Graduate thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

Hanne Opseth Rygg

State and parameter estimation for growth of the farmed kelp *Saccharina latissima*

Graduate thesis in Cybernetics and Robotics
Supervisor: Morten Omholt Alver
Co-supervisor: Ole Jacob Broch
June 2023

NTNU
Norwegian University of Science and Technology

SINTEF

Hanne Opseth Rygg

State and parameter estimation for growth of the farmed kelp *Saccharina latissima*

Graduate thesis in Cybernetics and Robotics
Supervisor: Morten Omholt Alver
Co-supervisor: Ole Jacob Broch
June 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of
Science and Technology

Preface

This report was written in TTK4900, Master's Project in Technical Cybernetics, a 30-credit program, in collaboration with Sintef Ocean. The course aims for students to specialize in a chosen area using scientific methodologies, gather additional information through literature reviews and other sources, and integrate this knowledge into a comprehensive report. This project builds upon the work completed in the fall of 2022 during the TTK4550 - Specialization Project. The master's thesis was conducted in the spring semester of 2023.

This thesis includes a selection of material from previously studied work, which has been restated to enhance readability and provide the necessary context for the reader. These restatements will be explicitly acknowledged in the text.

I want to thank Morten Omholt Alver for constructive feedback and technical guidance, and SINTEF Ocean with Ole Jacob Broch for the opportunity to work on this project and for being available for questions regarding kelp and the growth model.

The code used to carry out this project can be found here: `https://github.com/hanneory/State-and-parameter-estimation-kelp-growth.git`.

Trondheim, 01-06-2023

Hanne Opseth Rygg

Executive summary

This research aims to estimate the growth states of the cultivated kelp species *Saccharina latissima*.

The escalating trend of kelp farming in Norway calls for advanced knowledge and potential solutions in the areas of modeling and automation, intending to enhance production scalability and efficiency in the seaweed aquaculture industry. An established growth model for *Saccharina latissima* offers application potential in resolving aquaculture-related issues, including assessing nutrient absorption capabilities and evaluating kelp's potential as a bioenergy production resource. Enhancing the accuracy of state estimation for growth is a subject of interest, as it is crucial for making the growth model more robust and refining the precision of its predictions.

The growth states in this model are the area, nitrogen, and carbon content of the kelp. The outcomes of this thesis illustrate that an ensemble Kalman filter can beneficially be used to estimate these states. Moreover, the ensemble Kalman filter demonstrated proficiency in estimating some model parameters and enhancing the precision of the kelp growth predictions.

The twin experiment investigated the model's theoretical aspects and filter implementation. For precise growth estimates, all three states should be measured every other week. Once this is achieved, obtaining an even better estimate is possible by including parameter estimation for the key model parameters related to uptake, photosynthesis, and biomass development.

From a practical perspective, when only measurements of the frond area are available, it would be beneficial to include parameter estimation for the parameters related to biomass development. The precision of the prediction weakens towards the end of the growth period, suggesting the need for more frequent measurements in this part.

To summarize, the findings of this project indicate that state estimation via the ensemble Kalman filter effectively enhances the accuracy of growth estimates. This study can also serve as a starting point for real-time estimates of kelp growth for *Saccharina latissima*.

Samandrag

Denne forskinga har som mål om å estimere veksttilstandane til den dyrka tarearten *Saccharina latissima*.

Den auka trenden for taredyrking i Noreg stiller krav til avansert kunnskap og potensielle løysingar innan modellering og automatisering. Dette med intensjon om å forbetre produksjonskapasiteten og effektiviteten i oppdrettsnæringa av tang. Ein etablert vekstmodell for *Saccharina latissima* byr på moglegheiter for å løysa problemstillingar knyta til akvakultur, inkludert vurdering av næringsopptaksevner og evaluering av taren sin levedyktigheit som ein ressurs for bioenergiproduksjon. Det er ønskeleg å forbetre nøyaktigheita i estimata av tilstandane for vekst for å gjere modellen meir robust, og samt auke presisjonen i predikasjonane.

Tilstandane for vekst i denne modellen er areal, nitrogen og karbon. Resultata frå dette prosjektet viser at eit ensemble Kalman-filter kan fordelaktig bli brukt til å estimere desse tilstandane. I tillegg klarte filteret å estimere nokre av parameterane for modellen. Dette resulterte også til betre predikasjonar av tareveksten.

Tvillingeksperimentet vart nytta for å undersøke dei teoretiske aspekta ved modellen og implementasjonen av filteret. For nøyaktig predikasjon av vekt bør alle tre tilstandane målast annankvar veke. Når dette er oppnådd, er det mogleg å få eit endå betre estimat ved å inkludere parameterestimering for nøkkelparameterane relatert til opptak, fotosyntese og biomasseutvikling.

Frå eit praktisk synspunkt, når berre målingar av bladområdet er tilgjengelege, vil det vere fordelaktig å inkludere parameterestimering for parameterane relatert til biomassen. Presisjonen av predikasjonane svekkast mot slutten av vekstperioden, noko som tydar på at det er eit behov for hyppigare målingar i denne delen.

For å summere opp, indikerer funna i dette prosjektet at tilstandsestimering via ensemble Kalman-filter vil effektivt forbetra nøyaktigheita av vekstpredikasjonane. Dette studiet kan også fungere som eit utgangspunkt for sanntids tilstandsestimering av tarevekst for *Saccharina latissima*.

Table of Contents

List of Tables

List of Figures

1

Introduction

This section introduces the problem that has been investigated in this project. Parts of section 1.1 are restated material from the project thesis, with some minor modifications incorporated [37].

1.1 Background

Cultivation of kelp has become a new investment area within aquaculture in Norway, with the first permit issued in 2014 [16]. The UN Food and Agriculture Organization reports that global seaweed production reached 34.7 million tons in 2019, predominantly produced in Asian countries like China and Indonesia [12]. Compared to Asia, the production volume and the number of species in seaweed cultivation are far behind here in Norway [20]. Nevertheless, interest in seaweed cultivation is rapidly expanding. There are 539 permits for macroalgal cultivation distributed across 105 locations and 23 companies in Norway [16]. Predictions state that kelp farming in Norway alone will reach 20 million tons by 2050, marking a substantial increase from the 180 tons produced in 2021 [16][34].

There are several reasons for increasing kelp production. Kelp serves as a potential solution to climate change and ocean acidification, given its capability to absorb CO_2 and its role in mitigating coastal eutrophication by removing extra nutrients from water caused by agricultural, industrial, or fish farming activities. Cultivated kelp can also contribute to enhancing marine habitats and promoting biodiversity. Moreover, kelp finds several commercial applications, ranging from bioenergy to food, cosmetics, material for packaging, and pharmaceuticals [4][38].

In the past decade, expertise in cultivating and harvesting seaweed for various applications has significantly advanced [11][43]. The overarching objective has

been to establish a Norwegian bio-economy centered around cultivated seaweed [42]. Comprehensive understanding of the growth potential and quality of *S. latissima* along Norway's extensive coastline would aid farmers in making informed decisions regarding deployment and harvest locations and timings to maximize production while minimizing losses. Several initiatives have been implemented to achieve this goal, such as establishing the Norwegian Test Center for Seaweed Cultivation and Utilization Technologies, Norwegian Seaweed Technology Center (earlier RI seaweed), in 2021. This research infrastructure, operated by SINTEF AS and NTNU, aims to support the Norwegian industry and public sector in creating a new economy based on seaweed cultivation and processing. As part of this effort, test facilities will be built, three at sea and two on land. Norwegian Seaweed Technology Center aims to advance technology related to industrial seaweed cultivation, harvesting, processing, and applications in Norway [2][39][18].

Furthermore, in July 2022, SINTEF, DNV, Equinor, and Aker BP collaborated to initiate the world's first pilot project focused on active, nature-based carbon capture in marine environments. The objective is to develop technology and methodologies to capture several million tonnes of CO_2 utilizing kelp [14].

Additionally, a notable effort in Norway is displayed in the Green Platform (GP) Seaweed project, led by SINTEF Ocean. The Research Council of Norway granted this initiative 50 million NOK as part of the Green Platform Initiative. It aims to enhance technology and knowledge concerning the expansion and modernization of kelp farming and to foster the development of new, environmentally friendly products derived from farmed kelp for use in blue-green value chains. The objectives further include expanding kelp cultivation's growth and positioning Norway as a global leader in kelp and kelp cultivation. The main emphasis is on products within four market segments: food, animal feed, substitutes for biodegradable plastic, and fertilizers/biocarbon [41].

Numerous challenges must be tackled to establish a novel marine industry around cultivated kelp. To this date, Norwegian companies have yet to possess extensive expertise in kelp farming, making it crucial to bridge the knowledge gap in the initial stages. As the industry expands, there will be a growing demand for enhanced knowledge and automation in areas such as seedling production, deployment, monitoring, harvesting, and storing large quantities of biomass. Optimizing the location of kelp farming to maximize biomass production will also be crucial. This can be accomplished by utilizing kelp growth models and actual environmental measurements or by employing ocean models like SINMOD [27]. If the ideal location for kelp farming lies offshore, it will pose additional challenges and require more advanced technology [35]. To address the increasing demand for kelp and the associated knowledge requirements, developing effective solutions for modeling and automating kelp farming is vital for scaling up and optimizing

the industry.

1.2 Problem description

SINTEF Ocean has already developed a dynamic system that is being used to simulate growth and biomass in kelp aquaculture. It is known that some of the most important environmental variables for kelp growth are light, temperature, salinity, and nutrient concentrations [9]. This model can be run using measured and simulated environmental conditions as input data. During the project thesis conducted in autumn 2022, the model's sensitivity was examined regarding the initial values, parameters, and input values [37]. This thesis, however, aims to utilize an ensemble Kalman filter for model correction and parameter estimation through measurements. The choice of which parameters should be estimated will be based on the sensitivity analysis conducted in the project thesis. The objectives for achieving this task are listed below.

- Implementation of the ensemble Kalman filter and selection of parameters for estimation.
- Test the implementation using a twin experiment where synthetic observations are generated from running the kelp model.
- Test the estimator with data from real kelp experiments.
- Assessment of the results of the state and parameter estimation.

1.3 Delimitations

In this thesis, the focus is solely on improving the state and parameter estimation of the model without making any other adjustments.

The conclusions made apply solely to the specific growth model of *Saccharina latissima* and are not relevant for any other varieties of kelp.

It is important to note that the dataset utilized for this analysis includes only data from March to June, thus limiting the predictions to this specific time range.

The data represent Trøndelag exclusively, meaning that the conclusions drawn from this study only apply to this region and no other areas in Norway. While some similarities may exist, and a certain degree of knowledge can be transferred, this transferability is expected to be limited.

The ensemble Kalman filter was the only technique employed for estimating the states and parameters. Alternative methods were not explored, which may have yielded more accurate estimates.

Another delimitation of the project is that the perturbations introduced to the model are intended to account for the uncertainty in the model rather than addressing the biological diversity within kelp.

1.4 Structure of the report

The organization of this report is as follows. Section 2 provides the study's theoretical foundation, containing topics such as farmed kelp, *Saccharina latissima*, data assimilation, Monte Carlo simulation, ensemble Kalman filter, and the growth model.

Section 3 outlines the methodologies employed in this project, containing details about the software utilized, the dataset, the growth model implementation, and the configuration of the ensemble Kalman filter.

Section 4 presents the project outcomes, which are further analyzed in Section 5 along with recommendations for future research. Section 6 summarises the findings to conclude the report.

2

Theory

This section introduces the foundational theoretical concepts upon which this project is built. Section 2.1, 2.2.2, and 2.3 are restated material from the project thesis, with some modifications incorporated [37].

2.1 Kelp

Kelp is a brown algae made of stipes and lamina and can be found all along the coast of Norway, including Svalbard. The most common species include *Laminaria cloustonii*, *Laminaria digitata*, and *Saccharina latissima*. These submarine forests live in cool waters, close to shore, and from 1 meter to 30 meters depth [1]. The combination of high kelp productivity in environments with the potential for high disturbance rates can make these forests relatively short-lived. Entire kelp beds can be eliminated by thermal events, storms, or outbreaks of herbivores and disappear within a year, but can also return nearly as quickly [47].

2.1.1 *Saccharina latissima*

Saccharina latissima is among Norway's most common kelp species. It is distinct from its relatives in that it has a long lamina with folds along the edge and a relatively wide and uneven middle section. The lamina of *S. Latissima* can grow up to 4 meters in length and 30 cm in width [21]. Approximately 50% of the world's natural sugar kelp beds are located along the coast of Norway [32].

The sporophyte of *S. latissima* consists of a lamina, stipe, and haptera. The lamina grows from an intercalary meristem located between the lamina and the stipes [19]. The zoospores are released from the sporangia of the kelp and develop into male and female gametophytes. Male gametophytes are thin and heavily

branched, while females have thicker and less branched cells. Here, the kelp enters the reproductive phase induced by optimal conditions like light and temperature. The female gametophytes produce eggs, while the male gametophytes produce spermatozoids with two flagella. The egg release pheromone, which attracts the spermatozoids. It then swims over and fertilizes the egg [33]. The fertilized egg then develops into a sporophyte. The life cycle of *S. latissima* is visualized in Figure 2.1.

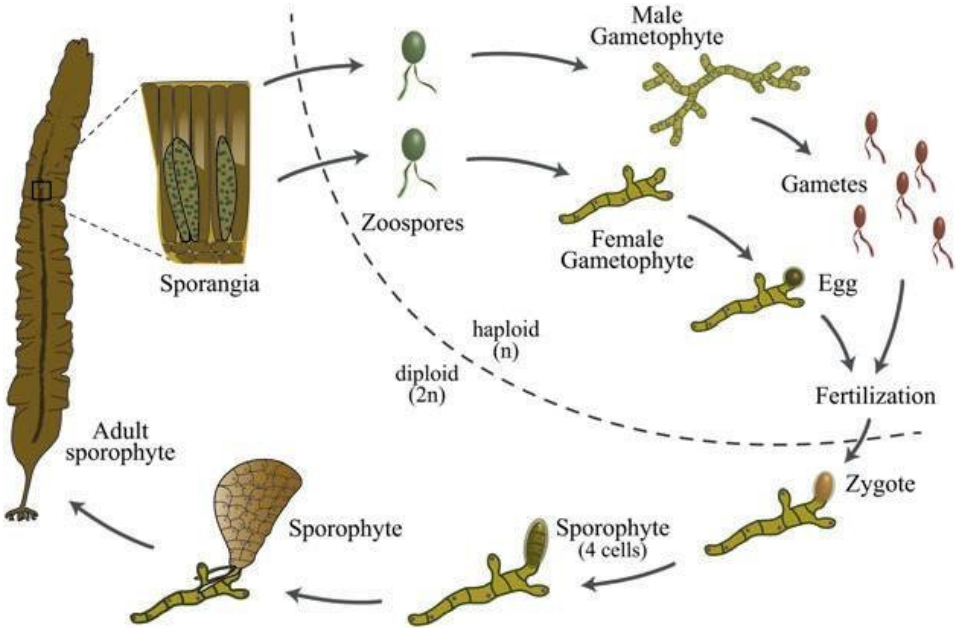


Figure 2.1: Life cycle of *Saccharina latissima* [45].

2.1.2 Farmed kelp

In farmed kelp, the first half of the life cycle mentioned in section 2.1.1 is executed in a controlled environment in the laboratory. This enhances the prospects for optimal growth and survival [17]. Spores are gathered from a mature sporophyte and then seeded onto ropes. In the study by Broch and Alver, the seedlings measured 1-5 mm before the ropes were deployed into the sea [6]. Typically, kelp is positioned between 2 and 15 meters deep in the ocean and harvested once it achieves the desired length. *S. Latissima* experiences its fastest growth rate from late winter to spring, at approximately 1.1 cm/day. As a result, kelp is frequently deployed into the sea in late autumn [46]. However, according to Forbord, cultivation in Norway is optimal during autumn, winter, and spring. *S. Latissima* is commonly deployed in September and February and harvested before summer, usually in May

or June [19]. A visualization of the production cycle of *Laminaria japonica* can be seen in Figure 2.2. This figure is to some extent alienable with *S. Latissima*. The phases of the production cycle of *S. Latissima* can however be seen in Figure 2.3

It is customary to take measurements to know when to harvest the kelp. The most common methods for assessing seaweed growth, according to Forbord, are frond elongation or biomass increase between two registration points. Frond elongation involves measuring length by tracking the movement of a point in the meristem over time or measuring the entire lamina's length from above the stipe to the distal end. Measuring frond elongation does not necessitate biomass destruction and can be performed on seaweed still attached to the cultivation lines. In contrast, measuring the total biomass yield requires weighing the cultivation, which is challenging while the biomass remains attached to the cultivation lines. Although assessing the dry weight of the kelp is more accurate, it necessitates harvesting and destroying the biomass, which is not always feasible [19]. The current method for measuring the carbon and nitrogen content in the kelp involves harvesting and destroying the biomass [44].

