlantic, *Canadian Journal of Fisheries and Aquatic Sciences*, *51*, 589–602, 1994.

Pham, D., J. Verron, and M. Roubaud, A singular evolutive extended Kalman Filter for data assimilation in oceanography, *J. Mar. Syst.*, *16*, 323–340, 1998.

Rabalais, N., E. Turner, R. Diaz, and D. Justic, Global change and eutrophication of coastal waters, *ICES Journal of Marine Science*, *66*, 1528–1537, 2009.

Radu, R., M. Deque, and S. Somot, Spectral nudging in a spectral regional climate model, *Tellus Series A*, *60*, 898–910, 2008.

Ramin, J. Z. M., V. Cheng, and G. Arhonditsis, Competition patterns among phytoplankton functional groups: How useful are the complex mathematical models?, *Acta Oecologica*, *33*, 334–344, 2008.

Retamal, L., W. Vincent, C. Martineau, and C. Osburn, Comparison of the optical properties of dissolved organic matter in two river-influenced coastal regions of the Canadian Arctic, *Estuarine, Coastal and Shelf Science*, *72*, 261–272, 2007.

Rost, B., I. Zondervan, and D. Wolf-Gladrow, Sensitivity of phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions, *Marine Ecology Progress Series*, *373*, 227–237, 2008.

Roy, S., HPLC-measured chlorophyll-type pigments during a phytoplankton spring bloom in Bedford Basin (Canada), *Mar. Ecol. Prog. Ser.*, *5*, 279–290, 1989.

Shi, D., Y. Xu, B. Hopkinson, and F. Morel, Effect of ocean acidification on iron availability to marine phytoplankton, *Science*, *327*, 676–679, 2010.

Song, H., R. Ji, C. Stock, K. Kearney, and Z. Wang, Interannual variability in phytoplankton blooms and plankton productivity over the Nova Scotian Shelf and in the Gulf of Maine, *Mar. Ecol. Prog. Ser.*, *426*, 105–118, 2011.

Stacey, M., J. Shore, D. Wright, and K. Thompson, Modeling events of sea-surface variability using spectral nudging eddy permitting model of the Northeast Pacific Ocean, *J. Geophys. Res.*, *111*, 2006.

Swain, D., G. Chouinard, R. Morin, and K. Drinkwater, Seasonal variation in the habitat associations of Atlantic Cod (*Gadus morhua*) and American plaice (*Hippoglossoides platessoides*) from the southern Gulf of St. Lawrence, *Canadian Journal of Fisheries and Aquatic Sciences*, *55*, 2548–2561, 1998.

Teo, S., R. Kudela, A. Rais, C. Perle, D. Costa, and B. Block, Estimating chlorophyll profiles from electronic tags deployed on pelagic animals, *Aquat. Biol.*, *5*, 195–207, 2009.

Thomas, M., C. Kremer, C. Kalusmeier, and E. Litchman, A global pattern of thermal adaptation in marine phytoplankton, *Science*, *338*, 1085–1087, 2012.

Thompson, D., and M. Fedak, How long should a dive last? A simple model of foraging decisions by breath-hold divers in a patchy environment, *Animal Behaviour*, *61*, 287–296, 2001.

Thompson, K., D. Wright, Y. Lu, and E. Demirov, A simple method for reducing seasonal bias and drift in eddy resolving ocean models, *Ocean Modelling*, *13*, 190–125, 2006.

Thompson, K., D. Wright, Y. Lu, and E. Demirov, Suppressing drift and bias of coastal circulation models through the assimilation of seasonal climatologies of temperature and salinity, *Cont. Shelf Res.*, *27*, 2007.

Topliss, B., L. Payzant, P. Hurley, J. Miller, and J. Freemantle, Interpretation of multi-season, multi-year colour imagery for a continental shelf region, *Limnol. Oceanogr.*, *14*, 1991.

Turner, E., and N. Rabalais, Coastal eutrophication near the Mississipi River Delta, *Nature*, *368*, 619–621, 1994.

While, J., K. Haines, and G. Smith, A nutrient increment method for reducing bias in global biogeochemical models, *J. Geophys. Res.*, *115*, 2010.

Wikle, C., and L. Berliner, A Bayesian tutorial for data assimilation, *Physica D*, *42*, 1521–1543, 2006.

Wildlife Computer USA, Mk10-AL Transmitting Linked Sensor Tag, `http://www.wildlifecomputers.com/products.aspx?id=13`, 2012, accessed: 06/01/2013.

Wozniak, B., and J. Dera, *Light absorption in sea water*, Springer, 2007.

Wuest, A., G. Piepke, and D. V. Senden, Turbulent kinetic energy balance as a tool for estimating vertical diffusivity in wind-forced stratified waters, *Limnology and Oceanography*, *45*, 1388–1400, 2000.

Yentsch, C., and D. Phinney, *Mapping Strategies in Chemical Oceanography*, ACS Publications, Washington, 1985.

Y.Kang, H. Choi, J. Noh, J. Choi, and I. Jeon, Seasonal variation of phytoplankton community structure in Northeastern Coastal Waters off the Korean Peninsula, *Algae*, *21*, 83–90, 2006.

York, D., N. Evensen, M. Martinez, and J. Delgado, Unified equations for the slope, intercept, and standard errors of the best straight line, *Am. J. Phys.*, *72*, 2004.