KEPLER Deliverable Report

Report on Deliverable D3.4

Deliverable name	Report on research gaps in terms of integration/assimilation of space-based and in-situ observations to fill in order to improve Arctic monitoring and forecasting capabilities.			
Scheduled delivery	**month:**	18	**date:**	June 2020
Actual delivery	**month:**	20	**date:**	August 2020
Report type	Public report			
Lead author	Thomas Kaminski, ILAB			

Contributing authors

F. Kauker (OASYS), M. Scholze (ULUND), L. Toudal-Pedersen (EOLAB), M. Voßbeck (ILAB), T. Lavergne (METNO), L. Bertino (NERSC), G. Garric (MERCATOR), S. Tietsche (ECMWF), and H. Chen (ULUND)

Context of deliverable within Work Package

This report evaluates several observational scenarios in terms of their performance in a data assimilation system. In the construction of these observational scenarios we put emphasis on the Copernicus Sentinel satellites with particular focus on the High Priority Candidate Missions (HPCMs) for expansion of the Sentinel fleet. One group of scenarios consists of observations of the Arctic sea ice-ocean system, while another group consists of observations of atmospheric CO_2.

Explanation of delays

Although the report was ready for the internal review in time the delivery was delayed due to a slow internal review process caused by Corona.

Co-funded by the Horizon 2020 programme
of the European Union

List of Acronyms

AET	Actual Evapotranspiration
AVISO	Archiving, Validation and Interpretation of Satellite Oceanographic data
BETHY	Biosphere Energy-Transfer Hydrology
BFB	Baffin Bay
BFS	Beaufort Sea
C3S	Copernicus Climate Change Service
CARMA	Carbon Monitoring for Action
CCDAS	Carbon Cycle Data Assimilation System
CCFFDAS	Carbon Cycle Fossil Fuel Data Assimilation System
CDR	Climate Data Record
CIMR	Copernicus Imaging Microwave Radiometer
CMCC	Euro-Mediterranean Center on Climate Change
CMEMS	Copernicus Marine Environment Monitoring Service
CO2M	Copernicus Anthropogenic Carbon Dioxide Monitoring
CRISTAL	Copernicus Polar Ice and Snow Topography Altimeter
CRU	Climatic Research Unit
DWD	Deutscher Wetterdienst (German Weather Service)
EASE	Equal-Area Scalable Earth
ECMWF	European Centre for Medium-Range Weather Forecasts
EO	Earth Observation
EPF	Error Parameterisation Formula
ESA	European Space Agency
ESS	East Siberian Sea

List of Acronyms

FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FFDAS	Fossil Fuel Data Assimilation System
GOSAT	Greenhouse gases Observing Satellite
HPCM	High Priority Candidate Mission
IEA	International Energy Agency
IR	infrared
IPCC	Intergovernmental Panel on Climate Change
LFB	Laser Freeboard
MODIS	Moderate Resolution Imaging Spectroradiometer
MPIOM	Max-Planck-Institute Ocean Model
NPP	Net Primary Productivity
NCEP	National Centers for Environmental Prediction
NSIDC	National Snow and Ice Data Center
NSR	Northwestern Sea Route
ONSI	Outer New Siberian Islands
OSI-SAF	Ocean and Sea Ice Satellite Application Facility
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis
PDF	Probability Density Function
PET	Potential Evapotranspiration
PFT	Plant Functional Type
QND	Quantitative Network Design
RFB	Radar Freeboard
SIC	Sea Ice Concentration
SIFB	Sea Ice Freeboard
SIT	Sea Ice Thickness
SIV	Sea Ice Volume

List of Acronyms

SMOS	Soil Moisture Ocean Salinity
SND	Snow Depth
SNV	Snow Volume
SST	Sea surface temperature
JRC-TIP	JRC Twostream Inversion Package
WLS	West Laptev Sea
XCO2	Column Integrated Atmospheric CO_2

Summary

- This report evaluates several observational scenarios in terms of their performance in a data assimilation system. In the construction of these observational scenarios we emphasise the Copernicus Sentinel satellites with particular focus on the High Priority Candidate Missions (HPCMs) for the expansion of the Sentinel fleet. One group of scenarios consists of observations of the Arctic sea ice-ocean system, while another group consists of observations of atmospheric CO_2.

- We employ the quantitative network design (QND) approach to assess the impact of these scenarios in a mathematically rigorous fashion through the reduction of uncertainties in a set of relevant target quantities. For the sea ice-ocean observations, our target quantities are 1-week to 4-week forecasts of sea ice volume (SIV) and snow volume (SNV) for selected regions along the Northern Sea Route and the Northwest Passage as well as for the entire Arctic. Our assessments assume observations are assimilated in April 2015, with the respective 1-week and 4-week forecasting periods starting on May 1. For the atmospheric CO_2 observations, our target quantities are the land-based fossil fuel emissions in the first week of June from several Arctic countries, namely Canada, Denmark, Finland, Island, Norway, and Sweden.

- As our focus is on the observational scenarios and model error is an aspect that is specific to our model and difficult to specify, we deliberately exclude it from the assessments (together with biases in the observations). For example, we assume that the penetration of radar and laser signals is treated correctly. The exclusion of model error emphasises the differences between observational scenarios and shows the upper limit for the impact that each scenario can achieve. We do, however, complement our assessments with estimates of the impact of model error on the simulated target quantities. As our reference is a prior case without any observations, the impact of a given observational data stream is also much higher than in a setup where it is added to an assimilation system that already assimilates a variety of other data sets, as it is the case for the Copernicus systems.

- Sentinel 3 (S3) radar freeboard (RFB) outperforms CryoSat-2 (CS-2) RFB in the selected target regions relevant for marine transportation in the Arctic because of the higher temporal coverage. The larger pole hole of S3 is irrelevant. While this is trivial for the selected target regions relevant for shipping (too far away), S3 outperforms CS-2 as well for the Arctic-wide assessment.

- When combined with CS-2 RFB, CIMR-like and CRISTAL-like snow depth (SND) products yield strong gain in forecast performance. The same holds for a typical product based on an atmospheric reanalysis. Although the differences

for these assessments are small, for the respective accuracies we have assumed CIMR shows the best performance among the three products.

- The combination of CS-2 RFB and ICESat2-like laser freeboard (LFB) shows the overall best performance for both, SIV and SNV. This is because the assumed accuracy of the LFB (2 cm) was higher than the accuracy of the SND products. Furthermore, assimilation of the raw freeboard product is more beneficial than the assimilation of a derived product.

- The performance of a CIMR-like SST product is better than that of a traditional infrared-based SST product. Although the infrared product is more accurate, the better spatial coverage (owing to its capability to penetrate clouds) renders CIMR attractive for predicting SIV and SNV along the shipping routes. We note that only CIMR and only over the target region in the Baffin Bay shows a strong impact, because for the other regions and the infrared the SST observations are too far away. The impact of SST is expected to be higher for Summer conditions when most of the target regions are at least partially ice-free.

- Assessments of several observational scenarios for atmospheric CO_2 in terms of their constraint on land-based fossil fuel CO_2 emissions in June show that an increase in the number of sites of a small surface network providing continuous in situ samples is more efficient than the reduction of its observational uncertainty or the addition of radiocarbon measurements. Further, combining the small surface network with a single CO2M satellite already provides a better constraint on land-based fossil fuel CO_2 emissions in June than increasing the number of continuous sampling sites.

- The provision of spatial and temporal uncertainty correlations with the EO products would be beneficial not only for QND assessments, but also for the assimilation of the products.

- The QND approach is ideally suited to assist the formulation of mission requirements or the development of EO products. In an end-to-end simulation it can translate product specifications in terms of spatio-temporal resolution and coverage, accuracy, and precision into a range of performance metrics. Alternatively, it can translate requirements on forecast performance into requirements on the respective observables. It can assess combinations of in situ and EO data (from multiple missions). This type of assessment can be performed for higher-level products (e.g. SIT or SIC) but also for rawer products (e.g. freeboard or brightness temperature).

Contents

List of Figures

List of Tables

1 Observational Scenarios

Forecasts of the sea ice and the ocean state are routinely performed by coupled sea ice-ocean models that are driven by prescribed atmospheric conditions. To derive reliable forecasts, uncertainties in the models' initial state, the atmospheric boundary conditions, and in the parameterisations of physical processes need to be minimised. Observations can help to reduce such uncertainties and, thus, to improve the forecast quality.

For example, the Copernicus Arctic Marine Forecasting Centre (ArcMFC) is currently running its real-time forecasting system at a resolution of 12 km regionally and assimilates

- in situ data T and S (Argo, ITPs) in situ TAC (thematic assembly centre),
- SST (OSTIA),
- along-track sea level anomaly (AVISO/CMEMS),
- SIC (OSI-SAF),
- SIT (SMOS), and
- SI drift (OSI-SAF)

over an assimilation window of 7 days to produce forecasts with a lead time of 10 days.

The Copernicus Global Marine Forecasting Centre (GloMFC) system is currently running at Mercator at a resolution of 8-9 km (1/12 degree) and assimilates

- in situ data T and S (Argo),
- SST (OSTIA),
- along-track sea level anomaly (AVISO/CMEMS), and
- SIC (OSI-SAF)

over an assimilation window of 7 days to produce forecasts with a lead time of 10 days.

The forecasts that are currently running at ECMWF use resolutions of 9-36 km of the atmospheric component and 0.25° of the ocean component and are based on the assimilation of the following oceanic variables:

- in situ profiles of T and S,
- SST (OSTIA),
- along-track sea level anomaly (AVISO/CMEMS), and

- SIC (OSTIA)

over an assimilation window of 5 days to produce forecasts with lead times from 10 days to 13 months. Within the Copernicus Climate Change Service (C3S) this is complemented with seasonal forecasts by the systems from Météo-France, the UK Met Office, DWD, CMCC, and NCEP.

This report evaluates several observational scenarios in terms of their performance in a data assimilation system. In the construction of these observational scenarios we emphasise the Copernicus Sentinel satellites with particular focus on the High Priority Candidate Missions (HPCMs) for expansion of the Sentinel fleet. Our first two observational scenarios address altimetric measurements from the CryoSat-2 (CS-2) and the Sentinel 3 (S3) missions. We focus on radar freeboard (RFB) products, because a preceding study (*Kaminski et al.*, 2018) had identified RFB as the data product with the highest impact compared to Sea Ice Thickness (SIT) or Sea Ice Freeboard (SIFB) products.

The next set of scenarios also builds upon a finding of *Kaminski et al.* (2018) which had indicated good complementarity of RFB with snow depth (SND) products. We hence constructed three hypothetical SND products and evaluated them in combination with the CS-2 RFB product. The first SND product is intended to look like an SND product to be expected from the Copernicus HPCM CRISTAL (Copernicus Polar Ice and Snow Topography Altimeter). The second SND product is intended to look like an SND product to be expected from the Copernicus HPCM CIMR (Copernicus Imaging Microwave Radiometer). The third SND product is intended to look like an SND product to be expected from a modelling approach that combines with the help of a numerical snow model an ice drift product with output from a numerical weather prediction model to calculate the temporal evolution of SND. A CRISTAL SND product would be derived from the difference of two freeboard measurements onboard the same platform (Ku-Ka band). An alternative is the direct assimilation of two freeboard products into the model. For this purpose we constructed an LFB product that looks like an ICESat-2 product, which we also evaluated in combination with CS-2.

A further observable we evaluate is SST. Our traditional scenario is based on the product retrieved by OSI-SAF from infrared (IR) measurements, and is contrasted with a scenario that uses an SST product to be expected from the HPCM CIMR.

For land, our observational scenarios focus on the detection of land-based fossil fuel emissions from Arctic countries with atmospheric CO_2 observations. We first assess a small surface network of 15 sites providing continuous in situ samples, and then four scenarios that extend/improve this network. The first scenario equips many more sites that are currently providing flask samples with continuous in situ analysers. The second scenario adds radiocarbon observations to the 15-site network. The third scenario employs the 15-site network with reduced observational uncertainty. The fourth scenario combines the 15-site network with a single satellite of the planned HPCM CO2M (Copernicus Anthropogenic Carbon Dioxide Monitoring) Mission.