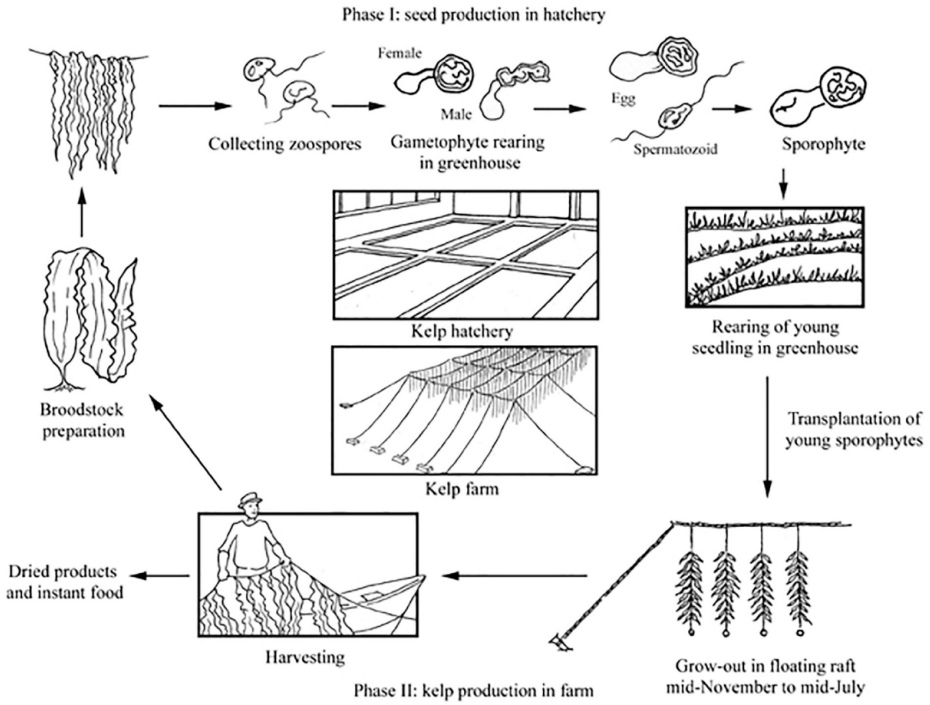


Figure 2.2: Production cycle of *Laminaria japonica* [13].

Figure 2.3: Phases of kelp production. Adapted from Broch [7].

2.2 Data assimilation

Data assimilation (DA) is a technique used to combine observation data and the underlying dynamical principles governing the system to provide an estimate of the system's state which is better than what could be obtained using just the data or the model alone. Several different data assimilation methods exist, such as the Cressman analysis method, the optimal interpolation method, the three-/four-dimensional variational analysis, and the Kalman filter (KF). In all these methods, the final estimate is chosen to minimize the uncertainty of the final estimate [49].

2.2.1 State and parameter estimation

Evensen states that the parameter estimation problem for a dynamical model can generally be formulated as determining the joint probability density function of the parameters and model state, given a set of measurements and a dynamical model with known uncertainties [15].

State estimation refers to employing a system's input and output measurements to conclude the current values of its internal variables, also called the system's state. It uses available data to infer the system's status at a specific point in time. State estimation can be a practical and cost-effective substitute for real measurements. The more information the controller has about the process it manages, the more precise its control can be. The Kalman Filter is a crucial algorithm for state estimation [25].

Parameter estimation involves determining the values of unknown parameters in a mathematical model based on observed data. It is frequently used to enhance a model's accuracy or better understand a system's behavior. This is particularly important when the true values of the parameters are unknown, and the model is used to make predictions or optimize an aspect of the system's performance [36].

2.2.2 Monte Carlo simulation

Monte Carlo simulation refers to system simulation using random sequences as inputs. As explained by Brown and Hwang, such methods are often helpful in un-

derstanding the behavior of stochastic systems that are not amenable to analysis by usual direct mathematical methods. These methods involve setting up a statistical experiment that matches the physical problem of interest, repeating the experiment repeatedly with typical sequences of random numbers, and finally, analyzing the experiment's results statistically [10]. It can also be used to estimate uncertainty in the problems using statistical sampling [15]. As an example, the input can be generated by N inputs from a Gaussian distribution (Equation 2.1) and further asses the outputs from the simulation by, for example, the expected value (Equation 2.2).

$$X \sim \mathcal{N}(\mu, \sigma^2) \tag{2.1}$$

$$E(X) = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{2.2}$$

2.2.3 Ensemble Kalman filter

The Kalman filter is an algorithm designed to estimate the state of a linear system based on measurements observed over time. This project exclusively utilizes an ensemble Kalman filter (EnKF) for state and parameter estimation due to the characteristics of the growth model. Consequently, the EnKF is the only filter that will be presented.

The EnKF is employed for large-scale nonlinear systems. The basic and extended Kalman filter pose challenges for these systems, often stemming from extensive computation time and complexity in managing nonlinear dynamics. The main limitation usually involves the computation of the error covariance matrix. Within the EnKF framework, the estimation of this covariance matrix (denoted as $\boldsymbol{P}$) is facilitated through both prediction and analysis based on ensemble statistics.

The EnKF is a method developed by Geir Evensen and is often used for weather forecasting, oceanography, and oil reservoir management. The theoretical formulation of the EnKF will be described in this section. A more detailed explanation can be found in Evensen's book [15]. This explanation, however, including equations, is based on the description of the filter provided by Houtkamer and Zhang [23].

The EnKF is a sequential data assimilation scheme that integrates observation data into a dynamical system obeying Kalman filter theory and the Monte Carlo method [15]. The EnKF can be regarded as an approximation of the Kalman filter, wherein the state distribution is characterized by an ensemble drawn from the distribution. This ensemble is advanced through time and updated as new data emerges. The ensemble representation is a dimensionality reduction technique, enabling computational feasibility for high-dimensional systems. EnKF maintains

and processes *N* vectors (ensemble members) with a length of *n*, and the estimated Kalman gain can be computed efficiently. In theory, as $N \to \infty$, the EnKF converges to the (exact) Kalman filter for linear Gaussian models, but large values for N are typically infeasible in real-world applications [26].

Within the EnKF framework, there are two primary approaches for parameter estimation. The first approach is state augmentation, which involves treating the parameters as time-varying quantities with slight artificial evolution noise. The states and parameters are combined in an augmented state vector, and an EnKF is run on this augmented state vector to obtain posterior estimates of states and parameters at each time, t [26].

The second approach for parameter estimation within the EnKF framework relies on constructing approximated likelihood functions using the output from the EnKF. Parameters are estimated through maximum likelihood or Bayesian methods. This method typically performs well for instances with relatively few parameters. Still, additional work is needed for scenarios where both parameters and states are high-dimensional [26].

The EnKF updates a previous estimate of the state, called forecast, denoted as $\boldsymbol{x}^f(t)$, at a certain time t, by incorporating new observational data, $\boldsymbol{y}^o$. This assimilation generates an improved estimate of the state vector, $\boldsymbol{x}^a(t)$, as depicted in Equation 2.3. The Kalman gain matrix, $\boldsymbol{K}$, assigns appropriate weight to the observations, which come with an error covariance $\boldsymbol{R}$, and to the forecast values, associated with an error covariance $\boldsymbol{P}^f$, as shown in Equation 2.4. The forward operator, $\mathcal{H}$, maps from the model space to the observation space. Finally, a forecast model, $\mathcal{M}$, is needed to transport the new estimate $\boldsymbol{x}^a(t)$ to the following analysis time as in Equation 2.5 [23].

$$\boldsymbol{x}^a(t) = \boldsymbol{x}^f(t) + \boldsymbol{K}[\boldsymbol{y}^o - \mathcal{H}\boldsymbol{x}^f(t)] \tag{2.3}$$

$$\boldsymbol{K} = \boldsymbol{P}^f\mathcal{H}^T(\mathcal{H}\boldsymbol{P}^f\mathcal{H}^T + \boldsymbol{R})^{-1} \tag{2.4}$$

$$\boldsymbol{x}^f(t+1) = \mathcal{M}[\boldsymbol{x}^a(t)] \tag{2.5}$$

In a Monte Carlo implementation, the ith element of an N-element analysis ensemble is derived by assessing Equation 2.3. This is done by applying a vector of observations, $\boldsymbol{y}_i^o$, that has been randomly altered and by utilizing a corresponding ensemble member from the forecast, as illustrated in Equation 2.6.

$$\boldsymbol{x}_i^a(t) = \boldsymbol{x}_i^f(t) + \boldsymbol{K}[\boldsymbol{y}_i^o - \mathcal{H}\boldsymbol{x}_i^f(t)], \quad i = 1, \ldots, N \tag{2.6}$$

Similarly, to obtain a member of the forecast ensemble valid at time $t+1$, Equation 2.5 can be used with a corresponding member of the analysis ensemble and a realization of the forecast model $\mathcal{M}$, seen in Equation 2.7.

$$\boldsymbol{x}_i^f(t+1) = \mathcal{M}_i[\boldsymbol{x}_i^a(t)], \quad i = 1, \ldots, N \tag{2.7}$$

The ensembles, which are produced by evaluating Equation 2.6 and 2.7, can be utilized to estimate the analysis error covariance matrix, denoted as $\boldsymbol{P}^a(t)$, and the forecast error covariance matrix, represented as $\boldsymbol{P}^f(t)$.

In an EnKF, there is no requirement for a comprehensive covariance matrix such as $\boldsymbol{P}^f$ in the model state space. Instead, when calculating the Kalman gain, $\boldsymbol{K}$, as displayed in Equation 2.4, the method relies on ensemble-based approximations of $\boldsymbol{P}^f\mathcal{H}^T$ and $\mathcal{H}\boldsymbol{P}^f\mathcal{H}^T$, as illustrated in Equation 2.8 and 2.9.

$$\boldsymbol{P}^f\mathcal{H}^T = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_i^f - \overline{\boldsymbol{x}^f})(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f})^T \tag{2.8}$$

$$\mathcal{H}\boldsymbol{P}^f\mathcal{H}^T = \frac{1}{N-1}\sum_{i=1}^{N}(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f})(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f})^T \tag{2.9}$$

Here $\overline{\boldsymbol{x}^f}$ and $\overline{\mathcal{H}\boldsymbol{x}^f}$ are given by Equation 2.10 and 2.11.

$$\overline{\boldsymbol{x}^f} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i^f \tag{2.10}$$

$$\overline{\mathcal{H}\boldsymbol{x}^f} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{H}\boldsymbol{x}_i^f \tag{2.11}$$

2.2.4 Twin experiment

In the context of EnKF, a twin experiment is a simulation-based study designed to evaluate the performance of the EnKF algorithm within a controlled setting, wherein the true system state is known. This involves running two parallel experiments that are identical, with the exception that one employs the actual system state while the other utilizes the estimated state derived from the EnKF [48].

In this study, the twin experiment is used to investigate the filter in a controlled environment, which is not possible in real-world scenarios because the true state of the system is not known. It makes it possible to investigate the parameter tuning of the assimilation algorithm to understand how different settings impact the results, as well as give a better understanding of how model and observation errors affect

the assimilation results. These insights can then be used to refine the filter and improve its performance in real-world applications.

2.3 The growth model

The mathematical model used in this project is a dynamical model for simulating the growth of the brown algae *Saccharina latissima*. It is developed by Broch and Slagstad [9]. All the following equations, states, and parameters are collected from their report unless specified otherwise, and a summary of these equations can be found in Table 2.1, as well as an overview of the model variables in Table 2.2. In this thesis, temperature, irradiance, nutrient concentration, and current are referred to as the environmental or input variables.

The estimations of the model parameters can be seen in Table 2.3. These values have been collected and estimated from the literature but may not be specifically tailored to the Norwegian environment. In some cases, there is limited data available on these values, which can contribute to uncertainty in the estimates. The same can be concluded about the initial values of the model. Some key parameters in the sensitivity analysis (SA) will be further presented in subsection 2.3.10.

A schematic representation of the model is shown in Figure 2.4. This illustration reveals that carbon and nitrogen reserves are the most important factors for kelp growth.

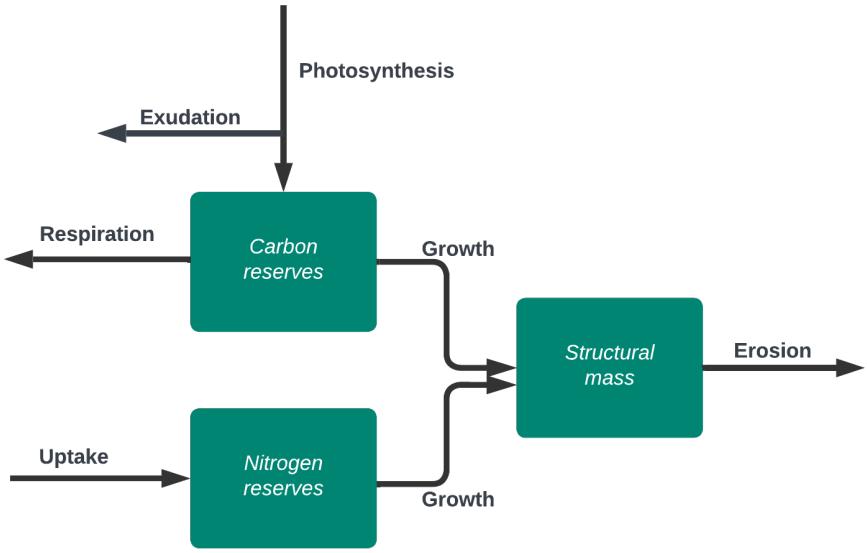


Figure 2.4: Schematic overview of the growth model. Adapted from Broch and Slagstad [9].

Table 2.1: Model variables and calculated quantities. Adapted from Broch and Slagstad [9].

Eq.		Description
2.12	$dA/dt = (\mu - \nu)A$	Rate of change of frond area
2.13	$\mu = f_{area} f_{photo} f_{temp} \min(1\text{-}N_{min}/N, 1 - C_{min}/C)$	Specific growth rate
2.14	$f_{area}(A) = m_1 \exp(-(A/A_0)^2) + m_2$	Effect of size on growth rate
2.15	$f_{temp}(T) = \begin{cases} 0.08T + 0.2 & \textit{for } -1.8 \leq T < 10 \\ 1 & \textit{for } 10 \leq T \leq 15 \\ 19/4 - T/4 & \textit{for } 15 < T \leq 19 \\ 0 & \textit{for } T > 19 \end{cases}$	Effect of temperature on growth rate
2.16	$f_{photo}(n) = a_1(1 + \text{sgn}(\lambda(n))\lvert\lambda(n)\rvert^{1/2}) + a_2$	Seasonal influence on growth rate
2.17	$\nu(A) = \frac{10^{-6}\exp(\varepsilon A)}{(1+10^{-6}(\exp(\varepsilon A)-1))}$	Frond erosion
2.18	$dN/dt = k_A^{-1} J - \mu(N + N_{struct})$	Rate of change in nitrogen reserves
2.19	$J = J_{max} \frac{X}{K_X + X} (\frac{N_{max} - N}{N_{max} - N_{min}})(1 - \exp(-U/U_{0.65}))$	Nitrate uptake rate
2.20	$dC/dt = k_A^{-1}(P(I,T)(1 - E(C)) - R(T)) - (C + C_{struct})\mu$	Rate of change in carbon reservese
2.21	$P(I,T) = P_S(1 - \exp(-\frac{\alpha I}{P_S}))\exp(-\frac{\beta I}{P_S})$	Gross photosynthesis
2.25	$R(T) = r_1 \exp(\frac{T_A}{T_1} - \frac{T_A}{T})$	Temperature dependent respiration
2.26	$E(C) = 1 - \exp(\gamma(C_{min} - C))$	Carbon exudation

Table 2.2: Model variables and calculated quantities. Adapted from Broch and Slagstad [9].

Symbol	Unit	Description
A	dm^2	Frond area, state variable
C	$gC(gsw)^{-1}$	Carbon reserves, relative to W_s, state variable
N	$gN(gsw)^{-1}$	Nitrogen reserve, relative to W_s, state variable
μ	day^{-1}	Specific growth rate (area), derived variable
W_w	g	Total wet weight of sporophyte, derived variable
W_d	g	Total dry weight, derived variable
W_s	g	Dry weight of the structural mass, derived variable
β	$gO_2dm^{-2}h^{-1}(\mu mol$ photons m^{-2} $s^{-1})^{-1}$	Photoinhibition parameter, auxiliary variable
P_s	$gO_2dm^{-2}h^{-1}$	Photosynthesis parameter, auxiliary variable
I	μmol photons $s^{-1}m^{-2}$	Irradiance (PAR), environmental variable
T	$^\circ$C	Water temperature, environmental variable
U	ms^{-1}	Water current speed, environmental variable
X	μmol N m^{-3}	Substrate nutrient concentration, environmental variable

Table 2.3: Model parameters. Adapted from Broch and Slagstad [9].