2 Approach

2.1 Introduction

The quantitative network design (QND) approach allows for an objective evaluation of the added value of observations for a given aspect of a model simulation or forecast. The technique originates from seismology (*Hardt and Scherbaum*, 1994) and was first applied to the climate system by *Rayner et al.* (1996), who optimised the spatial distribution of in situ observations of atmospheric carbon dioxide to achieve minimum uncertainty in inferred surface fluxes. After an initial QND study that demonstrated the feasibility of the approach for remote sensing of the column-integrated atmospheric carbon dioxide concentration (*Rayner and O'Brien*, 2001) QND is now routinely applied in the design of CO_2 space missions (e.g., *Patra et al.*, 2003; *Houweling et al.*, 2004; *Crisp et al.*, 2004; *Feng et al.*, 2009; *Kadygrov et al.*, 2009; *Hungershoefer et al.*, 2010; *Rayner et al.*, 2014; *Bovensmann et al.*, 2015).

For the western Arctic domain, the QND approach has been successfully applied to evaluate the impact of (hypothetical) airborne measurements of SIT/SND in improving sea ice predictions (*Kaminski et al.*, 2015). The study evaluated two idealised flight transects derived from NASA's Operation IceBridge airborne altimeter ice surveys in terms of their potential to improve ten-day to five-month forecasts of sea ice conditions, including for operational purposes. The study by *Kaminski et al.* (2018) presents QND assessments of the impact of various sea ice thickness and freeboard and snow depth remote sensing products on the performance of forecasts of ice and snow volume along the Northern Sea Route.

For the land surface *Kaminski et al.* (2010) apply the QND approach to assess the impact of measurements of atmospheric CO_2 in a Carbon Cycle Data Assimilation System (CCDAS *Rayner et al.*, 2005). *Kaminski et al.* (2012) presented an interactive Mission Benefit Analysis System that applies the QND approach to assess the impact of in situ samples of atmospheric CO_2 and other observations in a CCDAS.

2.2 Methods

2.2.1 Quantitative network design

The QND methodology is presented by *Kaminski and Rayner* (2017), partly based on *Tarantola* (2005) and *Rayner et al.* (2019). To keep it concise we provide a shortened form of the presentation by *Kaminski and Rayner* (2017). QND is a generic concept that is applicable beyond the context of Arctic modelling (see examples in Section 2.1), It is, hence, useful to provide a generic presentation in this subsection. The specific elements of our application will be described in the subsequent subsections.

2 Approach

As mentioned, the QND formalism performs a rigorous uncertainty propagation from the observations to a target quantity of interest relying on the indirect link from the observations to the target variables established by a numerical model. We distinguish between four sources of uncertainty in a model simulation:

1. Uncertainty caused by the formulation of individual process representations and their numerical implementation (structural uncertainty).
2. Uncertainty in constants (process parameters) in the formulation of these processes (parametric uncertainty).
3. Uncertainty in external forcing/boundary values (such as surface winds or precipitation) driving the relevant processes.
4. Uncertainty in the state of the system at the beginning of the simulation (initial state uncertainty).

The first category reflects the implementation of the relevant processes into the model (code) while the others can be represented by a set of input quantities controlling the behaviour of a simulation using the given model implementation. The QND procedure formalises the selection of these input quantities through the definition of a control vector, x. The choice of the control vector is a subjective element in the QND procedure. A good choice covers all input quantities with high uncertainty and high impact on simulated observations d_{mod} or target quantities y (*Kaminski et al.*, 2012; *Rayner et al.*, 2019).

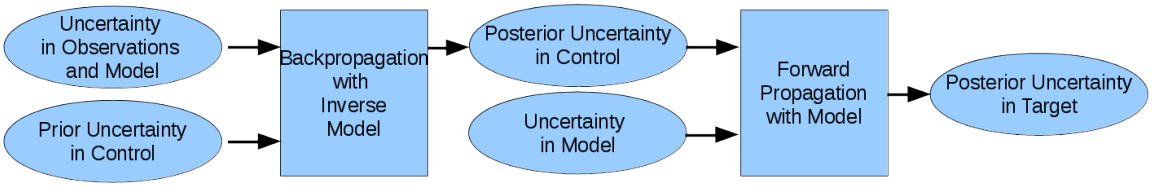


Figure 2.1: Data flow through the two-step procedure of QND formalism. Oval boxes denote data, rectangular boxes denote processing. "Backward propagation with inverse model" implements Equation (2.1), and "forward propagation with model" implements Equation (2.3). Figure is taken from *Kaminski and Rayner* (2017).

The target quantity may be any quantity that can be extracted from a simulation with the underlying model (in the current study regional integrals of predicted sea ice and snow volumes and fossil fuel emissions, see Section 2.2.2), but also any component of the control vector, for example, a process parameter such as the snow albedo. In

the general case, where the target quantity is not part of the control vector, the QND procedure operates in two steps (Figure 2.1). The first step (inversion step) uses the observational information to reduce the uncertainty in the control vector, i.e. from a prior to a posterior state of information. The second step (prognostic step) propagates the posterior uncertainty forward to the simulated target quantity.

Within the QND formalism, we present all involved quantities by probability density functions (PDFs). We typically assume a Gaussian form for the prior control vector and the observations, if necessary after a suitable transformation. The Gaussian PDFs' covariance matrices express the uncertainty in the respective quantities, i.e. $\mathbf{C}(x_0)$ and $\mathbf{C}(d_{\text{obs}})$ for the prior control vector and the observations. In the context of these PDFs we will use the term uncertainty to refer to its full covariance matrix in the case of a vector quantity, and in the case of a scalar quantity or a given vector component, it refers to the square root of the entry on the diagonal of the full covariance matrix corresponding to that particular vector component. In the latter case, the uncertainty refers to one standard deviation of the marginal PDF corresponding to that component, and we use the notation $\sigma(d_2)$ to denote, for example, the standard deviation of the second component of d.

For the first QND step, we use a mapping M from control variables onto equivalents of the observations. In our notation the observation operators that map the model state onto the individual data streams (see *Kaminski and Mathieu* (2017) and Section 2.2.5) are incorporated in M. M is in practice computed by a specific numerical model with specific inputs and outputs. Let us first consider the case of a linear model, for which we denote by $\mathbf{M}'$ the Jacobian matrix of M, i.e. the derivative of M with respect to x. In this case, the posterior control vector is described by a Gaussian PDF with uncertainty $\mathbf{C}(x)$, which is given by

$$\mathbf{C}(x)^{-1} = \mathbf{M}'^{T}\mathbf{C}(d)^{-1}\mathbf{M}' + \mathbf{C}(x_0)^{-1}\,, \tag{2.1}$$

where the data uncertainty $\mathbf{C}(d)$ is the combination of two contributions:

$$\mathbf{C}(d) = \mathbf{C}(d_{\text{obs}}) + \mathbf{C}(d_{\text{mod}}). \tag{2.2}$$

The term $\mathbf{C}(d_{\text{obs}})$ expresses the uncertainty in the observations and $\mathbf{C}(d_{\text{mod}})$ the uncertainty in the simulated equivalents of the observations $M(x)$. The first term in Equation (2.1) expresses the impact of the observations and the second term the impact of the prior information. In the non-linear case we use Equation (2.1) as an approximation of $\mathbf{C}(x)$.

The mapping N involved in the second, the uncertainty propagation step, is the mapping from the control vector onto a target quantity, y. The Jacobian matrix $\mathbf{N}'$ of the mapping N is employed to approximate the propagation of the posterior uncertainty in the control vector $\mathbf{C}(x)$ forward to the uncertainty in a target quantity, $\sigma(y)$ via

$$\sigma(y)^2 = \mathbf{N}'\mathbf{C}(x)\mathbf{N}'^{T} + \sigma(y_{mod})^2\,. \tag{2.3}$$

If the model were perfect, $\sigma(y_{mod})$ would be zero. In contrast, if the control variables were perfectly known, the first term on the right-hand side would be zero. The terms

$\mathbf{C}(d_{\mathrm{mod}})$ in Equation (2.2) and $\sigma(y_{mod})$ in Equation (2.3) capture the structural uncertainty as well as the uncertainty in those process parameters, boundary and initial values that are not included in the control vector. These two terms typically rely on subjective estimates. When comparing the effect of different data sets in the same setup, $\sigma(y_{mod})$ acts as an offset (for the respective variance) in Equation (2.3). For each observational scenario, we evaluate Equation (2.3) with $\sigma(y_{mod})$ set to zero, which sharpens the contrast between the products. For the sea ice-ocean assessments we also report a plausible estimate of $\sigma(y_{mod})$.

To conduct a valuable QND assessment, the requirement on the model is not that it simulates the target quantities and observations under investigation realistically, but rather that it provides a realistic *sensitivity* of the target quantities and observations under investigation with respect to a change in the control vector. (As a hypothetical example we can think of a perfect regional tracer model that is run with an offset in the initial or boundary conditions for a passive tracer. The simulated tracer concentration will carry this offset, but all sensitivities will be perfect.) If the sensitivities of the target quantities and observations (i.e. the Jacobians) are realistic, but the simulation of target quantities and observations is incorrect, we can always obtain a valuable QND assessment with appropriate model uncertainty. The result of the assessment may then be that a particular data stream is not useful in constraining a particular target quantity given current modelling capabilities. Under such circumstances, the QND system could be operated with reduced model uncertainty to explore the level of accuracy required of the model for a data stream to serve as a useful constraint on a given target quantity. In particular when it comes to newly available, unvalidated data streams and target quantities the accuracy of both, the simulation and the sensitivities, is hard to assess. In the case of a model that does not capture relevant processes we may expect errors in both the simulation and the sensitivities, and consequently also in the QND assessment.

Our performance metric is the (relative) reduction in posterior target uncertainty $\sigma(y)^2$ with respect to a reference. To compare against the case without any observations we first compute, as the reference, the prior target uncertainty, $\sigma(y_0)$, via

$$\sigma(y_0)^2 = \mathbf{N}'\mathbf{C}(x_0)\mathbf{N}'^T + \sigma(y_{mod})^2 \,. \tag{2.4}$$

The uncertainty reduction with respect to the prior,

$$\frac{\sigma(y_0) - \sigma(y)}{\sigma(y_0)} = 1 - \frac{\sigma(y)}{\sigma(y_0)}, \tag{2.5}$$

quantifies the impact of the entire network. A schematic illustration of the approach with the prior and posterior uncertainty ranges is shown in Figure 2.2. The observations d_1 and d_2 render a range of trajectories unlikely, which in the first step (Equation (2.1)) leads to a reduction of uncertainty in the control vector (from $C(x_0)$ to $C(x)$) and in the second step (Equation (2.3)) to a reduction in the target uncertainty (from $\sigma(y_0)$ to $\sigma(y)$). Note that the model trajectories in Figure 2.2 are for illustration purposes only; the algorithm employs the linearisations M' and N' for the uncertainty propagation.

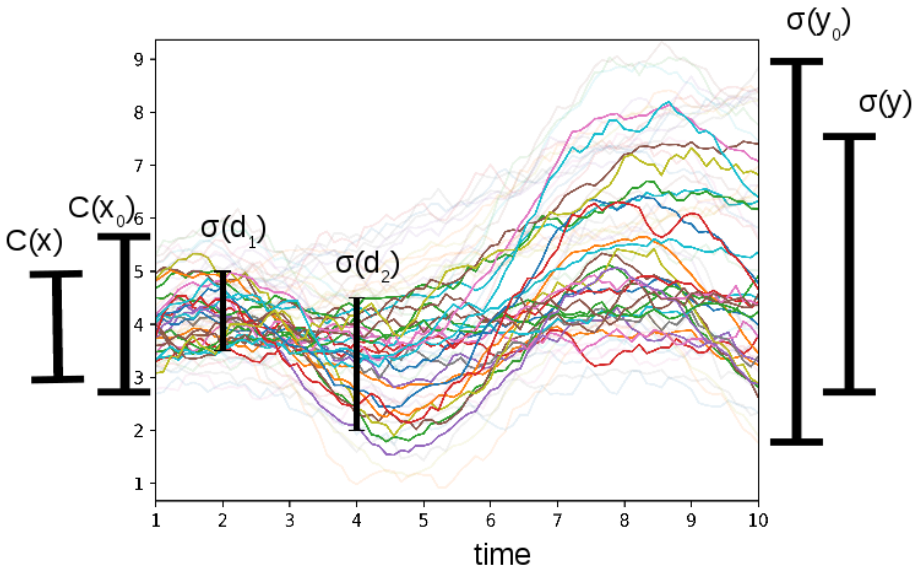


Figure 2.2: Schematic presentation of the QND procedure: Each coloured line illustrates a model trajectory that simulates for a given value of the control vector (x) counterparts of the observations (d_1 and d_2) and a target quantity (y). Through the model, the observations act as constraints on the control vector, which reduces its uncertainty from $C(x_0)$ to $C(x)$. This uncertainty reduction on the control vector translates into an uncertainty reduction in the target quantity from $\sigma(y_0)$ to $\sigma(y)$.

We note that (through Equation (2.1) and Equation (2.3)) the posterior target uncertainty solely depends on the prior and data uncertainties, the contribution of the model error to the uncertainty in the simulated target variable, $\sigma(y_{mod})$, as well as the observational and target Jacobians (quantifying the linearised model responses of the simulated observation equivalent and of the target quantities). Hence, the QND formalism can be employed to evaluate hypothetical candidate networks. Candidate networks are characterised by observational data type, location, sampling frequency and time, and data uncertainty but not the observational value. Here, we define a network as the complete set of the characterisation of observations, d, used to constrain the model. The term network is not meant to imply that the observations are of the same type or that their sampling is coordinated. For example, a network can combine different types of in situ and satellite observations.

In practice, for pre-defined target quantities and observations, model responses can be pre-computed and stored. A network composed of these pre-defined observations can then be evaluated in terms of the pre-defined target quantities without any further

model runs. Only matrix algebra is required to combine the pre-computed sensitivities with the data uncertainty. This aspect is exploited in our QND system.

For the land-based fossil fuel emissions, our inverse problem is characterised by a high-dimensional control vector combined with an observation vector of an even larger dimension. As a consequence, the size of M' is too large to solve Equation (2.1) on our computer. Hence, as *Rödenbeck* (2005), we restrict ourselves to the iterative computation of posterior uncertainties for several selected target quantities, which are described in Section 2.2.2.

2.2.2 Target quantities

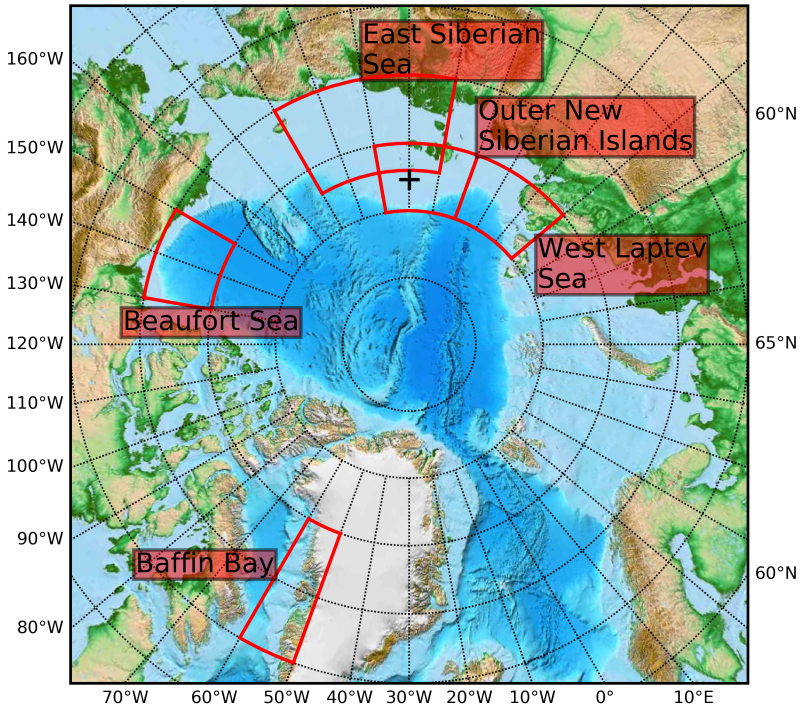


Figure 2.3: Target regions along the NSR and in the Northwestern Sea Route. Black cross indicates a location for further use in Figure 2.17.

For this study, we selected two sets of target quantities, y. The first set is particularly relevant for maritime transport, namely predicted sea ice volume (SIV) and snow volume (SNV) over three regions along the NSR, and two regions in the Northwestern Sea Route. These five regions are displayed in Figure 2.3 and respectively denoted as "West Laptev Sea" (WLS), "Outer New Siberian Islands" (ONSI), "East Siberian Sea" (ESS), "Beaufort Sea" (BFS), and "Baffin Bay" (BAB). As an additional target region we use the entire Arctic basin. We perform these predictions for May 28, 2015, a point in time at which there is still sufficient snow cover for our prediction to be relevant. We have two different assimilation windows (Figure 2.4). For freeboard/ice thickness/snow depth products we assimilate an average field over April, followed by a four-week forecast. For SST we assimilate an average field over the first week of April, followed by a seven-week forecast.

The second set of target variables are the land-based fossil fuel emissions from several Arctic countries, namely Canada, Denmark, Finland, Island, Norway, and Sweden (see

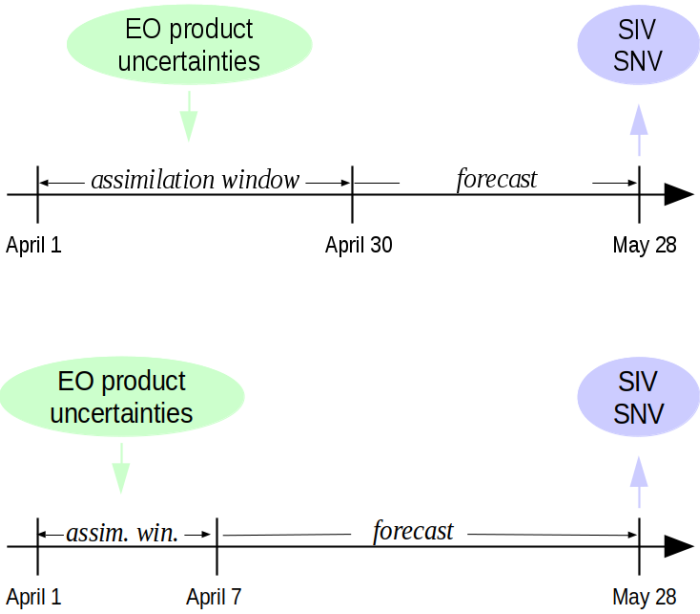


Figure 2.4: Time line of assimilation and forecast setups for freeboard/snow depth products (top) and SST (bottom).

Figure 4.2). We compute these for the first seven days in June 2008 by a sectorial fossil fuel emission model, which is described in Section 2.2.4. The chosen period is of particular interest because of the strong influence of biospheric fluxes during the growing season on atmospheric CO_2. In contrast to the sea ice-ocean setup, our observational constraint over land covers the target period, i.e. no prediction is made.

2.2.3 Sea ice-ocean model

To simulate observation equivalents (M in Equation (2.1)) and target quantities (N in Equation (2.3)) we employ a coupled model of the sea ice-ocean system. The model is required to provide realistic simulations of the sensitivity of observation equivalents and target quantities to changes in the control variables. In the present study we use the Max-Planck-Institute Ocean Model (MPIOM, *Jungclaus et al.*, 2012, 2013; *Haak et al.*, 2003), i.e. the sea ice-ocean component of the Max-Planck-Institute Earth System Model (MPI-ESM, *Giorgetta et al.*, 2013). MPI-ESM regularly provides climate projections for the Intergovernmental Panel on Climate Change (IPCC) in particular to the IPCC's 5th assessment report (*Stocker et al.*, 2013) and the upcoming 6th assessment report (AR6) and within the seasonal to decadal prediction system (*Müller et al.*, 2012). In the following, we provide a brief description of the current model development status, largely following *Jungclaus et al.* (2006) and *Niederdrenk* (2013).

The recent development of the ocean part of the model includes the treatment of horizontal discretisation which has undergone a transition from a staggered E-grid to an orthogonal curvilinear C-grid. The treatment of subgrid-scale mixing has been improved through the inclusion of a new formulation of bottom boundary layer slope convection, an isopycnal diffusion scheme, and a Gent and McWilliams style eddy-induced mixing parameterisation (*Gent and McWilliams*, 1990). Along-isopycnal diffusion is formulated following *Redi* (1982) and *Griffies* (1998). Isopycnal tracer mixing by unresolved eddies is parameterised following *Gent et al.* (1995). For the vertical eddy viscosity and diffusion the Richardson number–dependent scheme of *Pacanowski and Philander* (1981) is used. An additional wind mixing proportional to the cube of the 10-m wind speed (decaying exponentially with depth) compensates for too low turbulent mixing close to the surface. Static instabilities are removed through enhanced vertical diffusion.

Viscous–plastic rheology (*Hibler*, 1979) is used for the sea ice dynamics. Also, the parametrisation of the ice strength P is taken from *Hibler* (1979). We introduce it here explicitly for later reference:

$$P = \mathrm{hiccp} \cdot h_i \cdot \exp(-\mathrm{hibcc} \cdot (1 - c)), \tag{2.6}$$

where c denotes SIC, h_i SIT, and the parameters hiccp and hibcc are also known as pstar and cstar (see Table 2.1).

Sea ice thermodynamics are formulated using a *Semtner* (1976) zero-layer model relating changes in sea ice thickness to a balance of radiant, turbulent, and oceanic heat fluxes. In the zero-layer model the conductive heat flux within the sea ice/snow layer is assumed to be directly proportional to the temperature gradient across the sea ice/snow layer and inversely proportional to the thickness of that layer, i.e. the sea ice does not store heat. The effect of snow accumulation on sea ice is included, along with snow–ice transformation when the snow/ice interface sinks below the sea level because of snow loading (flooding). The effect of ice formation and melting is accounted for within the model assuming a sea ice salinity of $5g/kg$.

MPIOM allows for arbitrary placement of the model's poles on an orthogonal

curvilinear grid. In the setup used here (taken from *Niederdrenk* (2013); *Mikolajewicz et al.* (2015); *Niederdrenk et al.* (2016)) the poles are located over Russia and North America (Figure 2.5). Placement over land avoids numerical singularities that for poles over the ocean would be caused by the convergence of the meridians, and the non-diametric placement allows to reach high resolution (average of about 15 km) in the Arctic. This setup achieves a spatial resolution as high as that of the sea ice thickness and freeboard products we analyse (in fact over the target regions the model resolution is higher) without major computational constraints, which allows an evaluation of the full spatial information content provided by the respective EO products. Here, we will refer to this particular model configuration as Arctic MPIOM.

As forcing data at the ocean's surface, the model needs heat, freshwater, and momentum. These data are taken from ECMWF's ERA-Interim reanalysis (*Dee et al.*, 2011). ERA-Interim is a global atmospheric reanalysis (of the period from 1979 to present) that is produced by a 2006 release of the Integrated Forecasting System (IFS – version Cy31r2) and applies a 4-dimensional variational analysis with a 12-hour analysis window. The spatial resolution of the data set is approximately $80km$ (T255 spectral) on 60 vertical levels from the surface up to $0.1hPa$. ERA-interim surface variables that force Arctic MPIOM are 2-meter temperature, 2-meter dew point temperature (a surrogate of 2-meter specific humidity – not provided by ECMWF), 10-meter zonal and meridional wind velocity (to calculate the wind speed), total cloud cover and the following fluxes (provided in accumulated form over the 12-hourly forecast window): surface downward solar radiation, surface downward thermal radiation, total precipitation, zonal and meridional wind stress. Land runoff into the ocean is taken from the German Ocean Model Intercomparison Project (OMIP, *Röske*, 2001).