Symbol	Value	Unit	Description
A_0	6	dm^2	Growth rate adjustment parameter
α	3.75 $\times 10^{-5}$	gC $dm^{-2}h^{-1}$ (μmol photons m^{-2} $s^{-1})^{-1}$	Photosynthetic efficiency
C_{min}	0.01	$gC(gsw)^{-1}$	Minimal carbon reserve
C_{struct}	0.20	$gC(gsw)^{-1}$	Amount of carbon per unit dry weight of structural mass
γ	0.5	gCg^{-1}	Exudation parameter
ε	0.22	A^{-1}	Frond erosion parameter
I_{sat}	200	μmol photons m^{-2} s^{-1}	Irradiance for maximal photosynthesis
J_{max}	1.4$\times 10^{-4}$	$gNdm^{-2}h^{-1}$	Maximal nitrate uptake rate
k_A	0.6	gdm^{-2}	Structural dry weight per unit area
k_dw	0.0785		Dry weight to wet weight ratio of structural mass
k_C	2.1213	$g(gC)^{-1}$	Mass of carbon reserves per gram carbon
k_N	2.72	$g(gN)^{-1}$	Mass of nitrogen reserves per gram nitrogen
m_1	0.1085		Growth rate adjustment parameter
m_2	0.03		Growth rate adjustment parameter
μ_{max}	0.18	day^{-1}	Maximal area specific growth ratio
N_{min}	0.01	$gN(gw)^{-1}$	Minimal nitrogen reserve
N_{max}	0.022	$gN(gw)^{-1}$	Maximal nitrogen reserve
N_{struct}	0.01	$gN(gw)^{-1}$	Amount of nitrogen per unit dry weight of structural mass
P_1	1.22$\times 10^{-3}$	$gCdm^{-2}h^{-1}$	Maximal photosynthetic rate at $T = T^{\circ}_{P_1}$K
P_2	1.44$\times 10^{-3}$	$gCdm^{-2}h^{-1}$	Maximal photosynthetic rate at $T = T^{\circ}_{P_2}$K
a_1	0.85		Photoperiod parameter
a_2	0.3		Photoperiod parameter
R_1	2.785$\times 10^{-4}$	$gCdm^{-2}h^{-1}$	Respiration rate at $T = T_{R_1}$
R_2	5.429$\times 10^{-4}$	$gCdm^{-2}h^{-1}$	Respiration rate at $T = T_{R_2}$
T_{R_1}	285		Reference temperature for respiration
T_{R_2}	290		Reference temperature for respiration
T_{AP}	1694.4		Arrhenius temperature for photosynthesis
T_{APH}	25924		Arrhenius temperature for photosynthesis at high end of range
T_{APL}	27774		Arrhenius temperature for photosynthesis at low end of range
T_{AR}	11033		Arrhenius temperature for respiration
$U_{0.65}$	0.03	ms^{-1}	Current speed at which $J = 0.65J_{max}$
K_X	4	μmol N m^{-3}	Nitrate uptake half saturation constant

2.3.1 Frond area

Equation 2.12 represents the the rate of change in frond area A with respect to time. Function μ is a gross area-specific instantaneous growth rate, and function ν is the rate of frond loss.

$$\frac{dA}{dt} = (\mu(A, N, C, T, t) - \nu(A))A \tag{2.12}$$

2.3.2 Specific growth rate

The gross growth rate μ depends on size, carbon, nitrogen, temperature, and photoperiod and can be seen in Equation 2.13. Here f_{area} takes the effect of size into account, f_{temp} takes temperature into account, and f_{photo} takes time of the year into account.

$$\mu(A, N, C, T, t) = f_{area} f_{temp} f_{photo} \times \min(1 - N_{min}/N, 1 - C_{min}/C) \tag{2.13}$$

2.3.3 Effect of size

The gross specific growth rate depends on the size of the plant. Small sporophytes will grow relatively faster than larger ones. When the frond area is large, the growth rate is limited, and when the frond areas are close to 0, the growth rate stays high. A_0 determines in what area the specific growth rate will drop significantly.

$$f_{area}(A) = m_1 \exp(-(A/A_0)^2) + m_2 \tag{2.14}$$

2.3.4 Effect of temperature

The equation for the effect of temperature is given by Equation 2.15. Here the temperature is given in degrees Celsius.

$$f_{temp}(T) = \begin{cases} 0.08T + 0.2 & \textit{for } -1.8 \leq T < 10 \\ 1 & \textit{for } 10 \leq T \leq 15 \\ 19/4 - T/4 & \textit{for } 15 < T \leq 19 \\ 0 & \textit{for } T > 19 \end{cases} \tag{2.15}$$

2.3.5 Photoperiodic effect

f_{photo} decreases the growth rate whenever the day length decreases and increases when the day length increases. This is because the day length triggers changes in the growth pattern. Equation 2.16 takes this into account.

$$f_{photo}(n) = a_1(1 + \text{sgn}(\lambda(n))|\lambda(n)|^{1/2}) + a_2 \tag{2.16}$$

2.3.6 Apical frond loss

The two most dominating factors for apical frond loss are the age of tissue and water motion. Longer laminae more easily erode than short ones. In this case, the age of the tissue is not taken into account, but rather the size of the frond. When the frond is very small, the erosion is negligible. The model does not consider the water movement effect on frond erosion because little quantitative information is available.

$$\nu(A) = \frac{10^{-6}\exp(\varepsilon A)}{(1 + 10^{-6}(\exp(\varepsilon A) - 1))} \tag{2.17}$$

2.3.7 Nitrogen reserves and nutrient uptake

The total amount of nitrogen in the organism is the sum of structural and reserve nitrogen. A fixed fraction of the structural mass is devoted to nitrogen, which is the variable N_{struct}. Reserve nitrogen is spent on the growth of the structural mass and is the variable N. N has the maximal and minimal value N_{max} and N_{min}. The total minimal nitrogen content per unit structural mass is given by $N_{struct} + N_{min}$. The differential equation for the rate of change in nitrogen reserves is given in Equation 2.18.

$$\frac{dN}{dt} = k_A^{-1}J - \mu(N + N_{struct}) \tag{2.18}$$

The nutrient uptake rate per unit area, J, is given by Equation 2.19 where X is the external nutrient concentration, K_X is the half-saturation constant for nitrogen uptake, $U_{0.65}$ is the current speed at which the uptake is 65% of the optimal uptake and J_{max} is the maximal theoretical uptake rate under ideal conditions. The rightmost factor is a Holling type 2 functional response. The factor in the middle considers internal nutrient reserve concentration. The leftmost factor takes water current speed on the uptake rate into consideration.

$$J = J_{max}\frac{X}{K_X + X}\left(\frac{N_{max} - N}{N_{max} - N_{min}}\right)\left(1 - \exp\left(\frac{-U}{U_{0.65}}\right)\right) \tag{2.19}$$

2.3.8 Carbon reserves, photosynthesis, and respiration

The total amount of carbon is the sum of structural and reserve carbon. The fraction of structural carbon to structural dry weight is C_{struct}, with a minimum value

of C_{min} and no maximal value. The minimal carbon content is $C_{struct} + C_{min}$. The differential equation for carbon is given in Equation 2.20. The function P and R give the gross photosynthesis and respiration. E describes the exudation rate that combines exuded and leaked carbohydrates. μ is given by Equation 2.13 and k_A and C_{struct} are constants found in Table 2.3. While C is the carbon content in that respective time step.

$$\frac{dC}{dt} = k_A^{-1}(P(I,T)(1 - E(C)) - R(T)) - (C + C_{struct})\mu \tag{2.20}$$

Gross photosynthesis, P, is given by Equation 2.21. I is irradiance, and α is photosynthetic efficiency found in Table 2.3. β and $P_S(T)$ is given by Equation 2.22 and Equation 2.23.

The photoinhibition parameter, β, is calculated by solving Equation 2.23 by Newton's method, with start value $\beta_0 = 1 \times 10^{-9}$ and ten iterations. This value is then used to calculate Equation 2.22.

$$P(I,T) = P_S\left(1 - \exp\left(-\frac{\alpha I}{P_s(T)}\right)\right)\exp\left(-\frac{\beta I}{P_s(T)}\right) \tag{2.21}$$

$$P_S = \frac{\alpha I_{sat}}{\ln(1 + \alpha/\beta)} \tag{2.22}$$

$$P_{max} = \frac{\alpha I_{sat}}{\ln(1 + \alpha/\beta)}\left(\frac{\alpha}{\alpha + \beta}\right)\left(\frac{\beta}{\alpha + \beta}\right)^{\frac{\beta}{\alpha}} \tag{2.23}$$

The temperature also decides the maximal photosynthetic rate P_{max}. We therefore need to include Equation 2.24, when calculating β in Equation 2.23. P_1 is the maximal photosynthetic rate at reference temperature T_{P1}. T_{AP}, T_{APL} and T_{APH} are estimated Arrhenius temperatures. All of these are found in Table 2.3.

$$P_{max}(T) = \frac{P_1 \exp\left(\frac{T_{AP}}{T_{P1}} - \frac{T_{AP}}{T}\right)}{1 + \exp\left(\frac{T_{APL}}{T} - \frac{T_{APL}}{T_{PL}}\right) + \exp\left(\frac{T_{APH}}{T_{PH}} - \frac{T_{APH}}{T}\right)} \tag{2.24}$$

The respiration is also temperature dependent and obeys Equation 2.25. r_1 is the respiration rate at the reference temperature T_{R_1}. T_{AR} is the Arrhenius temperature estimated from respiration rates at T_{R_1} and T_{R_2}. These are constants found in Table 2.3.

$$R(T) = r_1 \exp\left(\frac{T_A}{T_1} - \frac{T_A}{T}\right) \tag{2.25}$$

The carbon exudation rate is given by Equation 2.26. This is deducted directly from photosynthesis. γ controls the rate at which carbohydrates are exuded, and both C_{min} and the exudation parameter are displayed in Table 2.3.

$$E(C) = 1 - \exp(\gamma(C_{min} - C)) \tag{2.26}$$

2.3.9 Nitrogen and carbon content

The nitrogen and carbon content, represented as a fraction of dry weight, are computed based on Equation 2.27 and Equation 2.28, respectively. The total carbon and nitrogen contents are calculated by Equation 2.29 and Equation 2.30. The dry weight, W_d, is calculated according to Equation 2.31, and the structural weight, W_s, is given by Equation 2.32.

$$N_{content} = \frac{N_{total}}{W_d} \tag{2.27}$$

$$C_{content} = \frac{C_{total}}{W_d} \tag{2.28}$$

$$N_{total} = (N + N_{struct})W_s \tag{2.29}$$

$$C_{total} = (C + C_{struct})W_s \tag{2.30}$$

$$W_d = k_A(1 + k_N(N - N_{min}) + N_{min} + k_C(C - C_{min}) + C_{min})A \tag{2.31}$$

$$W_s = k_A A \tag{2.32}$$

2.3.10 Model parameters

This model's photosynthetic efficiency, α, were selected based on literature from Lüning [28][29], with a value slightly higher than that used in his study.

The irradiance for maximal photosynthesis, I_{sat}, is reached when the maximal photosynthetic rate, P_{max}, is achieved. Photoinhibition occurs at irradiances higher than this value. The I_{sat} value is taken from Bartsch et al. [3], with an average of 215 μmol photons m^{-2} s^{-1}, although the values show considerable variability.

J_{max} represents the maximal theoretical uptake rate under ideal conditions. The maximal nitrate uptake rates for *S. latissima* reported in the literature vary significantly. Broch et al. highlight values ranging from 4.6 to 14.6 μmol g $dw^{-1}h^{-1}$

[9]. The growth model uses uptake rates per unit projected frond area, selecting the value 10 μmol g dw^{-1}h^{-1}, resulting in $J_{max} = 1.4 \times 10^{-4}$ g N dm^{-2}h^{-1}.

k_A is the structural dry weight per unit area. This value is chosen based on literate giving the dry weight per unit area. Values range from 0.18 to 1.35 g dw dm^{-2}. Broch et al. obtained values of 0.89-1.46 g dw dm^{-2} by subtracting the laminaran and mannitol from the results in Lüning [28]. The actual dry weight per area may be significantly greater.

The mass of carbon reserves per gram carbon, k_C, is estimated from the assumption that the carbon reserves are composed, in molar amounts, of 70% mannitol and laminaran, 10% alginate, and 20% of the reserve is stored as carbon. It is assumed that mannitol and laminaran exist in equal molar amounts, resulting in that 1 gram of carbon corresponds to 2.1213 g reserves.

N_{min} represents the minimal reserve nutrient (nitrogen) level, while N_{max} is the maximal nitrogen reserve. Nitrogen content lies within the range of 0.7-4.5% of dry weight, while carbon content ranges from 19-38% of dry weight [40][3].

3

Method

This section documents the work performed in this project and how the results were archived. Section 3.1, 3.2, 3.3.1 and 3.3.2 are restated material from the project thesis, with some minor modifications incorporated [37].

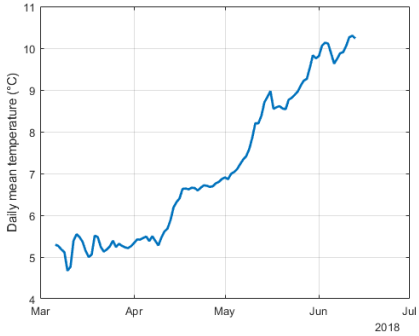
3.1 Software

In this thesis, the growth model and ensemble Kalman filter were implemented using MATLAB, a versatile and powerful programming language suitable for simulating complex systems. MATLAB's matrix calculations and data visualization capabilities make it an ideal choice for such tasks.

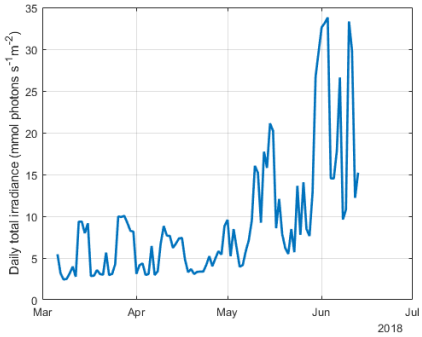
3.2 Dataset

SINTEF Ocean provided the measured values for temperature, irradiance, and nutrient concentration. These values were used in their GENIALG report from 2021 [8]. The dataset contained measured temperature, irradiance, and nutrient concentration taken every 15 minutes at four different depths, 2,5, 3, 4, and 5 meters, from March 6th 2018 to June 13th 2018. These values represent typical values for the growth season and are well-equipped for implementing the growth model and developing the EnKF. The measurements are from the Seaweed Solutions farm at Frøya, Trøndelag, Norway. In this project, only the results for 3 meters depth were used. The development of the environmental conditions can be seen in Figure 3.1. This figure represents the environmental conditions using daily mean temperature, nutrient concentration, and daily total irradiance. The measurements taken every 15 minutes were used in the simulation. As shown in the figure, temperature,

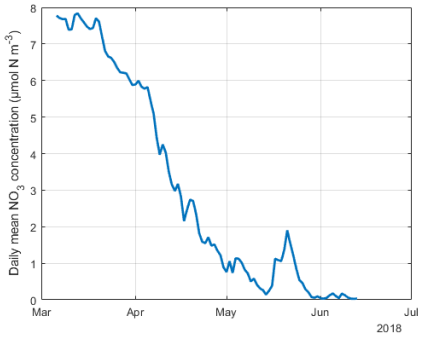
and irradiance increase as the season progresses while nutrient concentration decreases. This is due to higher temperatures promoting the growth of phytoplankton and other microorganisms that consume nutrients, as well as increased irradiance promoting the growth of photosynthetic organisms such as plants and algae, which consume nutrients during photosynthesis. This results in lower nutrient concentration in the waters [5].



(a) Daily mean temperature (°C).



(b) Daily total irradiance (mmol photons $s^{-1}m^{-2}$).



(c) Daily mean nutrient concentration (μmol N m^{-3}).

Figure 3.1: Daily environmental data at 3 meters depth from dataset.

The initial values for the frond area of the kelps were also provided by the GENIALG report by SINTEF Ocean [8]. The dataset contained measurements from five ropes, each containing ten kelp plants, resulting in 50 data points. Each data point represents the frond area of a single kelp plant. These values can be seen in Table 3.1.

Table 3.1: Measured values of frond area (dm^{-2}) of kelp.

Rope 1	Rope 2	Rope 3	Rope 4	Rope 5
0.6216	0.1757	0.8029	0.4144	0.444
0.296	0.2775	0.5698	0.8584	0.8954
0.09435	0.43735	0.1776	0.185	0.222
0.0555	0.6512	0.0888	0.4144	0.9657
0.3108	0.444	0.3552	0.1628	0.1332
0.1332	0.26825	0.1628	0.2664	0.68265
0.30535	0.111	0.4403	0.5661	0.58275
0.30525	0.666	1.073	1.2617	0.5957
0.2035	0.0888	0.3515	0.4736	0.259
0.2035	0.6993	0.3774	0.6808	0.7992

The measured values for the area throughout the season were also obtained from the GENIALG report by SINTEF Ocean [8]. The dataset comprised measurements from five ropes, with each rope containing ten kelp plants, yielding a total of 50 data points. These measurements were conducted six times between the beginning of March to mid-June. The progression of these 50 data points over time can be observed in Figure 3.2. The observed value in the EnKF was determined as the average value of these 50 data points.

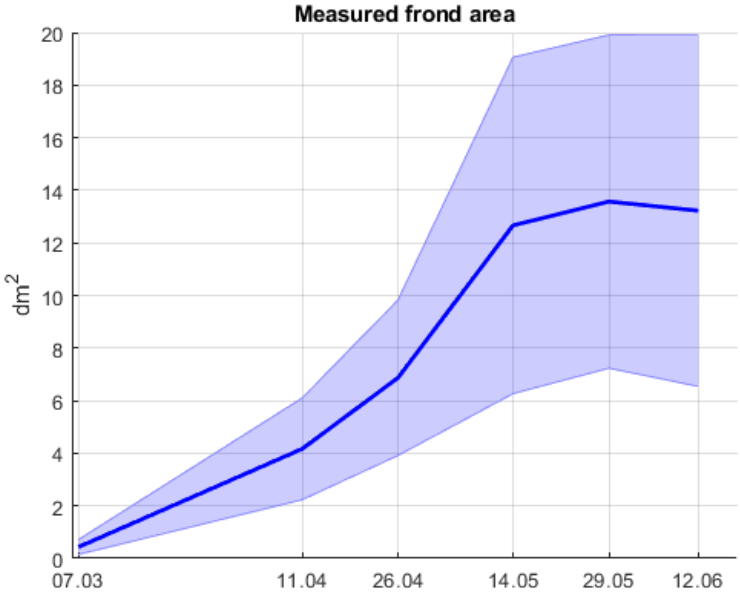


Figure 3.2: Development of the mean value and the standard deviation of the measured frond area of all of the five ropes.

3.3 Implementation

3.3.1 Growth model

The growth model described in section 2.3 was implemented in MATLAB. All parameters from Table 2.3 were initially defined. Since most variables were time-dependent, the script incorporated a loop that iterated over the duration of the growth period. In this manner, the current values for area, nitrogen, and carbon were employed to compute the time derivative of their respective values at each time step.

The initial values for the model are presented in Table 3.2. A_0 represents the mean of all values, as illustrated in Table 3.1. N_0 and C_0 were gathered from the GENIALG project [8]. The Forward Euler method was employed to solve the ordinary differential equations for A_{n+1}, N_{n+1}, and C_{n+1}.

Table 3.2: Initial values for the growth model.

State	Unit	Value
A_0	dm^2	0.4352
N_0	$gN(gw)^{-1}$	0.01
C_0	$gC(gw)^{-1}$	0.05

The Forward Euler formula can be seen in Equation 3.1. This is a first-order numerical method for solving ordinary differential equations with a given initial value. In this formula, Y_n denotes the current time step, F symbolizes the differential equation, and h represents the time step size. The current values of A, N, and C at each time step were used to compute their respective rates of change. These values and h were utilized to determine the next time step values for A, N, and C. The dataset mentioned in section 3.2 contained temperature, irradiance, and nutrient measurements at 15-minute intervals. The rates of change for area, nitrogen, and carbon were expressed in units of change per day. Consequently, h equaled 15/1440 (since there are 1440 minutes in a day). Calculations for the rates of change and the values of A, N, and C were carried out for each time step in the dataset, corresponding to 9505 dates.

$$Y_{n+1} = Y_n + hF(x_n, y_n) \tag{3.1}$$

In the present simulation, f_{photo} was assigned a constant value of 2. This parameter signifies the seasonal impact on the growth rate. Since only a short period was investigated, the minor variations in this value that may occur between March and June were insignificant and did not need to be considered.

To calculate the value of β discussed in subsection 2.3.8, the MATLAB function $fminsearch$ was employed to identify the minimum of the function derived by subtracting Equation 2.23 from Equation 2.24 and appending a negative sign. $fminsearch$ is a non-derivative technique for discovering the minimum of a multivariate, unconstrained function. It initiated with β_0 equal to 1 x 10^{-9} and located a local maximum for the function due to the sign alteration. The MATLAB code presented below illustrates the application of this concept, where P_{max} was defined according to Equation 2.24.

```
%% Finding the maximum value of beta
B = zeros(Nsample, 1);
B_0 = 1*10^(-9);
for i = 1:Nsample
  fun = @(B) -((alpha(i)*I_sat(i))/
  (log(1+alpha(i)/B)) * (alpha(i)/(alpha(i)+B)) *
  ((B/(alpha(i)+B))^(B/alpha(i))) - P_max(i));
  B_new = fminsearch(fun, B_0);
  B(i) = B_new;
end
```

3.3.2 Ensemble model

Ensemble simulation was accomplished by transforming the environmental variables and initial values into vectors containing N perturbed samples of the entire dataset. This approach enabled the creation of multiple system simulations by employing distinct input values, parameters, or initial conditions for each simulation. The EnKF utilizes a N ensemble of state estimates, each generated by randomly perturbing the initial estimate using a set of noise parameters. These perturbations are a method for sampling the potential range of solutions to the system's equations and representing the uncertainty in the model and the initial values. This is crucial when handling high-dimensional, nonlinear systems and is not there to address the biological diversity within the kelp.

In each time step for each of the N samples, the environmental variables temperature, irradiance, and nutrient concentration, not current, were perturbed by adding Gauss-Markov noise to the dataset, seen in Table 3.3. A Gauss-Markov process is characterized by a noise following a Gaussian distribution that accounts for the previous time step. One of the main properties of this process, and the primary reason for its use in this thesis, is its exponentially decreasing autocorrelation [10].

When representing uncertainty in input values for temperature, irradiance, and nutrient concentration, it was reasonable to assume that if an error occurred in interpreting a value, the same error was likely to happen in the future or had already occurred. The deviation in the Gauss-Markov process depends on the error in the previous time step, in contrast to white noise, where the value is independent. The autocorrelation in the Gauss-Markov process should match the expected rate of change of the error in the input values, which was related to the rate of change of the variables for nutrient concentration, irradiance, and temperature.

The equations for simulating the next value x_{t+1} of a Gauss-Markov process

with parameters β, σ, and time step Δt is shown in Equation 3.2. The equations for the Gauss-Markov process were taken from the book of Brown et al. [10].

Perturbations were only added to the irradiance data during the daytime, as the irradiance value is 0 at night.

$$x_{t+1} = f x_t + \sqrt{1 - f^2}\mathcal{N}(0, \sigma^2) \qquad (3.2)$$
$$f = e^{-\beta \Delta t}$$

Table 3.3: β, σ and Δt values for perturbations.

Environmental condition	β	σ	Δt
Temperature	0.4	0.708	0.0104 day^{-1}
Irradiance	0.4	19.634	0.0104 day^{-1}
Nutrient Concentration	0.4	0.305	0.0104 day^{-1}

In this study, the initial values and parameters were perturbed by incorporating Gaussian noise, as detailed in Table 3.4. The initial area was perturbed using the mean value obtained from the measured data presented in Table 3.1. For the parameters, a standard deviation (σ) corresponding to 20% of their initial values was utilized for perturbation.

Table 3.4: σ values for perturbation of parameters and states.

State	Non perturbed value	σ
A_0	0.4352	0.087
N_0	0.01	0.002
C_0	0.05	0.01
k_A	0.03	0.006
N_{min}	0.01	0.002
J_{max}	1.4 $\times 10^{-4}$	2.8 $\times 10^{-5}$
α	3.75 $\times 10^{-5}$	7.5 $\times 10^{-6}$
I_{sat}	200	40
k_C	2.1213	0.04243

3.3.3 Ensemble Kalman filter

In this study, the ensemble Kalman filter was implemented to be active at each time step where an observation was available. Without observations, the output

variables for area, nitrogen, and carbon were predicted solely based on the mathematical growth model presented in section 2.3, and no correction step was performed.

The article by Halvorsen et al. is used for inspiration to explain how the EnKF was implemented [22]. The forecast ensemble consisted of N ensemble members, each containing n state variables, which can be written as the $n \times N$ matrix, Equation 3.3. Here, the column vectors were the state vectors $\boldsymbol{x}_i$ of dimension n. The forecast step was given by adding the forecast-model error, $\boldsymbol{w_i}$, with the expectation value 0 and variance corresponding to the values in $\boldsymbol{Q}$. The values in $\boldsymbol{Q}$ were decided in advance. The ensemble mean was given by Equation 3.4.

$$\boldsymbol{x}^f = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N], \quad \boldsymbol{x}_i = \boldsymbol{x}_{i-1}^a + \boldsymbol{w}_i \quad \boldsymbol{w}_i \sim \mathcal{N}(0,\, \boldsymbol{Q}) \tag{3.3}$$

$$E(\boldsymbol{x}^f) = \frac{1}{N} \sum_{k=1}^{N} \boldsymbol{x}_k \tag{3.4}$$

The ensemble spread is defined by Equation 3.5. The vector $\boldsymbol{e}$, whose entries are all ones, was introduced as it is closer in terms of notation to how the averaging is implemented numerically.

$$\boldsymbol{A} = \boldsymbol{x}^f - E(\boldsymbol{x}^f) = \boldsymbol{x}^f - \frac{1}{N}(\boldsymbol{x}^f \boldsymbol{e}_{N\times 1})\boldsymbol{e}_{1\times N} \tag{3.5}$$

This gave the ensemble covariance matrix, shown in Equation 3.6

$$\boldsymbol{P}^f = \frac{\boldsymbol{A}\boldsymbol{A}^T}{N-1} \tag{3.6}$$

Measurements were organized in a vector, $\boldsymbol{d}$, with a size equal to the number of measurements, $m \times 1$. The corresponding measurement error covariance matrix, $\boldsymbol{R}$, was of size $m \times m$. Independent random perturbations were drawn from a normal distribution with expectation 0 and variance corresponding to the values in $\boldsymbol{R}$. The values in $\boldsymbol{R}$ were decided in advance. These perturbations were added to N instances of the measurement vector $\boldsymbol{d}$ to form a measurement matrix $\boldsymbol{D}$ with size $m \times N$, seen in Equation 3.7.

$$\boldsymbol{D} = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N], \quad \boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{v}_j, \quad \boldsymbol{v}_j \sim \mathcal{N}(0,\, \boldsymbol{R}) \tag{3.7}$$

Here $\boldsymbol{v}_j$ was the independent random perturbations with standard deviation corresponding to the uncertainty of each measurement. The analysis step was then given by

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \boldsymbol{P}^f \boldsymbol{H}^T (\boldsymbol{H}\boldsymbol{P}^f \boldsymbol{H}^T + \boldsymbol{R})^{-1} (\boldsymbol{D} - \boldsymbol{H}\boldsymbol{x}^f) \tag{3.8}$$

Here, $\boldsymbol{H}$ was the observation matrix with rows indicating the linear combination of state variables corresponding to each measurement.

To avoid computing $\boldsymbol{P}^f$ directly Equation 3.8 was rewritten to Equation 3.9.

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \frac{1}{N-1}\boldsymbol{A}(\boldsymbol{HA})^T\boldsymbol{P}^{-1}(\boldsymbol{D} - \boldsymbol{Hx}^f) \tag{3.9}$$

where

$$\boldsymbol{P} = \frac{1}{N-1}\boldsymbol{HA}(\boldsymbol{HA})^T + \boldsymbol{R} \tag{3.10}$$

$$\boldsymbol{HA} = \boldsymbol{Hx}^f - \frac{1}{N}((\boldsymbol{Hx}^f)\boldsymbol{e}_{N\times 1})\boldsymbol{e}_{1\times N} \tag{3.11}$$

A simplified illustration of the EnKF's implementation can be seen in Figure 3.3.

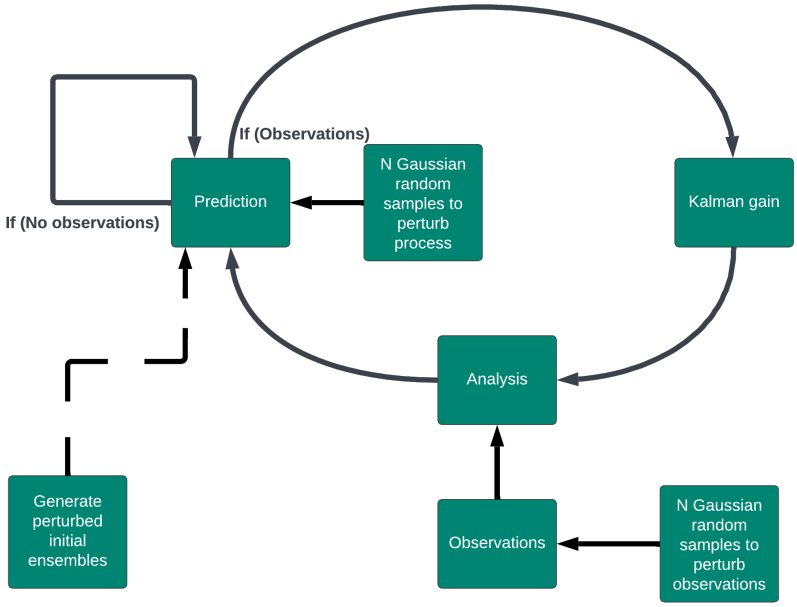


Figure 3.3: Streamlined depiction of the EnKF's workflow as applied to this project. The primary step is the initial ensembles. Visualization inspired by the flow chart for the EnKF in the book by Brown et al. [10].

Parameter estimation was implemented using a similar setup as described above; by incorporating the parameters into the state vector in Equation 3.3. Unlike the area, carbon, and nitrogen variables, the forecast values for the parameters were

solely determined by the previous parameter estimation, with perturbations, as the parameters did not possess derivatives like the model states. Only one parameter was examined at a time during the parameter estimation to prevent the parameters from influencing each other and yielding inaccurate results.

3.3.4 Twin experiment

The twin experiment was carried out by creating a duplicate of the model, referred to as the *true* value. This duplicate did not undergo any analysis or correction steps but solely relied on the calculations derived from the mathematical model, as explained in subsection 3.3.1. In the twin experiment, the true value is the measurement vector, corresponding to $\boldsymbol{d}$ in Equation 3.7. This approach provides observations for all three output values: area, nitrogen, and carbon. As for parameter estimation, it is not anticipated that these will be directly measurable; thus, no parameter observations were included.

During the twin experiment simulation, the actual value for k_A was set to be 0.039 gdm^2, while the estimated value was kept at the prescribed level of 0.03. This systematic error was deliberately introduced to ensure a distinction between the twin model and the estimation model. This discrepancy allows for a more in-depth understanding of the corrections introduced by the filter. The rationale behind embedding the error into the parameter k_A is due to its impact on the results of all three individual states, enabling conclusions to be drawn from each state. This feature was integrated into all simulations except for the parameter estimation phase.

The twin experiment has primarily been used to ascertain if any recommendations can be drawn. Assessing the twin experiment was employed to examine hypothetical cases, determining whether they could apply to real-world measurements. If a scenario is not theoretically possible, it is likely not feasible in reality; conversely, if it is theoretically possible, it may also be achievable in real-life situations.

3.3.5 Parameter estimation

The parameter selection used for parameter estimation was based on the sensitivity analysis conducted in the project thesis [37]. This sensitivity analysis reveals that the model's sensitivity varies depending on which output values, such as area, carbon, or nitrogen, are being examined. However, a general conclusion drawn from the analysis is that the model is susceptible to changes in k_A, N_{min}, N_{max}, J_{max}, α, I_{sat}, and k_C. Since the model is sensitive to these parameters, they significantly impact the system's output, making it crucial to obtain accurate values for these parameters. As mentioned in section 2.3, many parameters have been

collected and estimated from literature but may not be specially tailored to the Norwegian environment. In some cases, limited data is available for these values, which can contribute to uncertainty in the estimates.

More specifically, the predictions of the frond area are sensitive to the parameters N_{min}, k_A, J_{max}, and N_{max}. The values of N_{max} and N_{min} are used together in Equation 2.19; since they produce a ratio, the primary focus has been on estimating N_{min}. The nitrogen predictions are sensitive to the parameters α and I_{sat}. Carbon predictions are also sensitive to α as well as k_C.

3.4 Evaluation

This section will present various case studies explored to examine the application and derive insights from it. A summary of the tests for the twin experiment is presented in Table 3.8 and further elaborated in the following sections.

In testing the implementation, a distinction is made between measurements and twin experiments. Some tests have been conducted solely for the twin experiment, while others have been performed for both the twin experiment and with actual field data.

3.4.1 Metrics

Pearson's correlation coefficient was used to investigate the correlation between the different states, given in Equation 3.12. In this context, μ_A and σ_A denote the mean and standard deviation of A, while μ_B and σ_B signify the mean and standard deviation of B. The correlation coefficient of two random variables quantifies their linear interdependence [30]. This coefficient ranges from -1 to 1, with -1 suggesting a perfect negative correlation and 1 indicating a perfect positive correlation. In other words, if one variable increases, the other variable decreases in a negative correlation, while both variables increase in a positive correlation. A correlation coefficient of 0 signifies no observable relationship between the variables. The correlation coefficient was calculated for the entire ensemble at each time step to monitor the evolution of the correlation over time. The correlation analysis was restricted to the output of area and nitrogen and area and carbon and only applied to the estimations derived from field data, not the twin experiment.

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right) \tag{3.12}$$

Various approaches have been employed to assess the accuracy of the estimates obtained from the different cases. One method involves calculating the error between the mean value of the ensemble and the measurement. The formula for this

is shown in Equation 3.13, where y represents the true value or measurement value and $\hat{y}$ stands for the estimated value. In some cases, to assess the precision of the prediction, the absolute error was utilized, shown in Equation 3.14.

$$E = y - \hat{y} \tag{3.13}$$

$$AE = |y - \hat{y}| \tag{3.14}$$

Another approach to evaluate the accuracy of predictions is the root mean square error (RMSE). The equation for RMSE is given in Equation 3.15. RMSE quantifies the discrepancy between the true values and the predicted values by taking the square root of the average of the squared differences between the actual and predicted values [31]. Both RMSE and absolute error suggest that lower values indicate a better fit of the predictions to the actual model. A value of zero would imply a perfect prediction. Throughout this project, an estimation close to the true or measurement value has been used as a benchmark for assessing the accuracy of the prediction.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y_i})^2}{n}} \tag{3.15}$$

In order to represent the precision of the prediction with a single value, the mean value of the RMSE has been employed, as referenced Equation 3.16. The smaller this value is, the more accurate the prediction.

$$RMSE_{mean} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y_i})^2}{n}} \tag{3.16}$$

3.4.2 Measurement frequency

The frequency of measurements, or sampling strategy, was examined using the twin experiment. This was accomplished by altering the availability of the true values. The various cases investigated are presented in Table 3.5.