For the computation of the Jacobians M' and N' (introduced in Section 2.2.1) that is described in Section 2.3 we run Arctic MPIOM from a restart file for April 1, 2015. This restart file is in turn generated from a hindcast run of Arctic MPIOM that is initialised on January 1, 1979. This initialisation is based on a set of observations that consists of a topography data set (ETOPO5 5-minute gridded elevation data, *NOAA*, 1988), and a hydrographic climatological data set (Polar science center Hydrographic Climatology, PHC3; *Steele et al.*, 2001) containing potential temperature and salinity. The ocean initially is assumed to be at rest. Sea ice is assumed to be present if the sea surface temperature falls below the freezing temperature of sea water. 100% ice cover and a sea ice thickness of $2m$ is assumed where sea ice is present and sea ice is assumed to be at rest. From this initial state the model is integrated with the ERA-Interim surface forcing until March 31, 2015 (the beginning of our assimilation window). While a 36-year integration is certainly too short to spin up the deep ocean, it is sufficient for this study, because the spinup time of sea ice and the upper ocean (depth above about $500m$) is generally assumed to be only a few decades.

For a successful QND assessment it is essential that MPIOM provides realistic sensitivities of the observation equivalent and the target quantities to the changes in the control vector (Equation (2.1) and Equation (2.3)). However, observations are not available to validate these sensitivities. The only validation of MPIOM possible is

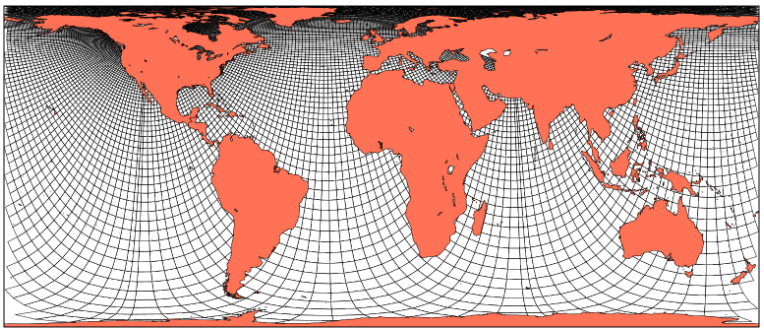
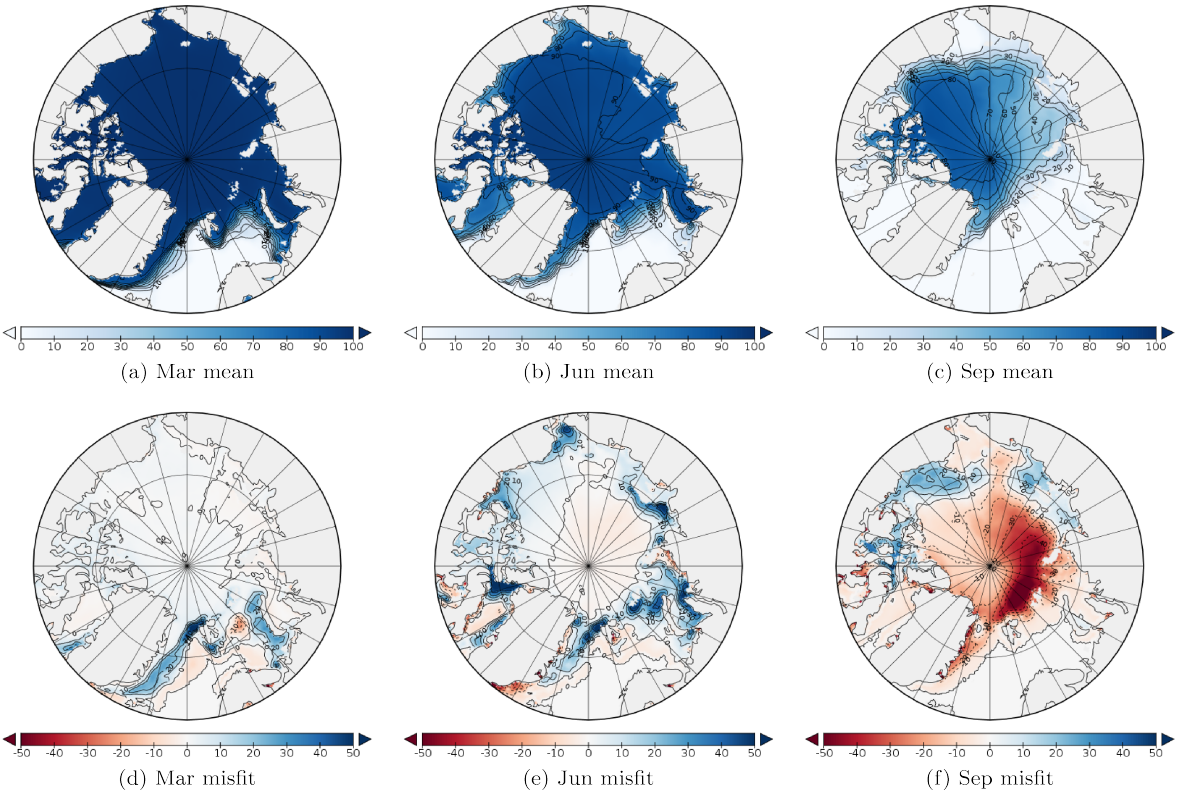


Figure 2.5: Model grid, mesh indicates groups of 4 by 4 grid cells.

against observations of the state of the sea ice and ocean. In the following we present comparisons with selected observation-based products first for the hindcasting period, and then for the assimilation window and the forecasting period.

The hindcast with Arctic MPIOM has been validated against remotely sensed ice concentration from the reprocessed Ocean and Sea Ice Satellite Application Facility (OSI-SAF) sea ice concentration product (*Eastwood et al.*, 2015) and against a combination of in situ and remotely sensed ice thickness observations. In situ observations of sea ice thickness still have high uncertainty, and each data source has its strengths and weaknesses. As of today the most reliable pan-Arctic sea ice thickness data set is derived from a combination of various sources of in situ observations and remotely sensed satellite sea ice thickness products (*Lindsay and Schweiger*, 2015).

The reprocessed OSI-SAF sea ice concentration product is available daily on a $10km$ spatial grid and includes spatially and temporally varying uncertainty estimates. For an assessment of the performance of the Arctic MPIOM, the sea ice concentration has been compared to the long-term means of the March, June, and September monthly means for the period 1990 to 2008 (Figure 2.6). In March (panel d) and June (panel e) only relatively small scale misfits to the OSI-SAF ice concentration are found but they can reach up to 50% (here and in the following we use the term for the model-observation difference). The sea ice margin in the Nordic Seas and the Barents Sea is captured well. The anomalies apparent in March correspond to the results of a study performed with the MPIOM version of the Max-Planck-Institute's Earth System model MPI-ESM-LR (*Notz et al.*, 2013), for which the MPIOM was forced with the same atmospheric forcing data set as used in our study (ERAinterim) (see panel f of their Figure 3). In September large misfits to the OSI-SAF sea ice concentration are obtained (Figure 2.6 panel f). Especially over the Eurasian basin the model's sea ice margin is located too far north, but also over the central Arctic the model underestimates the sea ice concentration. In our target regions the misfit remains relatively small. The aforementioned analysis by *Notz et al.* (2013) shows similar misfits (see panel f of their Figure 4) to a different sea ice concentration data set, namely NSIDC-CDR (National Snow and Ice Data Center Climate Data Record).



(a) Mar mean (b) Jun mean (c) Sep mean

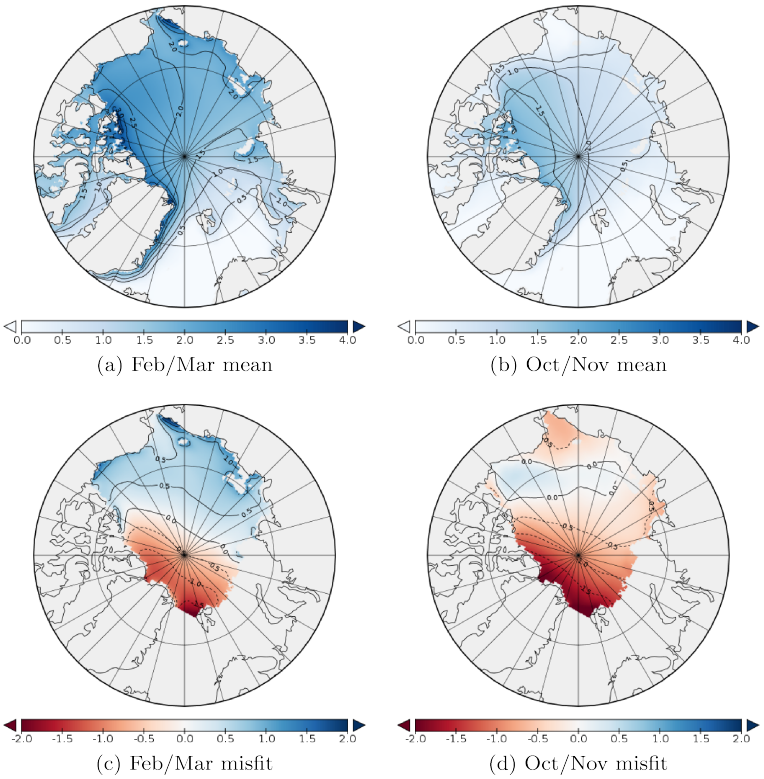
(d) Mar misfit (e) Jun misfit (f) Sep misfit

Figure 2.6: Long-term mean sea ice concentration [%] of the Arctic MPIOM for 1990 to 2008 for March, June and September (panel a to c) and the misfit to the OSI-SAF sea ice concentration (panel d to f). In panels d to f, red colours indicate underestimation and blue colours overestimation of sea ice concentration in the model.

An evaluation of the hindcast simulation with Arctic MPIOM with respect to the modelled SIT is much more difficult because the observation-based products exhibit large uncertainties reflecting the corrections imposed by the respective measurement principle. For example, electro-magnetic Air-EM measurements detect the air-snow interface, and not the interface between snow and sea ice, introducing significant errors in the SIT estimates that are corrected by assumptions or measurements of snow depth. Moored and submarine ULS measurements have to be corrected for the first return echo. Differences in the observed and measured spatial scales further complicate the comparison. The aforementioned study of *Lindsay and Schweiger* (2015) synthesises all available in situ and remotely sensed satellite SIT products in an ice thickness regression procedure (ITRP) for the period from year 2000 to 2012. Low order spatial and temporal polynomials are fitted to the available sea ice thickness measurements. The resulting sea ice thickness regression product describes the evolution in the central Arctic and is linear in time plus a quadratic time-dependent component, i.e. it does not contain year-to-year variability. Uncertainty ranges are deduced from the uncertainty

of the individual regression coefficients. The year-to-year variability is reflected in this uncertainty. *Lindsay and Schweiger* (2015) show for example that the ICESat ice thickness product from the Jet Propulsion Laboratory (ICESat-JPL, *Kwok and Cunningham* (2008)), which is widely used for model validation, has a large positive bias of about $40cm$. Here we compare the modelled long-term mean (2000 to 2012) sea ice thickness of the Arctic MPIOM hindcast to the ITRP sea ice thickness for the two-months periods February/March and October/November. We selected these two-month periods because the availability of the ICESat satellite product ensures high data coverage in the ITRP. The long-term mean sea ice thickness of the Arctic MPIOM hindcast simulation for February/March and October/November is depicted in Figure 2.7 (panel a and panel b) together with the misfit to the ITRP ice thickness (panel c and panel d). A prominent feature is a strong underestimation of the Arctic MPIOM sea ice thickness north and west of Fram Strait and in the strait itself. In the target areas along the NSR and the target region in the Beaufort Sea the misfit is moderate in February/March (overestimation of about 25%) with the exception around the New Siberian Islands where the misfit can reach more than one meter (overestimation of about 50%). In October/November the misfit is very moderate in these areas except for Bering Strait where Arctic MPIOM underestimates the sea ice thickness by more than $50cm$. In the Baffin Bay, no sea ice thickness situ data entered the ITRP.

Next, we address Arctic MPIOM performance over our assimilation and forecasting period (see Figure 2.4). We show the April mean and the May 28 mean of the modelled SIT and the misfit of the April mean thickness to that retrieved from CryoSat-2 (Figure 2.8). For a comparison of CryoSat-2 thickness to in situ observations we refer to *Haas et al.* (2017). The misfit to the CryoSat-2 ice thickness in April 2015 is similar to the misfit to the ITRP shown in Figure 2.7: a strong underestimation north of the Canadian Archipelago and north and west of Fram Strait and a moderate overestimation in the area of the target quantities of about or less than $50cm$ (about 25% relative error). Figure 2.9 depicts the April mean and the May 28 mean of the modelled snow depth and the misfit to the modified Warren climatology (*Warren et al.*, 1999) that is used in the CryoSat-2 retrieval (see Section 2.2.5). The main challenge for sea ice thickness retrieval with satellite altimeters is the parameterisation of snow depth on sea ice, which is still not measured routinely. The current CryoSat-2 retrieval uses a modified snow climatology that addresses shortcomings of the *Warren et al.* (1999) climatology that was based largely on data from drifting stations mainly on multi-year sea ice collected over the past decades, and hence is not reflective of a much younger, more seasonal Arctic ice cover. Given the increased fraction of first-year ice in the Arctic Ocean, the approach proposed by *Kurtz and Farrell* (2011) is used and the climatological snow depth values used in the retrieval are multiplied over first-year ice by a factor of 0.5. Note that on May 28 over the target regions a large fraction of snow cover has already melted. The misfit to the modified Warren climatology in the target area East Siberian Sea is on the order of about $10cm$ (50% relative error) but much less for the other target areas.



(a) Feb/Mar mean (b) Oct/Nov mean

(c) Feb/Mar misfit (d) Oct/Nov misfit

Figure 2.7: Long-term mean (2000 to 2012) of the simulated sea ice thickness $[m]$ for the two-month periods February and March and October and November (panel a and b) and the misfit (model – observations) to the ITRP (panel c and d). In panels c and d, red colours indicate underestimation and blue colours overestimation of sea ice thickness in the model.

Overall, the misfits of the Arctic MPIOM are acceptable in particular for our target regions along the NSR and the Northwestern Sea Route (Figure 2.3) and are comparable to misfits found in sea ice-ocean model intercomparison projects (e.g, *Chevallier et al.* (2017)).

The criteria for the choice of the control vector are presented in Section 2.2.1. The specification of prior, both mean (x_0) and uncertainty ($C(x_0)$), follows *Kaminski et al.* (2015), and is listed in Table 2.1. The largest possible control vector in our modelling system is the superset of initial and surface boundary conditions as well as all parameters in the process formulations, including the observation operators. As described in Section 2.3, the Jacobian computation requires an extra run for each additional component of the control vector. To keep our QND system numerically efficient, two and three-dimensional fields are partitioned into regions. More precisely, we divide the Arctic domain into nine regions (shown in Figure 2.10). In each of these regions we add a scalar perturbation to each of the forcing fields (indicated in Table 2.1 by "f" in the type column); the perturbation is applied for the entire simulation time. Likewise, we add a scalar perturbation to six initial fields indicated

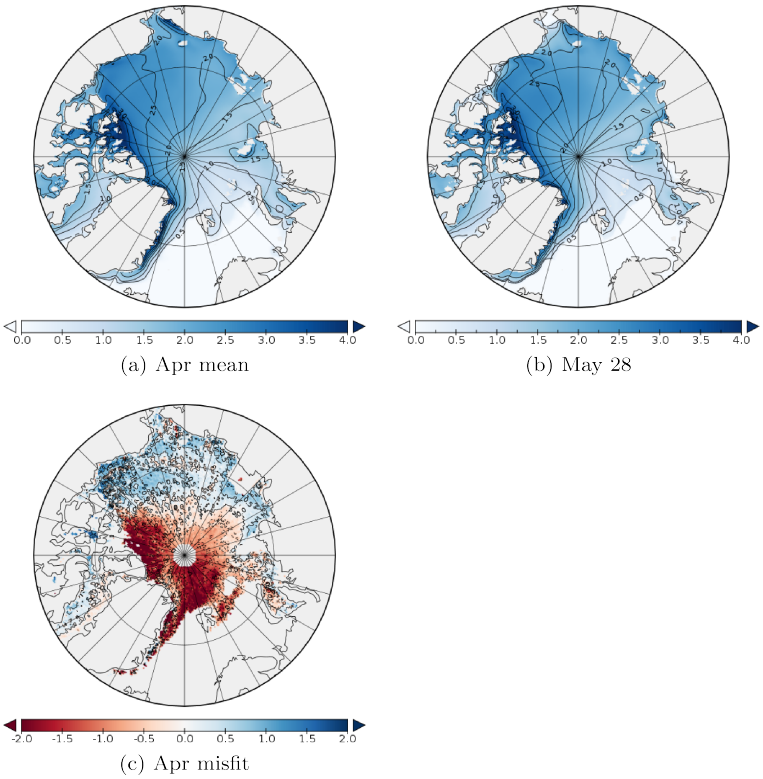
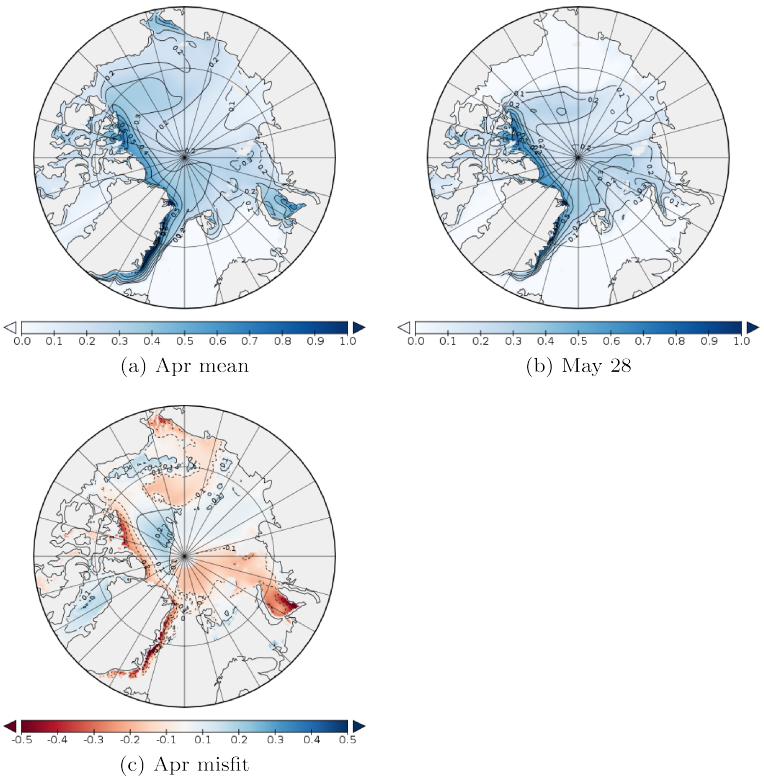


Figure 2.8: a) The modelled mean April 2015 sea ice thickness $[m]$, b) the modelled sea ice thickness on May 28 2015, and c) the mean April 2015 misfit of the modelled sea ice thickness relative to the CryoSat-2 sea ice thickness. In panel c, red colours indicate underestimation and blue colours overestimation of sea ice thickness in the model.

in Table 2.1 by "i" in the type column. For the ocean temperature and salinity the size of the perturbation is reduced with increasing depth (and zero below $500m$). Finally, we have selected 29 process parameters from the sea ice–ocean model plus two parameters from the observation operators for freeboard products (see Section 2.2.5 for details). This procedure results in a total of 157 control variables. We assume the prior uncertainty to have diagonal form, i.e. there are no correlations among the prior uncertainties relating to different components of the control vector. The diagonal entries are the square of the prior standard deviation. For process parameters this standard deviation is estimated from the range of values typically used within the modelling community. The standard deviation for the components of the initial state is based on a model simulation over the past 37 years and computed for the 37 member ensemble corresponding to all states on the same day of the year. Likewise, the standard deviation of the surface boundary conditions is computed for the 37 member ensemble corresponding to the April-October means of the respective year.

Table 2.1: Control variables. Column 1 lists the quantities in the control vector; column 2 gives the abbreviation for each quantity; column 3 indicates whether the quantity is a process parameter ("p"), an initial field ("i"), or an atmospheric boundary field (denoted by "f" for forcing); column 4 gives the name of each quantity; column 5 indicates the prior uncertainty (one standard deviation) and the corresponding units (unless unitless) and provides the mean parameter value in parenthesis, where applicable; and column 6 identifies the position of the quantity in the control vector – for initial and boundary values (which are differentiated by region) this position refers to the first region, while the following components of the control vector then cover regions 2 to 9.

Index #	Name	Type	Meaning	Prior uncertainty (Mean)	Start
1	hiccp	p	(alias pstar) ice strength (devided by density)	15(20) $[N/m^2/kg]$	1
2	hibcc	p	(alias cstar) ice strength depend. on ice conc.	5.0(20.0)	2
3	hicce	p	(alias eccen) squared yield curve axis ratio	0.5(2.0)	3
4	rlc1	p	extra lead closing (Notz et al., 2013)	0.2(0.25)	4
5	rlc2	p	extra lead closing (Notz et al., 2013)	1.0(3.0)	5
6	rlc3	p	extra lead closing (Notz et al., 2013)	1.0(2.0)	6
7	h_0	p	lead closing	1.0(0.5) $[m]$	7
8	hmin	p	mimimal ice thickness	0.04(0.05) $[m]$	8
9	armin	p	minimal ice compactness	0.15(0.15)	9
10	hsntoice	p	limit on flooding	0.45(0.45)	10
11	sice	p	salinity in sea ice	2.0(5.0)$[g/kg]$	11
12	albi	p	freezing ice albedo	0.1(0.75)	12
13	albm	p	melting ice albedo	0.1(0.70)	13
14	albsn	p	freezing snow albedo	0.1(0.85)	14
15	albsnm	p	melting snow albedo	0.1(0.70)	15
16	rhoice	p	density of sea ice	20(910)$[kg/m^3]$	16
17	rhosn	p	density of snow	20(330)$[kg/m^3]$	17
18	cw	p	ocean drag coefficient	$2.0 \times 10^{-3}(4.5 \times 10^{-3})$	18
19	av0	p	coefficient vertical viscosity	$1. \times 10^{-4}(2. \times 10^{-4})[m^2/s]$	19
20	dv0	p	coefficient vertical diffusitivity	$1. \times 10^{-4}(2. \times 10^{-4})[m^2/s]$	20
21	aback	p	background coefficient vertical viscosity	$3. \times 10^{-5}(5. \times 10^{-5})[m^2/s]$	21
22	dback	p	background coefficient vertical diffusitivity	$1. \times 10^{-5}(1.05 \times 10^{-5})[m^2/s]$	22
23	cwt	p	vertical wind mixing parameter tracers	$2.0 \times 10^{-4}(3.5 \times 10^{-4})[m^2/s]$	23
24	cwa	p	vertical wind mixing parameter momentum	$0.4 \times 10^{-3}(0.75 \times 10^{-3})[m^2/s]$	24
25	cstabeps	p	vertical wind mixing stability parameter	0.03(0.06)	25
26	cdvocon	p	coefficient for enhanced vertical diffusivity	0.1(0.15)	26
27	bofric	p	linear bottom friction	$2. \times 10^{-4}(3. \times 10^{-4})[m^2/s]$	27
28	rayfric	p	quadratic bottom friction	$0.5 \times 10^{-3}(1. \times 10^{-3})[m^2/s]$	28
29	jer_a	p	jerlov atten - ocean-water types	0.04(0.08)	29
30	jer_b	p	jerlov bluefrac - ocean-water types	0.20(0.36)	30
31	albw	p	open water albedo	0.05(0.1)	31
32	SIT	i	initial ice thickness	0.5 $[m]$	32
33	SIC	i	initial ice concentration	0.1	41
34	SND	i	initial snow thickness	0.2 $[m]$	50
35	TEMP	i	initial ocean temperature	0.5 $[K]$ (vertically decreasing)	59
36	SAL	i	initial salinity	0.5 $[g/kg]$ (vertically decreasing)	68
37	SLH	i	initial sea level elevation	0.08 $[m]$	77
38	CLD	f	cloud cover	0.07	86
39	PREC	f	total precipitation	0.4×10^{-8} $[m/s]$	95
40	SWR	f	solar downward radiation	6. $[W/m^2]$	104
41	TDEW2	f	2 m dew point temperature	1.1 $[K]$	113
42	TEMP2	f	2 m air temperature	1.2 $[K]$	122
43	WND10	f	10m scalar wind speed	0.6 $[m/s]$	131
44	WIX	f	zonal wind stress x component	0.02 $[Nm^2]$	140
45	WIY	f	meridional wind stress y component	0.02 $[Nm^2]$	149



(a) Apr mean (b) May 28

(c) Apr misfit

Figure 2.9: a) The modelled mean April 2015 snow depth $[m]$, b) the modelled snow depth on May 28 2015, and c) the mean April 2015 misfit of the modelled snow depth relative to the modified Warren climatology used in the CryoSat-2 sea ice thickness retrieval. In panel c, red colours indicate underestimation and blue colours overestimation of snow depth in the model.