Table 3.5: Sampling strategies.

Sampling
Once a day
Every other day
Once a week
Every other week
Once a month
Once a month to mid-May, then once a week after mid-May
Once a week after mid-May
Every other day after mid-May

3.4.3 Available measurements

Another scenario investigated involved varying the availability of measurements. This was done to comprehend how the measures influence the estimates of the different states. The cases examined are listed in Table 3.6 and were only carried out on the twin experiment.

Table 3.6: The different cases for available measurements.

State
Area
Area and carbon
Area and nitrogen
Area, nitrogen, and carbon
Nitrogen
Nitrogen and carbon
Carbon

3.4.4 Consistency in measurements

To investigate the importance of maintaining consistency in the measurements, the filter was simulated ten times with the true value made available from 0 to 2 days before and after the original measurement date, which was set precisely seven days apart. This test was conducted solely on the twin experiment.

3.4.5 Parameter estimation

In the twin experiment, the precision of parameter estimation, as discussed in subsection 3.3.5, was examined in two ways. Firstly, the stability of parameter estimation was assessed by verifying whether the estimate tracked the true value from the twin when the parameter value for the twin was altered. The modified values are presented in Table 3.7. The parameter value for the twin experiment was increased by 30% of the value specified in Table 2.3 in mid-April.

Table 3.7: The parameter values for the twin experiment conducted to evaluate parameter estimation.

Parameter	Original value	Modified value
N_{min}	0.01	0.013
J_{max}	1.4 $\times 10^{-4}$	1.82 $\times 10^{-4}$
k_A	0.03	0.039
α	3.75 $\times 10^{-5}$	4.875 $\times 10^{-5}$
I_{sat}	200	260
k_C	2.1213	2.7577

The second approach to examine the parameter estimation outcome involved analyzing the results of the estimates of the states. The question was whether allowing parameter estimation led to more precise state estimates. This case was tested by setting the parameter value for the twin experiment equal to the modified value listed in Table 3.7 for the whole simulation. Measurements were only available up until the 1st of May. The simulation results with parameter estimation were compared to those with only state estimation. As mentioned in subsection 3.3.3, only one parameter was investigated at a time.

In the context of field data, parameter estimation was explored similarly to the twin experiment. The first aspect of this investigation was to examine the stability of the parameter estimation, ensuring that the parameter did not display divergence. Subsequently, the study focused on whether the parameter estimation yielded more precise predictions than what could be obtained without including the parameter in the state vector.

3.4.6 Standard deviation for measurement error

This study also examined the filter's reaction to variations in the standard deviations for the measurement error. This analysis was done by changing the respective value and was undertaken only when actual field data for the frond area was available.

Table 3.8: Summary of the different cases investigated on the twin experiment.

		σ_R			σ_Q			
Case		A	N	C	A	N	C	Time step
Sampling Strategy	Once a day	0.5	5×10^{-4}	0.01	0.2	0.0001	0.005	-
	Every other day							
	Once a week							
	Every other week							
	Once a month							
	Once a month to mid-May, then once a week after mid-May							
	Once a week after mid-May							
	Every other day after mid-May							
Available measurements	Area	0.5	-	-	0.2	0.0001	0.005	Every other week
	Area and nitrogen	0.5	5×10^{-4}	-				
	Area and carbon	0.5	-	0.01				
	Area, nitrogen and carbon	0.5	5×10^{-4}	0.01				
	Nitrogen	-	5×10^{-4}	-				
	Nitrogen and carbon	-	5×10^{-4}	0.01				
	Carbon	-	-	0.01				
Consistency in measurements		0.5	5×10^{-4}	0.01	0.2	0.0001	0.005	Once a week $\pm$ 0-2 days
Parameter estimation	N_{min}	0.5	0.001	0.01	0.2	0.0001	0.005	Once a week to 1st of May
	J_{max}							
	k_A							
	α							
	I_{sat}							
	k_C							

4

Results

In this section, the results from the simulations of the model are presented, along with the ensemble Kalman filter's state and parameter estimation, both with the twin experiment and with the use of field data.

4.1 Twin experiment

The outcome of solely estimating the states and incorporating measurements for the states' area, nitrogen, and carbon every 14 days is illustrated in Figure 4.1. The blue shaded bars represent the estimated states using the EnKF, while the black line indicates the true values derived from the twin experiment. The green shaded bars result from only using the growth model to predict the states. The bold blue and green line signifies the mean value across the ensemble, and the shaded error bar denotes the ensemble's standard deviation.

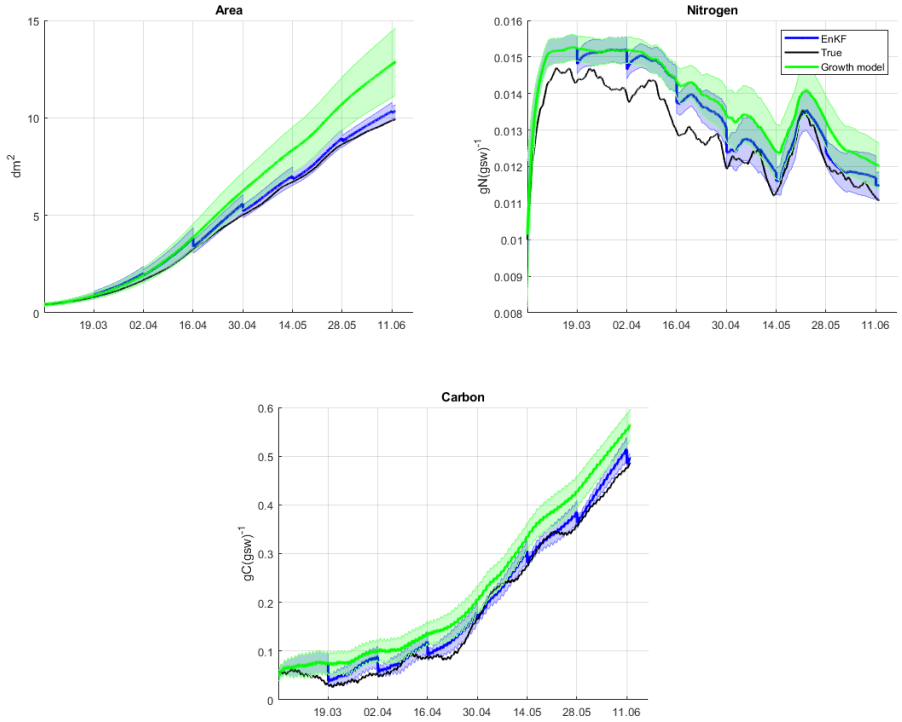


Figure 4.1: Mean value and standard deviation of estimated area, nitrogen, and carbon with and without the use of the EnKF. The black line represents the true value obtained from the twin experiment.

4.1.1 Sampling

Figure 4.2 displays the error for the three different output values using various sampling strategies. The error represents the deviation between the estimates' true and mean values. The figure illustrates the eight sampling approaches in Table 3.5.

From Figure 4.2, it is evident that the smallest error for area, nitrogen, and carbon is achieved with a daily sampling strategy. More significant deviations from the true value are observed when the sampling rate is reduced. Nonetheless, as the sampling rate increases toward the end of the simulation, the filter rapidly converges toward the true value.

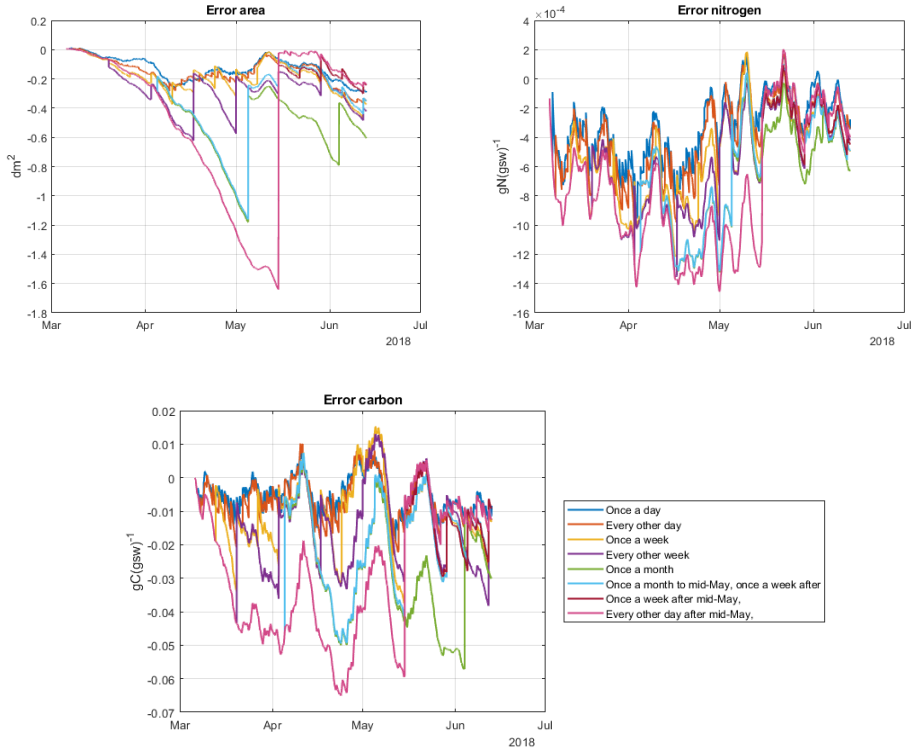


Figure 4.2: The error associated with various sampling strategies for measurements.

The impact of having measurements for only a subset of values requiring estimation is illustrated in Figure 4.3. The graph reveals that measuring at least the area is necessary to obtain a more accurate estimate of the frond area. The lowest error across the ensemble is achieved for carbon when the carbon and at least one additional variable are measured. In the case of nitrogen, the deviations between the different scenarios are not as substantial as in carbon and area. Still, the highest accuracy is when nitrogen and at least one of the other states are measured.

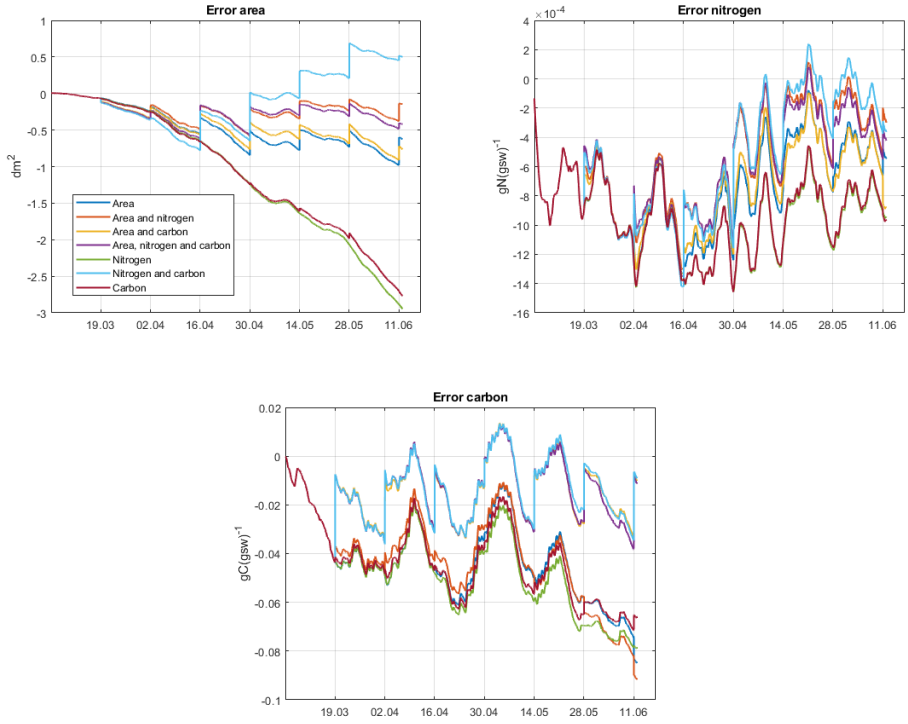


Figure 4.3: The discrepancy between the true values and the estimates of various output values when different measurements were available.

The significance of measurement consistency can be observed in Figure 4.4. In this context, consistency refers to providing measurements precisely every seventh day. The figure illustrates that all ten simulations yield approximately the same degree of error.

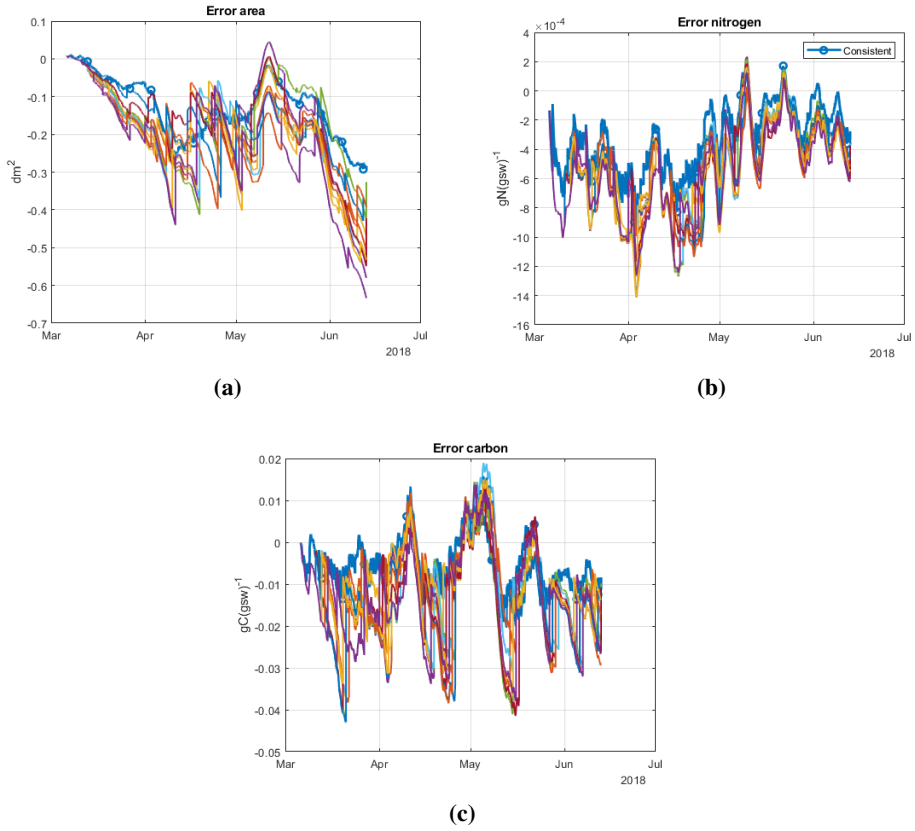


Figure 4.4: Examination of the impact of consistent measurements.

4.2 Parameter estimation

Figure 4.5 illustrates the response of various parameters when incorporating parameter estimation in the EnKF. The black line represents the true value, while the blue shaded error bar indicates the filter's response. The blue line is the true value being changed throughout the simulation. The green-shaded error bar displays the EnKF's response to this change. As observed, all parameters, except for k_C, do not diverge but follow the true value.

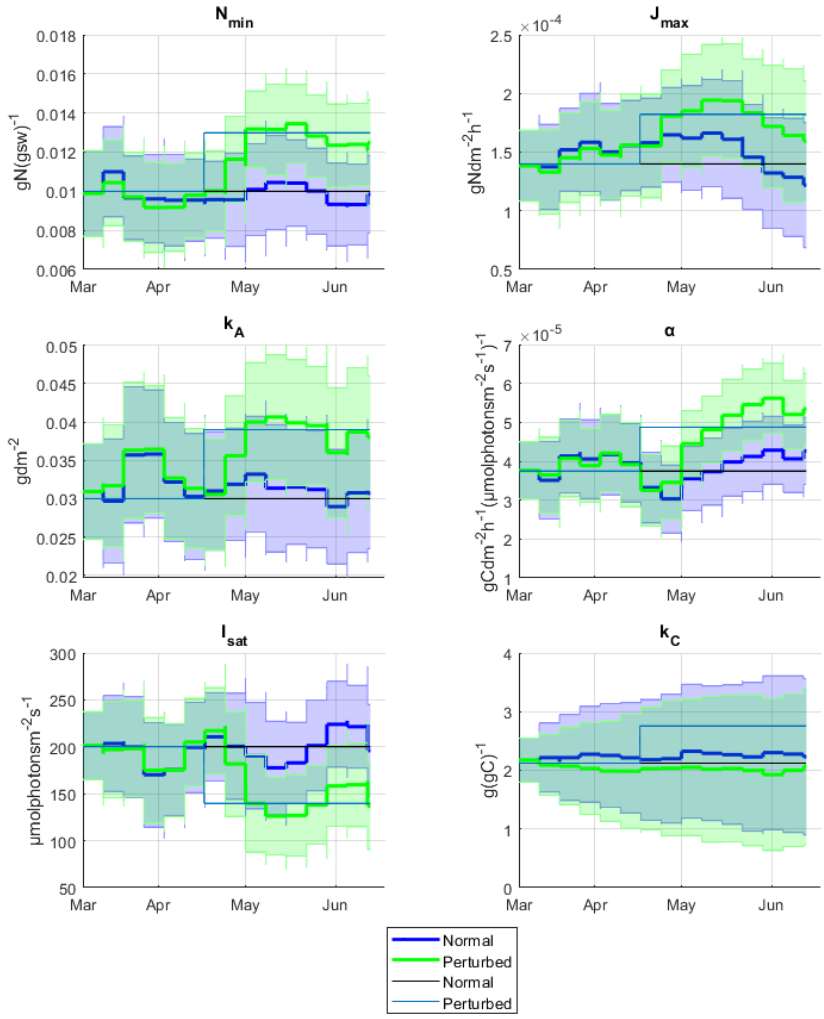


Figure 4.5: Evolution of the various parameters during the simulation when EnKF was implemented with parameter estimation.