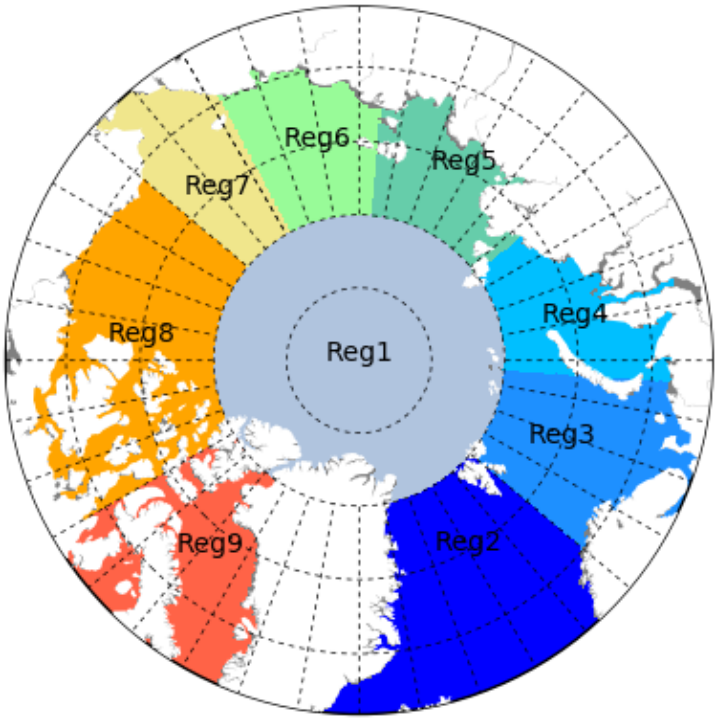


Figure 2.10: Sub-regions for spatial differentiation of initial and boundary values in the control vector. 1 (light plum): central Arctic; 2 (dark blue): North Atlantic; 3 (blue) Barents Sea; 4 (light blue) Kara Sea; 5 (green) Laptev Sea, 6 (light green) East Siberian Sea; 7 (yellow): Bering Strait/Chukchi Sea; 8 (orange): Beaufort Sea; 9 (red): Baffin Bay.

2.2.4 Modelling Chain for Atmospheric CO_2

For the simulation of atmospheric CO_2 we use a modelling chain that combines models of land-based fossil fuel emissions and natural surface fluxes with an atmospheric transport model to simulate atmospheric CO_2 concentrations. The transport model together with the atmospheric CO_2 data sets are described in Section 2.2.6. This modelling chain relies on pre-existing components and was developed in the ESA-funded Carbon Cycle Fossil Fuel Data Assimilation (CCFFDAS) study (http://ccffdas.inversion-lab.com/).

The objective of the Copernicus HPCM CO2M, which aims at observing total column CO_2 concentrations, is to constrain fossil fuel emissions of CO_2, in particular to assist the verification of the Paris agreement. Here we analyse the constraint of CO2M and the surface in situ network to constrain the sectorial national total fossil fuel emissions for a set of Arctic countries (see Section 2.2.2).

For the fossil fuel emissions, we use the model underlying the Fossil Fuel Data Assimilation System (FFDAS) by *Asefi-Najafabady et al.* (2014), which defines two sectors: a power generation sector and all other emissions denoted as the "other" sector. Through the Kaya identity (*Nakicenovic*, 2004) the emissions F from the other sector at a point x in country c are given by

$$F(x, c) = pP(x)g(x)ef(c) \tag{2.7}$$

where lower case variables denote elements of the control vector (see Table 2.2). In summary, these are a scalar multiplier for the population density (p), a pointwise estimate of the per capita GDP ($g(x)$), a global constant for the energy intensity of the economy (e), and a per-country estimate for the carbon intensity of energy production ($f(c)$). For the carbon intensity we apply an uncertainty of 10 times it's prior value (denoted as f_0 in Table 2.2) with a floor value of 10^{-6} kgC/MJ. $P(x)$ is the population density. The population dataset is based on the gridded population density provided by the LandScan Global Population Database (*Bhaduri et al.*, 2007).

We use nightlights as an additional observation to constrain the fossil fuel emissions model. The data originate from the Defense Meteorological Satellite Program (DMSP) and are archived by the US National Oceanic and Atmospheric Administration at the National Geophysical Data Center (NGDC) (http://ngdc.noaa.gov/eog/index.html). At NGDC nightlights for the years 2006 and 2010 are available. We used the products for these two years and applied a linear interpolation to obtain nightlights for our study year 2008.

Uncertainties in the nightlight observations arise from errors in the instrument as well as from aggregating the native 30" observation to the 0.1° grid used here. To estimate these uncertainties we use the same approach as in *Asefi-Najafabady et al.* (2014), given by

$$\sigma_{\mathrm{NL}} = (0.5 + 0.1x_{\mathrm{NL}})1.25 \tag{2.8}$$

where x_{NL} is the observed nightlight value. The uncertainty is inflated by 25% (the factor 1.25 in Equation (2.8)) to allow for interpolation error.

For the power generation sector we use the emissions from a list of power plants in each country. As in *Asefi-Najafabady et al.* (2014) emissions associated with each power plant were mapped onto the FFDAS grid based on the power plant location and assigning the emission value to the entire grid cell with no subgrid cell spatial distribution. When location information was not available for a power plant, the power plant emissions were proportionally distributed to the rest of the grid cells containing power plants in that country. Control variables in the power generation sector are the emissions from each listed power station (see Table 2.2). These emissions are based on the CARMA database (see http://carma.org).

Uncertainties on the CARMA power plant emissions are estimated from the average of the reported standard error from plants listed in the US Energy Information Administration and/or the US EPA Clean Air Market datasets. For a detailed description of the emission and emission uncertainty estimates from the world's power plants we refer to the supporting material of *Asefi-Najafabady et al.* (2014). We use a floor value of 0.01 MtC for the prior uncertainty of an individual power plant. This results in relative prior uncertainties well above 100% for small power plants. Using this procedure, for four countries (China, India, South-Africa, and USA) the prior uncertainties aggregated to country-scale were below the uncertainty that *Asefi-Najafabady et al.* (2014) assume for the national totals provided by the International Energy Agency (IEA; *International Energy Agency* (2011)). In each of these four countries the prior uncertainty for each of the power plants were upscaled to match the uncertainty in the national total.

Table 2.2: List of control variables for the fossil fuel emissions model.

Description	Symbol	Resolution	Prior Uncertainty
Per capita GDP	g	gridcell	1.25 dollar/yr/person
Energy intensity	e	global	10^{-6} MJ/dollar
Carbon intensity	f	country	$\max(10 f_0, 10^{-6}\text{kgC/MJ})$
Population density scaling	p	global	10^{-6}
Nightlight scaling	n	global	see Equation (2.8)
Power plant emissions (CARMA)	u	individual power plants	see text

The terrestrial biosphere model we used to calculate the natural terrestrial CO_2 exchange fluxes is conceptionally similar to the one used by *Kaminski et al.* (2002) and based on *Knorr and Heimann* (1995). The model was previously employed for the assimilation of column integrated atmospheric CO_2 (XCO2) (*Kaminski et al.*, 2017a) in a Carbon Cycle Data Assimilation System (CCDAS). It operates on a 0.5° global grid and divides the global terrestrial biosphere into 8 land cover classes (in the following also referred to as Plant Functional Types, PFTs) based on the MODIS land cover classification (*Friedl et al.*, 2010) (see Figure 2.11):

1. cropland/urban/natural vegetation mosaic,

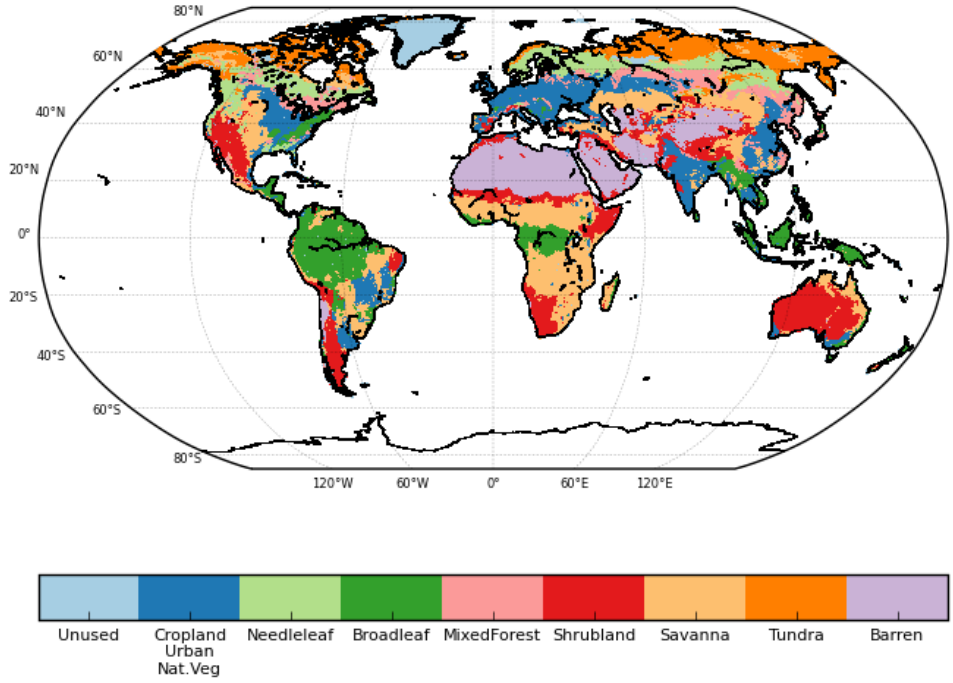


Figure 2.11: PFT groupings by land cover classes cropland/urban/natural vegetation mosaic, needle-leaf forest, broadleaf forest, mixed forest, shrubland, savanna, tundra, barren/sparsely vegetated.

2. needleleaf forest,
3. broadleaf forest,
4. mixed forest
5. shrubland,
6. savanna or grassland,
7. tundra,
8. barren or sparsely vegetated.

The model calculates the uptake of CO_2 by photosynthesis (expressed as Net Primary Productivity, NPP) using a light-use efficiency approach:

$$\mathrm{NPP}(x,t) = \epsilon \cdot \alpha(x,t) \cdot \mathrm{FAPAR}(x,t) \cdot \mathrm{srad}(x,t) \tag{2.9}$$

where ϵ denotes the PFT-specific light-use efficiency parameter, α a water stress factor, FAPAR the fraction of absorbed photosynthetically active radiation, and srad the incident solar radiation driving photosynthesis.

Heterotrophic respiration $\mathrm{R_{het}}$ is calculated following a $\mathrm{Q_{10}}$ functional relationship with temperature T and is, as the photosynthesis calculation, modulated by the water stress factor:

$$\mathrm{R_{het}}(x,t) = n(x) \cdot \alpha(x,t) \cdot Q_{10}^{T/10} \tag{2.10}$$

The PFT-specific parameter $\mathrm{Q_{10}}$ expresses the ratio of respiration at T + 10 to that at T, with T measured in °C. The control variables for the terrestrial biosphere model component are the parameters ϵ and $\mathrm{Q_{10}}$ (see Table 2.3). The spatially varying normalisation factor $n(x)$ is the ratio of the temporal integrals of $\mathrm{NPP}(x,t)$ and $\mathrm{R_{het}}(x,t)$ computed for prior parameter values over the entire model simulation period from 2004 to 2008. The normalisation factor thus ensures a balanced prior biosphere over the model simulation period but allows for non-zero posterior net flux. This is different from the study by *Kaminski et al.* (2002) which assimilated the observed seasonal cycle and enforced an annually balanced posterior biosphere.

In summary, the model uses as driving data global fields of temperature, FAPAR, and the water stress factor. The temperature fields are taken from the CRU data set (*Harris et al.*, 2014) and the plant water-stress factor, taken as monthly climatological AET/PET (actual divided by potential evapotranspiration) values per grid-cell, is computed with the BETHY model of *Knorr* (2000) for the years 1989 to 2012 using the same CRU data set to provide meteorological input data for BETHY. We use the FAPAR product (*Pinty et al.*, 2011) derived with JRC-TIP (*Pinty et al.*, 2007) directly on the 0.5° model grid, to avoid significant errors which would be imposed by spatial aggregation of the FAPAR product derived on a finer grid; for details we refer to *Kaminski et al.* (2017b).

Table 2.3: List of control variables for the terrestrial biosphere fluxes model.

Description	Symbol	Resolution	Prior Uncertainty
Light use efficiency	ϵ	PFT	100% of prior
Respiration relation to temperature	$\mathrm{Q_{10}}$	PFT	100% of prior

Emissions from biomass burning are derived from a combination of burned area data provided by *Chuvieco et al.* (2016) with *Giglio et al.* (2013) and exchange fluxes with the ocean from a combination of *Takahashi et al.* (1999) climatology with interannual variability taken from *Buitenhuis et al.* (2010), as described in *Kaminski et al.* (2017a).

The values for the prior model parameters (ϵ and Q_{10}) have been selected such that they are consistent with XCO2 observations obtained from GOSAT *Kaminski et al.* (2017a). Their relative prior uncertainties were set to 100%.

2.2.5 sea ice-ocean data sets and observation operators

The AWI (*Ricker et al.*, 2014) retrieves radar freeboard, sea ice freeboard, and sea ice thickness products from observations provided by the CryoSat-2 mission. Of these variables, a preceding QND study (*Kaminski et al.*, 2018) identified radar freeboard (RFB) as the one with the highest impact. We thus use AWI's RFB data set as the starting point for our assessments. In addition we evaluate several snow-on-sea-ice (snow depth, SND) data sets, a laser freeboard (LFB) data set, a data set that characterises the RFB information to be expected from Sentinel 3 (S3), and SST data sets. Below, we describe these data sets and the simulation of their model equivalents, i.e. the respective observation operators that provide the links from the model's state variables to the respective data sets (*Kaminski and Mathieu*, 2017). Recall that the (combination of) data set(s) enters the QND algorithm through its uncertainty $C(d)$ and that the observation operator is incorporated in the model M (see Section 2.2.1).

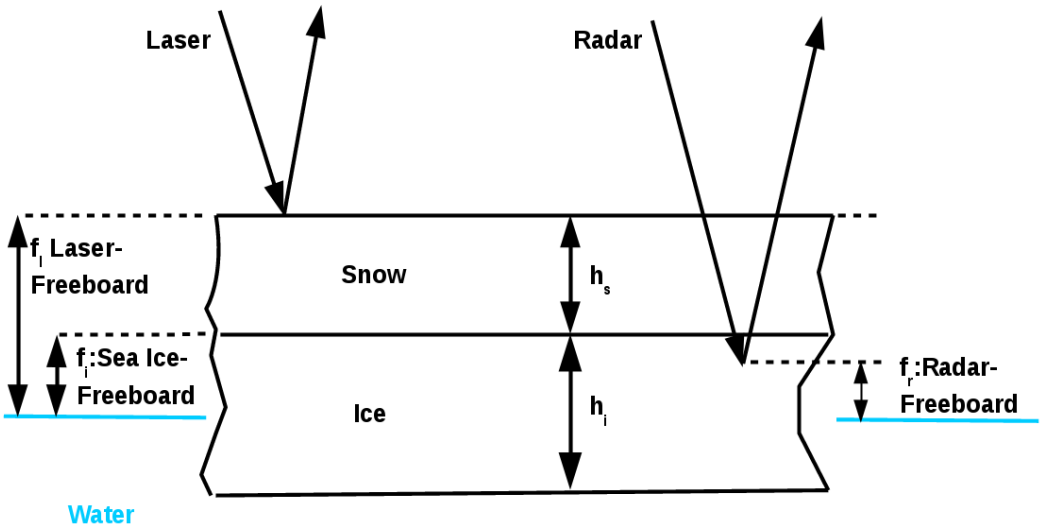


Figure 2.12: Schematic illustration of sea ice thickness and different freeboard variables.

The three products derived by AWI from CryoSat-2 are SIT (h_i), SIFB (f_i), and RFB (f_r). Their definition is illustrated in Figure 2.12 together with that of LFB (f_l).

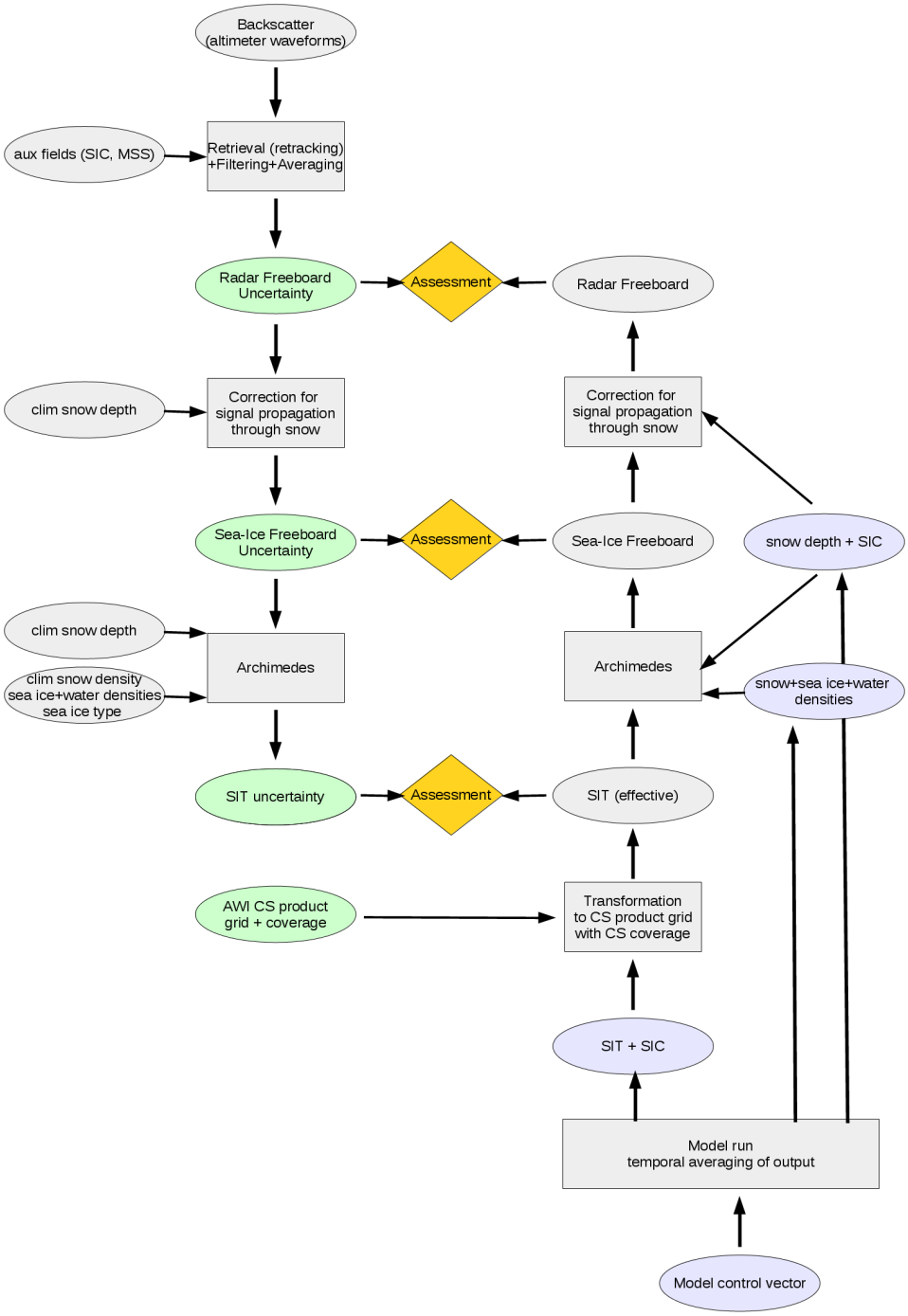


Figure 2.13: Overview on the processing chain for CryoSat-2 (CS) product retrievals (left-hand side) and the chain for modelling product equivalents from the control vector ($M(x)$, right-hand side). Oval boxes denote data and rectangular boxes processing steps. Green colour emphasises remote sensing products and violet colour model variables. Yellow diamonds mark the assessment of the EO products with the QND algorithm. MSS: mean sea surface height.

The retrieval chain is described in detail by *Ricker et al.* (2014) and *Hendricks et al.* (2016). Recall that for each product, to run an assessment, we need the spatio-temporal coverage as well as the uncertainty ranges. The left-hand side of Figure 2.13 summarises the main steps in the retrieval chain, starting with the rawest (lowest-level) product (RFB) on top. When descending from RFB via SIFB to SIT each step adds further assumptions, which contribute to the product uncertainty. The other element required to evaluate a given product is the observational Jacobian (M'), i.e. the sensitivity of the model simulation to a change in the control vector. The right-hand side of the graph illustrates how the equivalents of the respective products are simulated from the relevant model variables, which are denoted in violet colour. On this side of the graph, the complexity increases from bottom to top, i.e. from SIT via SIFB to RFB. For example, in the assessment of the SIT product, the uncertainty in quantities needed to apply the Archimedes' principle to the ice floe (including that of snow depth derived from climatology) is contained in the retrieval product, whereas the observation operator that extracts the product equivalent from the model is relatively simple (Archimedes' principle is described, for example, by *Guerrier and Horley* (1970)). We note that, while retrieved SIT is the effective SIT ($h_{i,eff}$), i.e. refers to the average over the ice-covered area of a grid cell, simulated SIT refers to the grid-cell average, i.e. for the Jacobian calculation it has to be divided by the simulated sea ice concentration (SIC, denoted by c):

$$h_{i,eff} = h_i/c. \tag{2.11}$$

Likewise for snow depth:

$$h_{s,eff} = h_s/c. \tag{2.12}$$

At the level of RFB, by contrast, it is the observation operator that includes on the modelling branch, the application of Archimedes' principle for which it requires simulated snow depth and the densities of snow (ρ_s), sea ice (ρ_i), and water (ρ_w), while the retrieval product is relatively raw. In particular this retrieval product is not affected by uncertainties due to assumptions concerning the snow depth, ρ_s, ρ_i, and ρ_w.