The impact of incorporating parameter estimation in the EnKF can be observed in Figure 4.6, as tested under the approach described in subsection 3.4.5. The blue graph represents the case where the parameter was included in the state vector, while the green graph represents the filter without parameter estimation, focusing solely on state estimation. As evident from Figure 4.6, including estimation of N_{min}, J_{max}, and k_a results in significantly more accurate estimates of the area, whereas the same cannot be said for α, I_{sat}, and k_C. Estimating k_A, I_{sat}, and

α leads to more precise estimates of carbon, while estimating N_{min} and J_{max} enhances the accuracy of nitrogen estimates.

A summary of the diverse scenarios explored in the twin experiment, along with the mean value of the RMSE across the ensemble and the entire simulation, can be found in Table 4.1 and 4.2.

Table 4.1: Mean value of RMSE for the different cases studied in the twin experiment.

Case		RMSE$_{mean}$ area	RMSE$_{mean}$ nitrogen	RMSE$_{mean}$ carbon
Sampling strategies	Once a day	0.1837	4.9012 $\times 10^{-4}$	0.0102
	Every other day	0.2319	5.8081 $\times 10^{-4}$	0.0127
	Once a week	0.3172	7.3226 $\times 10^{-4}$	0.0201
	Every other week	0.4241	7.9999 $\times 10^{-4}$	0.0247
	Once a month	0.5937	8.8804 $\times 10^{-4}$	0.0372
	Once a month to mid-May, once a week after	0.4862	8.4719 $\times 10^{-4}$	0.0305
	Once a week after mid-May	0.7124	9.5574 $\times 10^{-4}$	0.0381
	Every other day after mid-May	0.6671	9.1714 $\times 10^{-4}$	0.0361
Measurements on different values	Area	0.6009	9.1915 $\times 10^{-4}$	0.0494
	Area and nitrogen	0.4067	7.9964 $\times 10^{-4}$	0.0485
	Area and carbon	0.5595	9.0398 $\times 10^{-4}$	0.0245
	Area, nitrogen and carbon	0.3172	7.3226 $\times 10^{-4}$	0.0201
	Nitrogen	1.3569	0.0011	0.0535
	Nitrogen and carbon	0.6937	8.1428 $\times 10^{-4}$	0.0247
	Carbon	1.3188	0.0011	0.0500
Consistent measurements	1	0.3220	7.3225 $\times 10^{-4}$	0.0207
	2	0.3325	7.3743 $\times 10^{-4}$	0.0201
	3	0.3453	7.4311 $\times 10^{-4}$	0.0212
	4	0.3213	7.3788 $\times 10^{-4}$	0.0210
	5	0.3369	7.3940 $\times 10^{-4}$	0.0206
	6	0.3227	7.3263 $\times 10^{-4}$	0.0207
	7	0.3189	7.2913 $\times 10^{-4}$	0.0211
	8	0.3481	7.4229 $\times 10^{-4}$	0.0200
	9	0.3343	7.3886 $\times 10^{-4}$	0.0197
	10	0.3552	7.5143 $\times 10^{-4}$	0.0212

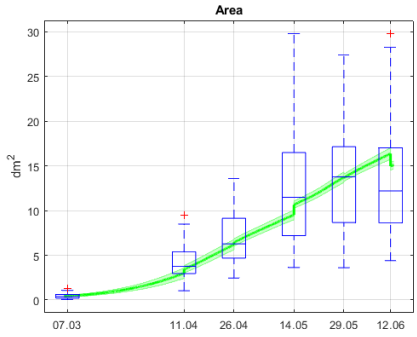
Figure 4.6: Estimation error for area, nitrogen, and carbon compared to the true values of the states, with measurements available until May 1st.

Table 4.2: Mean value of RMSE for the parameter estimation in the twin experiment. Measurements until May 1st.

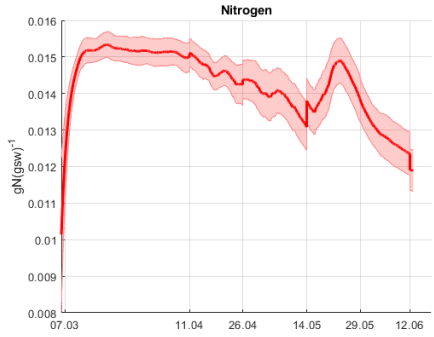
	RMSE_{mean} area		RMSE_{mean} nitrogen		RMSE_{mean} carbon	
Parameter	With	Without	With	Without	With	Without
N_{min}	0.5424	0.9860	0.0017	0.0024	0.0299	0.0201
J_{max}	0.5469	0.6985	9.8055×10^{-4}	0.0012	0.0247	0.0265
k_A	0.4380	0.7217	7.8032×10^{-4}	8.6786×10^{-4}	0.0292	0.0374
α	0.3774	0.3828	7.2740×10^{-4}	6.9035×10^{-4}	0.0465	0.0557
I_{sat}	0.350	0.3873	6.2539×10^{-4}	6.8717×10^{-4}	0.0481	0.0491
k_C	0.3829	0.4330	6.2109×10^{-4}	6.0612×10^{-4}	0.0207	0.0233

4.3 Measurements

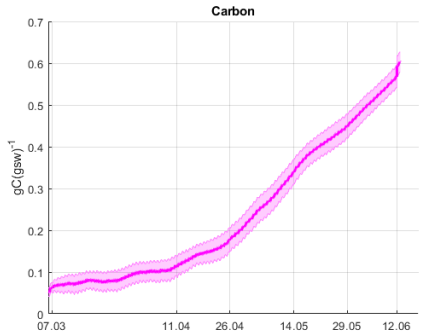
In this section, the results of the EnKF with available measurements are presented. The measures from the GENIALG report, mentioned in section 3.2, provided measurements only for the area, resulting in Figure 4.7. The mean value of this data is used as the measured value. The box plot is used to visualize the measured area. On each box, the central mark denotes the median, while the bottom and top edges represent the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually using the plus sign. Unless specified otherwise, the standard deviation for the measurement error used was 0.5 dm^2, and the standard deviation for the model error was 0.2 dm^2 for the area, 0.0001 $gN(gsw)^{-1}$ for nitrogen, and 0.005 $gC(gsw)^{-1}$ for carbon.



(a) Estimation of frond area. Box plot of the observations.



(b) Nitrogen, no observations



(c) Carbon, no observations.

Figure 4.7: Estimations of area, nitrogen, and carbon with measurements available for the area.

The progression of the correlation coefficient between area and nitrogen and area and carbon across each ensemble at every time-step is visualized in Figure 4.8.

The impact of varying the standard deviation of the measurement errors can be observed in Figure 4.9. The figure displays a heat map of the absolute error between the mean of the measured frond area and the mean of the estimated frond area when the standard deviation of the error in the measurement in the EnKF is altered. Corrections were made until April 26th, with the last measure used for comparing the results. The absolute error exhibits notable differences between the various standard deviation for measurement errors. In the heat map, the x-axis represents the dates, while the y-axis displays the outcomes of the different standard deviations. The data with the most significant error in each time step is marked by the most intense color, whereas the faintest color indicates the data with

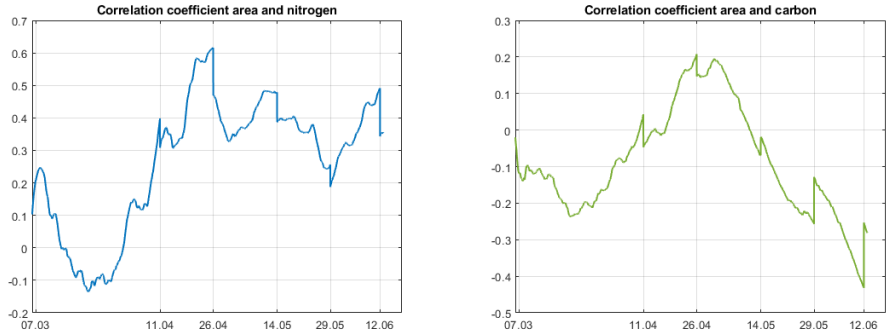


Figure 4.8: The progression of the correlation coefficient between area and nitrogen, and area and carbon.

the slightest error.

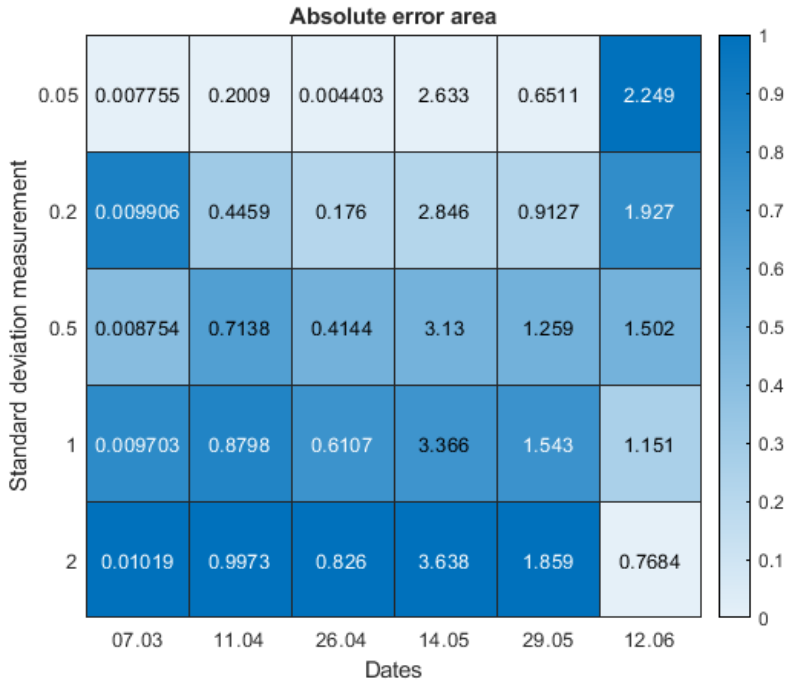
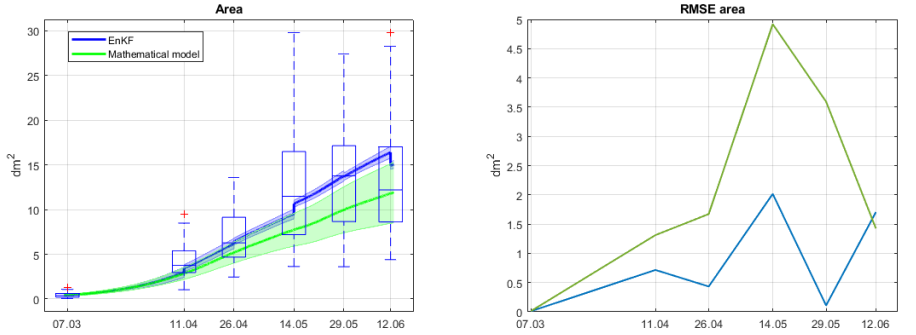


Figure 4.9: The absolute error between measured and estimated area of different standard deviation for measurement errors. The last correction was on April 26th.

Figure 4.10 illustrates the distinction between the predictions of the frond area

utilizing the ensemble Kalman filter and those without its application. The blue graph displays the estimate derived from the EnKF and the mathematical model, while the green graph represents the prediction generated solely by the growth model.



(a) Estimation of the frond area with and without using EnKF together with the growth model. Box plot of the observations.

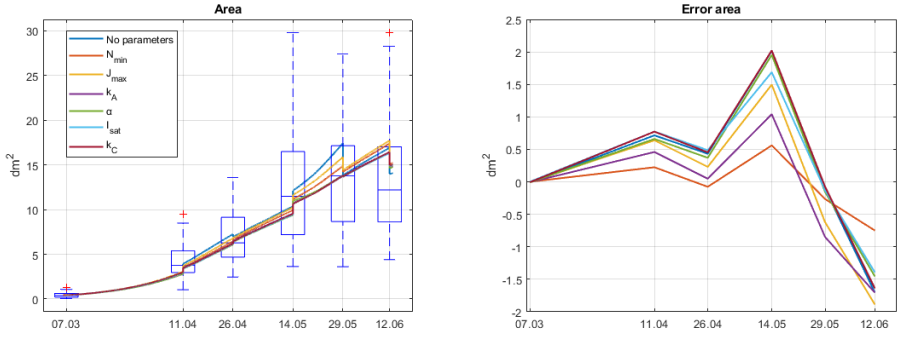
(b) Root mean square error between the estimates and the mean value of the observations.

Figure 4.10: Comparison between the estimates of the area with and without the use of the ensemble Kalman filter.

4.3.1 Parameter estimation

The results of the individual parameter estimations are illustrated in Figure 4.11 and Figure 4.12. As observed in Figure 4.12, the values for k_C, I_{sat}, and α diverge throughout the simulation. This is not the case for N_{min}, J_{max}, and k_A.

Figure 4.11 illustrates that the estimation of N_{min}, J_{max}, and k_A decreases the error between the observed and estimated frond area, especially in the initial stage of the simulation. Incorporating α, I_{sat}, and k_C in the estimation yields results approximately equivalent to those obtained with state estimation alone.



(a) Frond area under different parameter estimation scenarios. Box plots are the observed frond areas.

(b) Error between estimated and measured values when incorporating parameter estimation.

Figure 4.11: Outcomes for frond area with varying parameters incorporated in the parameter estimation process.

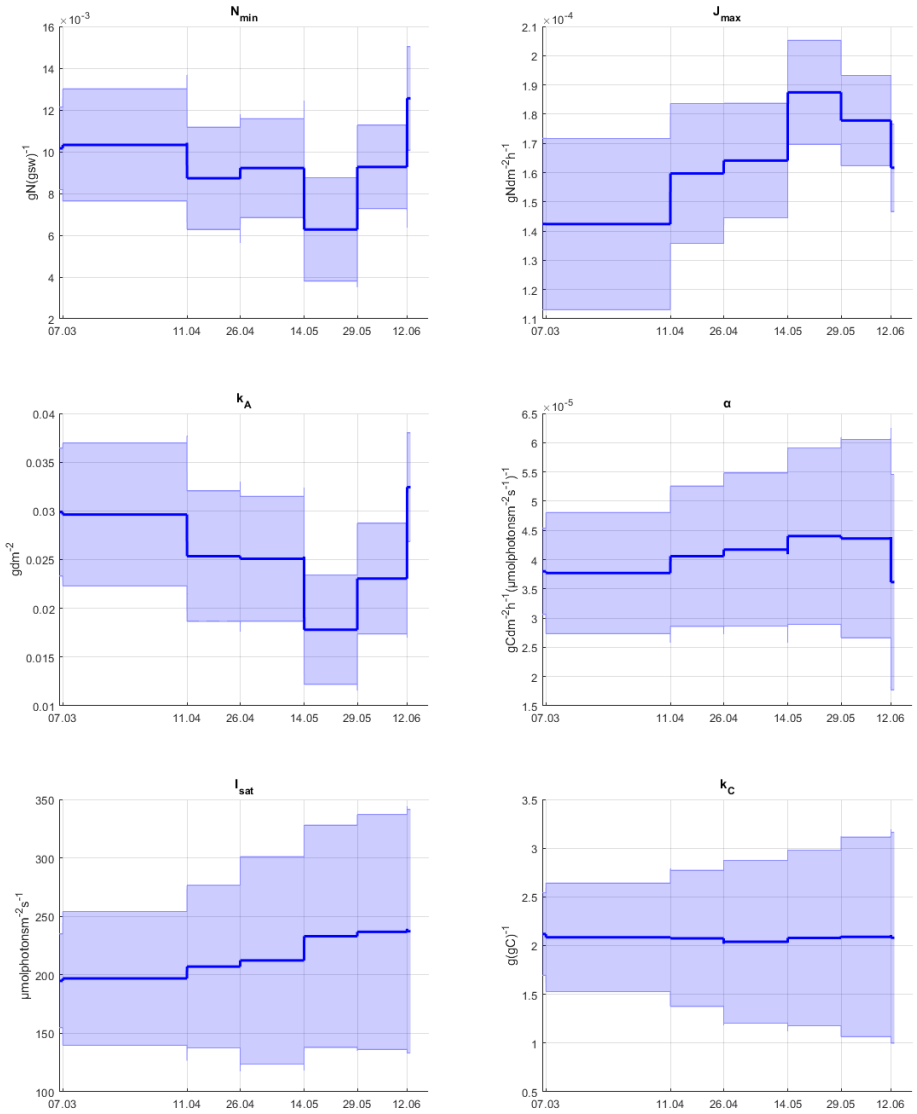


Figure 4.12: Parameter estimation for ensemble Kalman filter with measurements for the area.

A summary of the mean value of the RMSE over the ensemble and the duration of the study for the various scenarios explored in the application of the EnKF when measurements are available is provided in Table 4.3.

Table 4.3: The mean value of RMSE for the cases studied when measurements were available.

Case		RMSE Area
Different σ, measurement end April 26th	0.05	1.0448
	0.2	1.1607
	0.5	1.3249
	1	1.4733
	2	1.7277
Measurements on different values	Area	0.9505
Parameter estimation	N_{min}	0.5684
	J_{max}	0.9268
	k_A	0.8406
	α	0.8990
	I_{sat}	0.8696
	k_C	0.9570

5

Discussion

In this section, the results of the work are explained and discussed. Possible system weaknesses and sources of errors will be considered, as they could affect the result. The primary focus for discussion is the frond area of the kelp, as it represents a key state variable that is relatively easy to measure. Moreover, it directly reflects the growth and productivity of the kelp population and, in this dataset, is the only state with actual field data.

To discuss the results from the work, primarily the error graphs and the tables for the mean value of the RMSE are used. The graphs are advantageous because they display temporal progression. On the other hand, the RMSE tables provide an efficient way to assess the estimates' precision quickly. However, they are insufficient when specific simulation parts are of interest. Additionally, these tables are primarily utilized to support the conclusions.

5.1 Twin experiment

In Figure 4.1, it can be observed that when using the standard deviations for the measurement and model error provided in Table 3.8 as the default values for the standard deviation for this project, the estimates for all three states are reasonably accurate. The estimates for area and carbon demonstrate a high level of precision, while the estimate for nitrogen is not as accurate. This trend is consistent across other cases examined in this study. Despite reducing the standard deviation of the nitrogen measurement error, achieving a precise estimate remains challenging. Nonetheless, using the EnKF leads to more precise estimations for all three states compared to what could be achieved without its use.

In Figure 4.1, it can also be observed that the accuracy of the area estimation

diminishes towards the latter part of the simulation. This observation suggests that increasing the measurement frequency during the final stages of the kelp's growth period could be beneficial.

5.1.1 Sampling strategies

Figure 4.2 indicates that some specific sampling strategies outperform others. As anticipated, the filter demonstrates enhanced performance with more frequent measurements. Since kelp farms are currently situated at sea, obtaining the desired measurements of area, nitrogen, and carbon for kelp is challenging and resource-demanding. Consequently, it is essential to identify a more practical measure strategy that continues to yield valid results.

Regarding the estimation of area, the error in the estimate is more significant when the measurement strategy is every other week or less frequently. This can be seen from the mean value of the RMSE in Table 4.1. When choosing a strategy, the importance of estimation precision must be considered. As mentioned in section 1.1, if the aim is to aid kelp farmers in estimating the frond area, thus enabling them to identify the best harvest time, which is when the frond area is at its largest and thereby providing the highest biomass yield. In this case, the end of the simulation is of the highest importance.

For this scenario, a less frequent measurement strategy may be adequate. As seen in Figure 4.2, this could involve measurements taken once a month or no measures until mid-May, followed by more frequent measurements. Suppose the accuracy of the result at the end of the growth period is crucial. In that case, measurements should be conducted every other day after mid-May, which, towards the end of the simulation, yields the same outcome as measuring every other day throughout the whole simulation period. However, if the precision of the predictions for the whole simulation period is essential, this is not sufficient.

Table 4.1 shows that the different measurement strategies result in different mean values for the RMSE. Drawing from the manual of seaweed farming for the red algae *Eucheuma*, which may apply to other types of kelp, it is recommended that plants be measured either weekly or every 15 days to minimize stress on the plants [24]. With this in mind, and with a sufficiently low RMSE, it can be reasoned that a sustainable measurement strategy might involve measurements taken every other week, which may also lead to adequate resources being put into the task.

Referring to nitrogen in Figure 4.2, it is evident that the error between the estimated and true value is relatively small across all cases. However, it is clear from Table 4.1 that some sampling strategies are better than others. If maintaining precise predictions throughout the simulation is crucial, then a frequent sampling

strategy should be emphasized, preferably once a day. Nonetheless, every other week, the same sampling strategy used for the area may be sufficient. On the other hand, if accuracy towards the end of the simulation is of utmost importance, the findings from Figure 4.2 suggest that a more frequent strategy in the later stages, such as once a week after mid-May, is appropriate.

For carbon, a more considerable discrepancy between various strategies is observed throughout the simulation. It exhibits a response akin to the area where the error for measurements daily to every other week is smaller than the other cases. However, similar to the area, the error towards the end of the simulation remains relatively consistent. This implies that the measurement strategy chosen for the area, every other week, might also be appropriate for carbon.

As evident from Figure 4.3, obtaining an accurate estimate of the area requires at least measuring the area itself. Including carbon or nitrogen measurements can further improve the estimation. The most precise estimate is achieved when all states are measured. This is further confirmed by the mean of the RMSE in Table 4.1. The table also indicates that only measuring nitrogen and carbon can have almost the same accuracy as measuring only the area. However, the resulting precision is still too low to be considered a valid measurement strategy.

Among the measurements, the area is the simplest to measure as it can be determined for seaweed still attached to the cultivation line, as mentioned in subsection 2.1.2. Therefore, the optimal scenario would involve obtaining the most accurate estimate using only area measurements without considering the two other states. This approach would minimize the required resources and time investment. However, Figure 4.3 demonstrates that this is not the case, necessitating further exploration of alternative measurement strategies.

The imprecision resulting from measuring area alone is evident in the prediction of carbon. It is essential to include measurements of carbon and at least one of the other states to obtain reliable carbon estimates. The same response can also be seen for nitrogen. The measurement strategies that yield estimates closest to the true value involve measuring nitrogen alongside at least one of the other states.

This suggests that each state must be measured to obtain a reliable estimate for growth. A different approach could be not to measure all the conditions every time, instead alternating the states measured throughout the growth period. However, this particular scenario has not been investigated in this thesis.

Regarding the significance of measurement consistency, as shown in Figure 4.4, there is little to no difference between the various scenarios, which is further confirmed in Table 4.1. This suggests that measuring the states precisely on the planned date is not crucial, which is advantageous because the state measurements depend on weather and other unforeseen events.

5.1.2 Parameter estimation

From the inclusion of parameter estimation in the twin experiment, as illustrated in Figure 4.5, it becomes apparent that all parameters to which the states are particularly sensitive can be effectively estimated, except for k_C. This may also indicate that the model is not especially sensitive to this parameter after all; thus, estimating it is unnecessary. This inference for the estimable parameters can be made since the variance is constant over time, which signifies that the distribution's dispersion remains stable. The values do not diverge as observed for k_C. Additionally, when the value of a specific parameter in the twin experiment is altered, it can be noticed that the estimations track this change, again with the exception of k_C. This further reinforces that parameter estimation is advantageous and feasible.

The advantage is further confirmed in Figure 4.6, which illustrates that permitting the estimation of specific parameters yields improved state estimates compared to those obtained without parameter estimation. This finding is additionally verified in Table 4.2. However, given the interest in examining the response after the ending of the correction steps, the most critical aspect of this experiment is the simulation's end. Consequently, the charts are more interesting to investigate.

Estimating N_{min} results in more precise forecasts, particularly for area and nitrogen, but not for carbon. The same enhancement in the estimation can be observed when estimating J_{max}. This aligns with the findings obtained from the SA. Despite N_{min} and J_{max} not being the parameters that nitrogen is most sensitive to, it still plays crucial roles in predicting the nitrogen content in the kelp [37]. N_{min} is the parameter the area is the most sensitive to, which is reflected in both Table 4.2 and Figure 4.6. Conversely, according to the SA, carbon is not sensitive to J_{max}, but an enhancement in this state estimate is observed when the estimation includes this parameter. This suggests that J_{max} has a more significant role in predicting the carbon content than initially anticipated.

Estimating k_A yields better predictions for area, nitrogen, and carbon. α estimations do not improve area and nitrogen prediction but significantly enhance carbon prediction. The same effect is observed for I_{sat}. The results for k_A are consistent with those from the SA. However, a heightened response in nitrogen was anticipated when incorporating estimates of I_{sat} and α, considering these are among the parameters to which nitrogen is most responsive [37]. Nevertheless, this response was absent. This may indicate that the model state of nitrogen is not that sensitive to this parameter after all. It is also worth noting that the SA is conducted on the nitrogen and carbon content of the kelp, whereas this thesis solely focuses on the kelp's carbon and nitrogen. As shown in section 2.3, the computation of nitrogen and carbon content contains additional factors beyond the mere computation of solely nitrogen and carbon. Thus, using carbon and nitrogen content in this study might have yielded a different result.

As mentioned earlier for k_C, estimating this parameter does not yield better estimates for area, nitrogen, and carbon than scenarios without parameter estimation.

When the difference between the models, with and without parameter estimations, is minimal, it is challenging to confirm whether this outcome is random—suggesting that in subsequent simulations, the model with parameter estimation could produce a higher mean RMSE than the one without—or whether this outcome is accurate. It would have been beneficial to run multiple simulations to assess the randomness of the results.

5.2 Estimation using field data

In Figure 4.7, it is evident that an estimate consistent with the measurements can be obtained given a sufficiently large ensemble size and a standard deviation of 0.5 for the measurement error of the frond area and 0.2 for the area, 0.0001 for nitrogen, and 0.005 for carbon for the model error. However, due to the absence of corresponding measures, limited conclusions can be drawn for nitrogen and carbon. Despite this, some correction steps reveal updates in nitrogen and carbon. This suggests a correlation between nitrogen, carbon, and the area, a conclusion that is further supported in Figure 4.8. It is noticeable that the area and nitrogen primarily exhibit a positive correlation that fluctuates over time, with varying degrees of correlation ranging from non-existent to weak and even strong. Conversely, the area and carbon mainly display a negative correlation, which is not as pronounced as between the area and nitrogen.

This update comes from the Kalman gain. When assimilating observations, the EnKF computes the Kalman gain, which is used to update the state estimates. The Kalman gain considers the covariances between state variables and the uncertainties in both the model predictions and the measurements. As a result, when the filter updates one variable based on new observations, it also adjusts other correlated variables, even if there are no direct measurements for those variables.

As depicted in Figure 4.9, the selection of the configuration values plays a crucial role. An ensemble size of 200 has been chosen based on several considerations. Firstly, a larger ensemble enables a more comprehensive representation of the uncertainties inherent in the system, thereby leading to more accurate state and parameter estimates. The increased ensemble size enhances the filter's ability to contain the full spectrum of possible outcomes and provides a more reliable estimation. Moreover, it enhances the robustness of the filter, as the large ensemble size helps mitigate the impact of stochastic errors that may arise from perturbations throughout the simulation.

Another crucial adjustable parameter is the standard deviation of the measure-

ment error. If the measurement error is underestimated, the filter may rely too heavily on the observations, leading to overfitting and poor performance in the presence of noisy data. Conversely, if the measurement error is overestimated, the filter may not effectively assimilate the observations, resulting in less accurate state and parameter estimates.

The standard deviation of the model error is also a critical configurable value. If the model error is overestimated, it can make the filter put too much weight on the measurements, resulting in overly responsive noisy estimates. However, an underestimated value can cause the filter to put too much weight on the model prediction, making it slow in response to changes in the measurements. There is a trade-off between these two errors that is crucial when estimating the states. Numerous model and measurement error combinations were explored to ascertain an optimal balance. It was decided to assign a slightly smaller value to the model error than the measurement error, hence moderately favoring the estimates over the measurements. This approach was considered appropriate due to the anticipated high degree of uncertainty in the measurements and also in the model.

As displayed in Figure 4.9, the absolute error throughout the ensemble increases with the standard deviation of the measurement error. Interestingly, the heat map also suggests that larger measurement errors yield the most precise estimates towards the end of the simulation, possibly hinting at a flaw in the mathematical growth model. Despite achieving accuracy in the latter stages of the simulation, the average RMSE across the entire simulation remains notably high for the highest standard deviation values. Nevertheless, the most commonly used measurement error had a standard deviation of 0.5 dm^2.

Figure 4.10 demonstrates that the EnKF generally offers superior estimates of the frond area compared to the mathematical model alone. However, toward the end of the simulation, the mathematical model delivers a more accurate forecast. This may also expose a potential weakness in the mathematical growth model. Ideally, for the predictions made by the EnKF, it is preferable that the prediction of the frond area levels off once the area becomes sufficiently large, a feature currently lacking in this model. This observation also implies that incorporating additional measurements could be beneficial due to the non-linear growth of kelp. If the objective is to maximize biomass yield, the conclusion of the simulation and growth period becomes even more crucial. As the frond area increases, frond erosion also increases, making it advantageous to harvest the kelp before erosion becomes excessive.

The current frequency of area measurements near the end aligns with the bi-weekly recommendation from the twin experiment. However, this does not yield adequate estimates, suggesting a revised measurement strategy is needed. This could involve increasing the frequency of measurements early in the process to

follow the recommendations of measurements every other week or making measurements more frequently towards the end. Alternatively, it could be insightful to test whether the existing measurement strategy produces similar results when applied to a different dataset.

5.2.1 Parameter estimation

In Figure 4.12, it is evident that when only frond area measurements are available, it is possible to estimate N_{min}, J_{max}, and k_A. This finding aligns with the sensitivity analysis results, which identified the growth model's frond area as being most sensitive to N_{min}, N_{max}, k_A, and J_{max} [37]. In this case, the inability to estimate α, I_{sat}, and k_C can be concluded due to the divergence in their estimation. It is clear from the figure that in each correction step, the parameter value are only perturbed, not corrected. The noticeable variation in N_{min}, k_A, and J_{max} at each correction step suggests that avoiding constant parameter values and including variability may enhance the Ensemble Kalman Filter's performance.

As mentioned in subsection 2.3.10, there is considerable variation in specific parameter values found in the literature, particularly for J_{max} and k_A. In light of this, combined with the outcomes from the parameter estimation, it is possible that the current parameter values used in the growth model may not be optimal. Consequently, employing alternative values might enhance the accuracy of both the model and its predictions.

As observed in Figure 4.11, the most accurate frond area estimate is achieved when N_{min} is included in the state vector. Improved estimates can also be noticed when estimating J_{max} and k_A, but this is primarily limited to the initial part of the simulation. As discussed in section 5.1, this may suggest that having more frequent measurements towards the end of the growth period is advantageous.

When estimating α, I_{sat}, and k_C, the resulting estimates are roughly as accurate as solely estimating the states. This is reasonable, given that these parameters are crucial for carbon and nitrogen. As no available measurements exist for these states, the effect of estimating these parameters is not present.

5.3 Limitations

One limitation of these results is that this model's current was set to a constant of 0.06 ms^{-1}, which is not representative in a real system. Therefore, a time-varying current could affect the model's predictions and estimates.

Another area for improvement is the accuracy of the measurements. In this thesis, only observations for the area were available, but this is hard to measure and is expected to contain much uncertainty.

Another limitation is the need for carbon and nitrogen measurements, which limits the conclusions made in the thesis. It would be advantageous also to have measurements for carbon and nitrogen to investigate if some of the conclusions drawn from the twin experiment would be transferable.

In this thesis, nitrogen and carbon have been investigated. However, it would have been better to study the carbon and nitrogen content of the kelp instead since these are measurable units and the units upon which the sensitivity analysis was conducted.

Another limiting factor is the statistical foundation. Even though a sufficiently large ensemble size should theoretically eliminate any stochastic errors, this may not always be the case, and the result might represent stochastic values rather than the actual model.

Lastly, there might be an issue with the configurational values. They might not have been selected optimally, thereby failing to present an accurate representation of the ensemble Kalman filter.

5.4 Further work

Based on the results and limitations in this thesis, future points on research and areas of improvement are suggested.

First, a measurement program for cultivated kelp should be devised, informed by the recommendations gleaned from this study on sampling strategies.

One of the other vital areas for future work is to use the results from this growth model with an ensemble Kalman filter to estimate the states and parameters of real-time kelp growth. As mentioned in section 1.1, a considerable motivation is to help kelp farmers achieve the largest biomass. It would be interesting to investigate if this growth model possibly is a tool that could be used for this application.

A further intriguing prospect for future evaluation is obtaining measurements of carbon and nitrogen content. This could verify if some of the inferences derived from the twin experiment could be realized in real-life models. Additionally, collecting data from geographical locations beyond Trøndelag would be beneficial to ascertain if the conclusions are transferable to other areas.

Subsequent efforts might consider enabling the estimation of multiple parameters concurrently and examining additional model parameters, in contrast to how it is implemented in this thesis. It could also be beneficial to explore whether there are alternative methods to EnKF that may be more proficient in estimating the states and parameters of this model.

It may also prove helpful to incorporate current measurements and examine if the growth model with the EnKF can be applicable beyond the temporal scope addressed in this thesis.

The findings of this project are specific to this context, yet the approach employed for the growth of kelp is general. It could potentially apply to species beyond *Saccharina latissima*. Further exploration into the feasibility of this application would be fascinating.

With time, it may become possible to estimate kelp's area, nitrogen, and carbon content optically. This method could offer more regular measurements of the kelp as it is less resource-intensive. Consequently, it could lead to more precise estimates, enabled by a more frequent measurement strategy.

6

Conclusion

This study integrated an ensemble Kalman filter for state and parameter estimation in a growth model for cultivated *Saccharina latissima*. A sensitivity analysis was used to select the model parameters for estimation, and the implementation's effectiveness was evaluated through a twin experiment and real-world kelp experiment data.

The findings from this thesis suggest that the ensemble Kalman filter can be effectively employed for estimating the states of growth - area, nitrogen, and carbon - as well as the key parameters N_{min}, J_{max}, k_A, α, and I_{sat}. The twin experiment, in particular, demonstrated that estimating these parameters resulted in more accurate predictions. Furthermore, the findings suggest an optimal measurement strategy involves measuring all three output states bi-weekly, with allowances for minor schedule inconsistencies.

Practically, by employing measures focused exclusively on the area from actual kelp experiments, the growth estimate accuracy can be enhanced by including parameter estimation for N_{min}, J_{max}, and k_A. Despite these promising results, a significant decline in the precision of estimates was observed towards the end of the growth period, indicating a potential system limitation. This limitation can be restrained with more frequent measurements towards the end of the growth period.

In conclusion, it would be favorable to incorporate field data on nitrogen and carbon content to ensure accurate state estimates and validate the conclusions drawn from the twin experiment. This could be vital in optimizing biomass production and identifying ideal locations for new kelp farms. Further, it is recommended to establish a measurement program for cultivated kelp, based on the recommendations presented in this thesis, and to apply the filter on real-time kelp growth to optimize harvest timing.

Bibliography

[1] Andersen, G.S., 2021. *Tang og tare (brunalger)*. Available from: https://snl.no/tang_og_tare_-_brunalger (Accessed: February 26th 2023).

[2] Bakken, A., 2021. *1,4 milliarder kroner til laboratorier, utstyr og annen forskningsinfrastruktur*. Regjeringen.no. Available from: https://www.regjeringen.no/no/aktuelt/penger/id2891691/ (Accessed: May 16th 2023).

[3] Bartsch, I., Wiencke, C., Bischof, K., Buchholz, C.M., Buck, B.H., Eggert, A., Feuerpfeil, P., Hanelt, D., Jacobsen, S., Karez, R., Karsten, U., Molis, M., Roleda, M.Y., Schubert, H., Schumann, R., Valentin, K., Weinberger, F., Wiese, J., 2008. *The genus Laminaria sensu lato : recent insights and developments*. European Journal of Phycology, 43(1), p. 1-86. doi:https://doi.org/10.1002/aqc.3270010208.

[4] Benjaminsen, C., 2022. *Denne plasten er laget av tare*. Gemini. Available from: https://gemini.no/2022/09/denne-plasten-er-laget-av-tare/ (Accessed: May 4th 2023).

[5] Bermejo, P., Durán-Romero, C., Villafañe, V.E., Helbling, E.W., 2020. Influence of fluctuating irradiance on photosynthesis, growth and community structure of estuarine phytoplankton under increased nutrients and acidification. *Journal of Experimental Marine Biology and Ecology*, 526, p. 151-348. doi:10.1016/j.jembe.2020.151348.

[6] Broch, O., Alver, M., Bekkby, T., Gundersen, H., Forbord, S., Handå, A., Skjermo, J., Hancke, K., 2018. Kelp cultivation potential in coastal and

offshore regions of norway. *Frontiers in Marine Science* doi:10.3389/fmars.2018.00529.

[7] Broch, O., Skjermo, J., Handå, A., 2016. The potential for macroalgal cultivation in møre and romsdal (in norwegian, english summary) doi:10.13140/RG.2.2.21137.02406.

[8] Broch, O.J., Skjermo, J., Ellingsen, I., Knutsen, , Jiang, L., Soetaert, K., Jansen, H., Neves, L., Borrero, A., Fariñas-Franco, J.M., Fort, A., Sulpice, R., 2021. *Growth, nutrient uptake and carrying capacity of kelp (Saccharina latissima) cultures in Europe*. SINTEF Ocean. Available from: https://genialgproject.eu/results/deliverables/. (Accessed: February 15th 2023).

[9] Broch, O.J., Slagstad, D., 2012. Modelling seasonal growth and composition of the kelp saccharina latissima. *Journal of Applied Phycology*, 24, p. 759-776 doi:10.1007/s10811-011-9695-y.

[10] Brown, R.G., Hwang, P.Y., 2012. *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises*. Industrial Network Security (4th Edition). John Wiley Sons Inc, p. 77-81, 128-130, 262-285.

[11] Buschmann, A., Camus, C., 2019. An introduction to farming and biomass utilisation of marine macroalgae, *Phycologia*, 58, p. 443 doi:10.1080/00318884.2019.1638149.

[12] Cai, J., 2021. *Global status of seaweed production, trade and utilization*. Food and Agriculture Organization of the United Nations. Available from: https://www.fao.org/ (Accessed: February 15th 2023).

[13] Chen, J., 2004. *Cultured Aquatic Species Information Programme. Laminaria japonica.* FAO Fisheries and Aquaculture Division. Available from: http://firms.fao.org/fi/website/FIRetrieveAction.do?dom=culturespecies&xml=Laminaria_japonica.xml (Accessed: March 8th 2023).

[14] Danielsson, S.A., 2022. *Starter karbonfangst med sukkertare*. DNV. Available from: https://www.dnv.no/news/starter-karbonfangst-med-sukkertare-228139 (Accessed: April 16th 2023).

[15] Evensen, G., 2006. Data assimilation. the ensemble kalman filter. (2nd Edition), p. 136, 157, 178-196 doi:10.1007/978-3-540-38301-7.

[16] Fiskeridirektoratet, 2022. *Akvakulturstatistikk: alger*. Available from: https://www.fiskeridir.no/Akvakultur/Tall-og-analyse/Akvakulturstatistikk-tidsserier/Alger (Accessed: April 10th 2023).

[17] Flavin, K., Flavin, N., Flahive, B., 2013. *Kelp Farming Manual a Guide to the Processes, Techniques, and Equipment for Farming Kelp in New England Waters*. Available from: https://maineaqua.org/wp-content/uploads/2020/06/OceanApproved_KelpManualLowRez.pdf (Accessed: Februrary 7th 2023).

[18] FollaAlger, 2022. *RI SEAWEED*. Available from: https://www.follaalger.no/aktuelt/ri-seaweed (Accessed: May 27th 2023).

[19] Forbord, S., 2020. *Cultivation of Saccharina latissima (Phaeophyceae) in temperate marine waters*. Thesis for the Degree of Philosophiae Doctor, Norwegian University of Science and Technology, Faculty of Natural Sciences, Department of Biology. Available from: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2683614 (Accessed: May 22nd 2023).

[20] Forbord, S., Matsson, S., Brodahl, G., Bluhm, B., Broch, O., Handå, A., Metaxas, A., Skjermo, J., Steinhovden, K., Olsen, Y., 2020. Latitudinal, seasonal and depth-dependent variation in growth, chemical composition and biofouling of cultivated Saccharina latissima (Phaeophyceae) along the Norwegian coast, *Journal of Applied Phycology*, 32. doi:10.1007/s10811-020-02038-y.

[21] Gulden, A.S., 2022. *Sukkertare — Saccharina latissima*. Norsk Fiskerinæring. Available from: https://norskfisk.no/2022/05/25/sukkertare-saccharina-latissima/?r (Accessed: February 26th 2023).

[22] Halvorsen, D.Ø., Dallolio, A., Alver, M.O., 2022. Assimilation of heterogeneous measurements at different spatial scales in the Arctic ocean and Norwegian sea. doi:10.1109/OCEANS47191.2022.9977376.

[23] Houtkamer, P.L., Zhang, F., 2016. Review of the ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144(12), p. 4489-4532 doi:10.1175/MWR-D-15-0440.1.

[24] Juanich, G.L., 1988. *Manual on seaweed farming eucheuma spp.* ASEAN/UNDP/FAO Regional Small-Scale Coastal Fisheries Development Project. Available from: https://www.fao.org/3/ac416e/ac416e00.htm (Accessed: May 10th 2023).

[25] Kandepu, R., Foss, B., Imsland, L., 2008. Applying the unscented Kalman filter for nonlinear state estimation. doi:https://doi.org/10.1016/j.jprocont.2007.11.004.

[26] Katzfuss, M., Stroud, J.R., Wikle, C.K., 2016. Understanding the ensemble kalman filter. *The American Statistician Taylor and Francis*, 70(4), p. 350–357. doi:10.1080/00031305.2016.1141709.

[27] Knutsen, Ø., . *SINMOD*. SINTEF. Available from: https://www.sintef.no/sintef-ocean/satsinger/sinmod/ (Accessed: May 31st 2023).

[28] Luning, K., 1979. *Growth Strategies of Three Laminaria Species (Phaeophyceae) Inhabiting Different Depth Zones in the Sublittoral Region of Helgoland (North Sea).* Marine Ecology Progress Series, 1, p. 195-207.

[29] Luning, K., 1991. Seaweeds: their environment, biogeography, and ecophysiology. *Wiley Intersience*, p. 189. doi:10.1002/aqc.3270010208.

[30] MathWorks, . *Correlation coefficients*. Available from: https://se.mathworks.com/help/matlab/ref/corrcoef.html#buty8g7 (Accessed: May 25th 2023).

[31] Moody, J., 2019. *What does RMSE really mean?* Towards Data Science. Available from: https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e (Accessed: May 8th 2023).

[32] Moy, F., Alve, E., Bogen, J., Christie, H., Green, N., Helland, A., Steen, H., Skarbøvik, E., Stålnacke, P., 2006. *Sukkertareprosjektet statusrapport nr. 1*. Norsk institutt for vannforskning. Available from: https://niva.brage.unit.no/niva-xmlui/bitstream/handle/11250/213313/5265-2006_200dpi.pdf?sequence=2&isAllowed=y (Accessed: May 10th 2023).

[33] Müller, D., Gassmann, G., Lüning, K., 1979. *Isolation of a spermatozoid-releasing and -attracting substance from female gametophytes of Laminaria digitata*, 279, p. 430–431.

[34] NIVA, 2019. *Researcher believes in kelp farmings*. Norwegian Institute for Water Research. Available from: https://www.niva.no/en/news/researcher-believes-in-kelp-farming (Accessed: February 7th 2023).

[35] Norderhaug, K.M., Skjermo, J., Kolstad, K., Broch, O.J., Ergon, Å., Handå, A., Horn, S.J., Lock, E.J., Øverland, M., 2020. *Toward a new marine industry for kelp?* Havforskningsinstituttet. 1.3 Available from: https://www.hi.no/templates/reporteditor/report-pdf?id=37885&87464199. (Accessed: February 20th 2023).

[36] North, G.R., Pyle, J., Zhang, F., 2015. Numerical models — parameter estimation. *Encyclopedia of Atmospheric Sciences (Second Edition)*, p. 181-186 doi:10.1016/B978-0-12-382225-3.00494-1.

[37] Rygg, H.O., 2022. *Sensitivity analysis of growth model of the farmed kelp Saccharina latissima*. Norwegian University of Science and Technology.

[38] SES, 2009. *Saccharina latissima (Sugar Kelp / Kombu)*. Seaweed Solutions. Available from: https://seaweedsolutions.com/sugar-kelp-kombu-saccharina-latissima (Accessed: February 9th 2023).

[39] SINTEF, 2021. *Norwegian Seaweed Technology Center*. SINTEF Ocean. Available from: https://www.sintef.no/en/ocean/initiatives/norwegian-seaweed-technology-center/ (Accessed: February 15th 2023).

[40] Sjøtun, K., 1993. Seasonal Lamina Growth in two Age Groups of Laminaria saccharina (L.) Lamour. in Western Norway. *Botanica Marina*, 36, p. 433-441. doi:10.1515/botm.1993.36.5.433.

[41] Skjermo, J., . *GP Seaweed*. SINTEF. Available from: https://www.sintef.no/prosjekter/2023/gp-seaweed/ (Accessed: May 12th 2023).

[42] Skjermo, J., Aasen, I.M., Arff, J., Broch, O.J., Carvajal, A.K., Christie, H.C., Forbord, S., Olsen, Y., Reitan, K.I., Rustad, T., Sandquist, J., Solbakken, R., Steinhovden, K., Wittgens, B., Wolff, R., Handå, A., 2014. *A new Norwegian bioeconomy based on cultivation and processing of seaweeds: Opportunities and RD needs* , SINTEF Fisheries and Aquaculture. Available from: https://sintef.brage.unit.no/sintef-xmlui/handle/11250/2448125 (Accessed: May 22nd 2023).

[43] Stévant, P., Rebours, C., Chapman, A., 2017. Seaweed aquaculture in norway: recent industrial developments and future perspectives, *Aquaculture International*, 25. doi:10.1007/s10499-017-0120-7.

[44] Venolia, C., 2018. *KELP MEASUREMENTS: COLLECTING, DRYING AND GRINDING SPECIMENS*. University of Rhode Island. Available from: http://ahumphrieslab.com/kelp-measurements. (Accessed: May 22nd 2023.

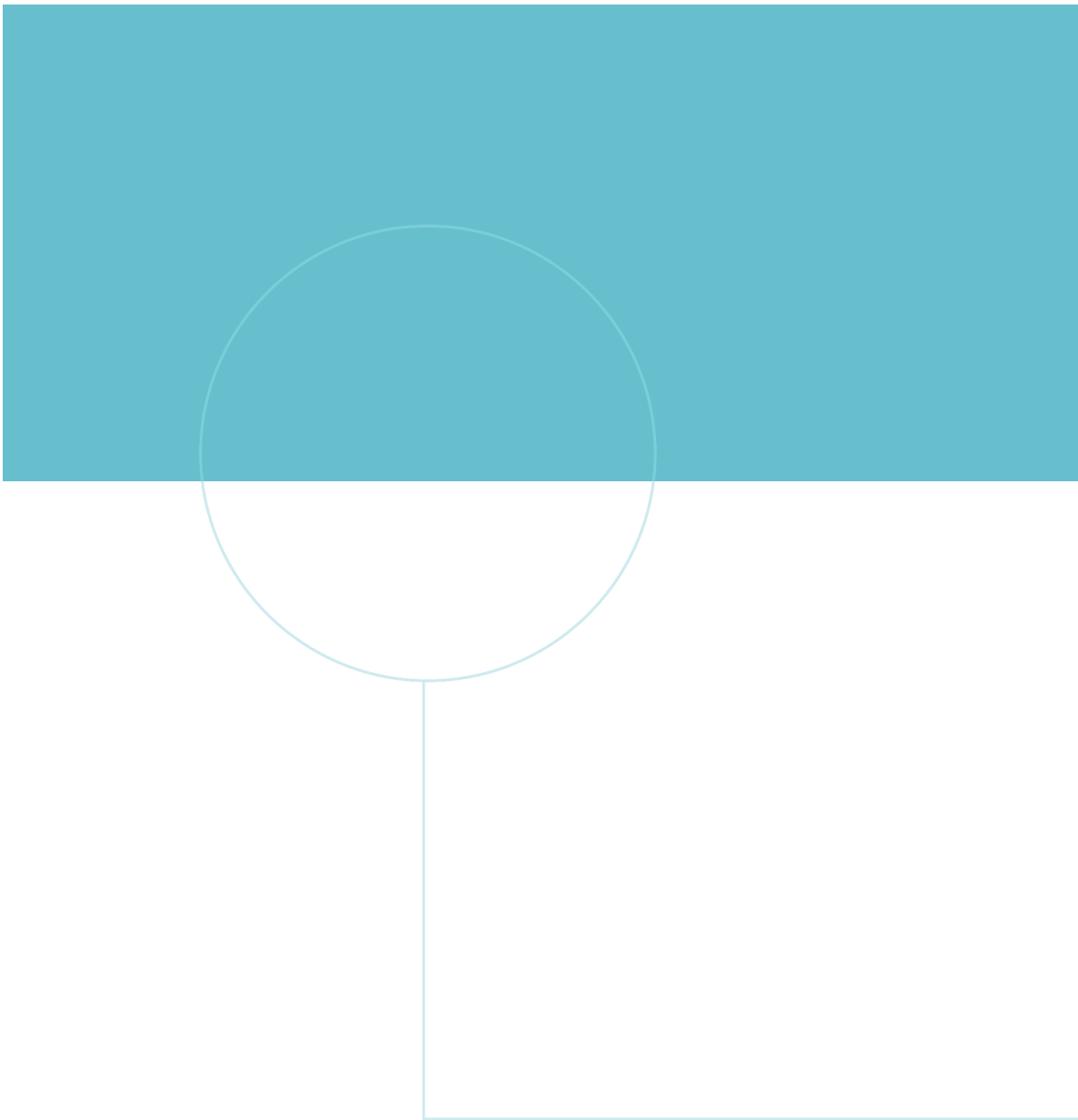
[45] Visch, W., Rad-Menéndez, C., Nylund, G.M., Pavia, H., 2019. Underpinning the Development of Seaweed Biotechnology: Cryopreservation of Brown Algae (Saccharina latissima) Gametophytes. *Biopreservation and Biobanking*. doi:10.1089/bio.2018.0147.

[46] White, N., Marchall, C., 2007. *Sugar kelp (Saccharina latissima)*. Frontiers in Marine Science. Available from: https://www.marlin.ac.uk/species/detail/1375 (Accessed: February 7th 2023).

[47] Witman, J.D., 1987. *Subtidal Coexistence: Storms, Grazing, Mutualism, and the Zonation of Kelps and Mussels*. Ecological Monographs, 57(2), p.167-187. Available from: http://www.jstor.org/stable/1942623 (Accessed: March 5th 2023).

[48] Zhang, Y.F., Bitz, C., Anderson, J., Collins, N., Hendricks, J., Hoar, T., Raeder, K., Massonnet, F., 2018. Insights on sea ice data assimilation from perfect model observing system simulation experiments. *Journal of Climate*, 31, p. 5911-5926. doi:10.1175/JCLI-D-17-0904.1.

[49] Zhang, Z., Moore, J.C., 2015. Chapter 9 - Data Assimilation. *Mathematical and Physical Fundamentals of Climate Change*, p. 291-311. doi:10.1016/B978-0-12-800066-3.00009-7.



NTNU
Norwegian University of
Science and Technology

SINTEF