The observation operators for f_i, for f_r, and for f_l are:

$$\begin{aligned} f_i &= h_i/c - (\rho_i h_i/c + \rho_s h_s/c)/\rho_w \\ &= (1 - \rho_i/\rho_w)h_i/c - (\rho_s/\rho_w)h_s/c \end{aligned} \tag{2.13}$$

(2.14)

$$\begin{aligned} f_r &= f_i - 0.22 h_s/c \\ &= (1 - \rho_i/\rho_w)h_i/c - (0.22 + \rho_s/\rho_w)h_s/c \end{aligned} \tag{2.15}$$

$$\begin{aligned} f_l &= f_i + h_s/c \\ &= (1 - \rho_i/\rho_w)h_i/c + (1 - \rho_s/\rho_w)h_s/c. \end{aligned} \tag{2.16}$$

The term $-0.22 h_s/c$ in Equation (2.15) adds to the simulated f_i the correction for the difference in propagation speed of the radar signal in snow compared to air, which

is affecting f_r (*Hendricks et al.*, 2016). This is the reason why f_r is located below f_i in Figure 2.12. We note that, in these three observation operators, f_i, f_r, and f_l have the same sensitivity to h_i, but sensitivities to h_s and c differ. The sea ice component of the MPIOM uses constant densities of snow, sea ice, and water. As simulated freeboard is relatively sensitive to densities of snow and sea ice, we have, however, included these quantities as parameters of the observation operator in the control vector (see Section 2.2.3). For ρ_i=910.0kg/m^3, ρ_s=330.0kg/m^3, ρ_w=1025.0kg/m^3, the sensitivity of f_i, f_r, and f_l to a change in h_i/c is $a = 0.112$, and the respective sensitivities to a change in h_s/c are $b = -0.322$, $b = -0.542$, and $b = 0.678$.

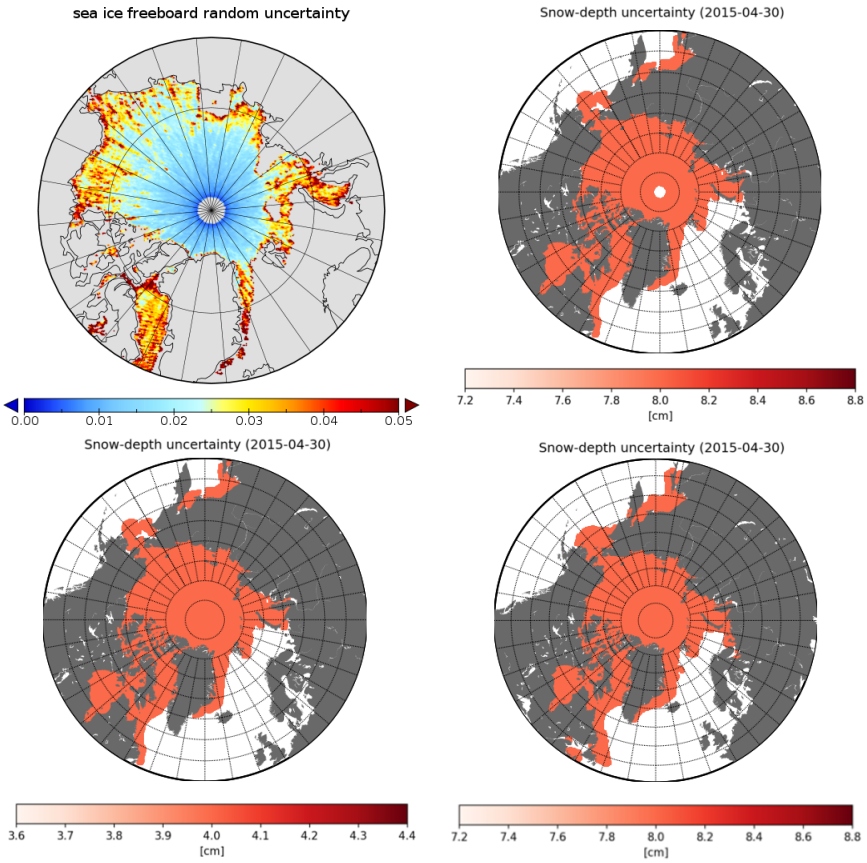


Figure 2.14: Uncertainty ranges [m] and spatial coverage for April 2015: CryoSat-2 RFB (top left), a CRISTAL-like SND product (top right), a CIMR-like SND product (bottom left), a product from a reanalysis based approach (bottom right)

Table 2.4 provides an overview on the products we assess and Figure 2.14 shows product uncertainties and coverage. As mentioned, out of the AWI products derived from CryoSat-2 the RFB product showed the highest benefit on forecast performance in a preceding QND study (*Kaminski et al.*, 2018) and we thus use this product as the starting point for our assessments. It is provided on the 25 km Equal-Area Scalable

Earth Grid (EASE) 2.0 grid, with per-pixel uncertainty ranges, which correspond to the random component of the uncertainty specified for the SIFB product (for details see *Hendricks et al.*, 2016, and references therein). It is crucial for the uncertainty estimate to include any source of uncertainty in the penetration depth of the radar signal. Recall, that we assume uncertainties to be uncorrelated in space. With two satellites in orbit, S3 provides a higher sampling frequency than CryoSat-2. Assuming 4 times more independent measurements than CryoSat-2 over the Arctic, we use the CryoSat-2 RFB uncertainty divided by a factor of $\sqrt{(1/4)} = 0.5$ as uncertainty estimate for an April S3 RFB product with a pole hole north of 81.5N.

The QND study by *Kaminski et al.* (2018) has also indicated good complementarity with SND products. We hence constructed three hypothetical SND products for April 2015. The first one is intended to look like an SND product to be expected from the Copernicus HPCM CRISTAL. We assume an orbit like that of CryoSat-2 with a pole hole north of 88N and retrievals on the 25 km EASE grid wherever simulated SIC exceeds 70%. Following *Lawrence et al.* (2018) we assume an uncertainty of 8 cm for the monthly product. The second SND product is intended to look like an SND product to be expected from the Copernicus HPCM CIMR. We assume retrievals on the 25 km EASE grid wherever simulated SIC exceeds 70%. Following *Braakmann-Folgmann and Donlon* (2019) (including the information provided by the L-Band radiometer) we assume an uncertainty of 4 cm for the monthly composite. The third SND product is intended to look like an SND product to be expected from a modelling approach that combines an ice drift product with output from a numerical weather prediction model to calculate the temporal evolution of SND. We assume the product to be available on the 25 km EASE grid wherever simulated SIC exceeds 1%. Following *Petty et al.* (2018) we assume an uncertainty of 8 cm for the monthly composite product.

A CRISTAL SND product would be derived from the difference of two freeboard measurements onboard the same platform. An alternative is the direct assimilation of two freeboard products into the model. For this purpose we constructed an LFB product that looks like an ICESat 2 product. Following *Farrell et al.* (2009) we assume an uncertainty of 2 cm for the monthly composite product.

A further observable we evaluate is SST (see Figure 2.15). Our first product is derived from the product retrieved by OSI-SAF from infrared measurements (*Eastwood*, 2011a), which cannot penetrate clouds. The OSI-SAF product is available every 12 hours on a 5 km grid, and we assume a composite for the first week of April is available for each MPIOM grid cell north of 40N with simulated SST > -1 K, for which at least 25% of the potential samples are available. For this composite we assume an uncertainty of 0.5 K. Our second product is intended to look like an SST product to be expected from the Copernicus HPCM CIMR. We assume a composite for the first seven days in April has an uncertainty of 1 K and is available on a 12.5 km EASE grid, for each grid cell north of 40N with simulated SST > -1 K.

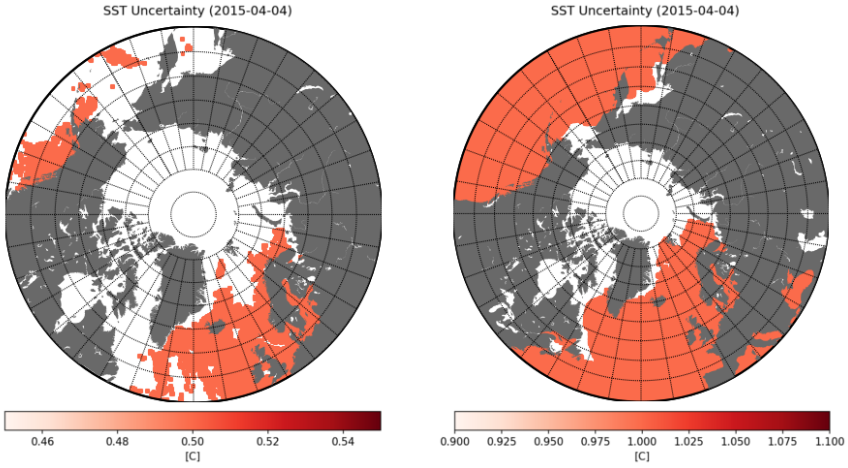


Figure 2.15: Uncertainty ranges $[K]$ and spatial coverage of SST products for April 1-7, 2015: Derived from OSI-SAF products based in infrared measurements (left), a CIMR-like product (right)

Table 2.4: Overview on data sets used for assessments.

Product	Grid	Coverage	Uncertainty
RFB (CS-2)	25 km EASE	product	product
RFB (S3)	25 km EASE	as RFB(CS-2), < 81.5N (pole hole)	0.5 * RFB(CS-2)
SND (CRISTAL)	25 km EASE	SIC > 70%, < 88N (pole hole)	0.08 m
SND (CIMR)	25 km EASE	SIC > 70%	0.04 m
SND (reanalysis)	25 km EASE	SIC > 1%,	0.08 m
LFB (ICESat-2)	25 km EASE	SIC > 70%	0.02 m
SST (IR)	MPIOM grid	SST > -1 C, > 25% sampled, > 40N	0.5 K
SST (CIMR)	12.5 km EASE	SST > -1 C, > 40N	1.0 K

2.2.6 CO_2 Data sets and observation operators

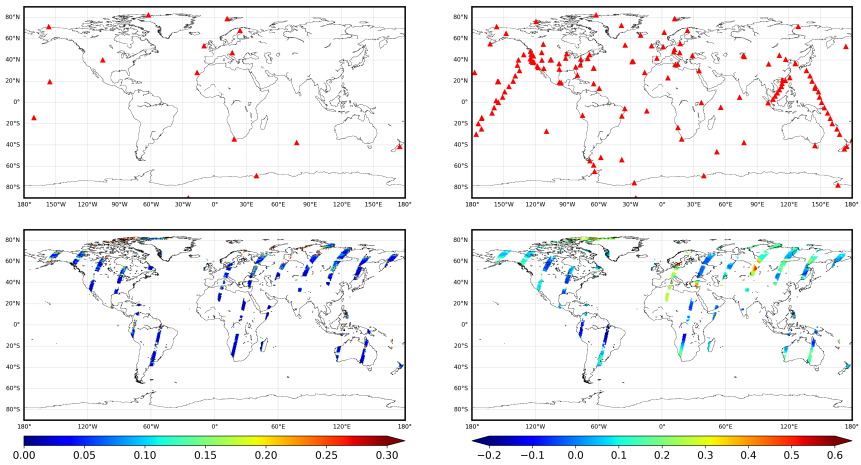


Figure 2.16: Coverage of Atmospheric CO_2 data sets. Networks of 15 (top left) and 141 (top right) sites providing continuous samples and simulated random (bottom left) systematic (bottom right) error of one Copernicus HPCM CO2M satellite for June 1, 2008, aggregated to a 0.5° grid.

Table 2.5: Overview on atmospheric CO_2 data sets used for assessments.

Description	Coverage	Uncertainty
surface network 15 sites	see Figure 2.16, top left	1 ppm
surface network 141 sites	see Figure 2.16, top right	1 ppm
surface network 15 sites plus radiocarbon	see Figure 2.16, top left	0.5 ppm
surface network 15 sites	see Figure 2.16, top left	0.5 ppm
one CO2M satellite	see Figure 2.16, bottom	see figure

An overview of the atmospheric CO_2 data we use for our assessments is provided in Table 2.5. The starting point is the in situ network consisting of the monitoring sites providing continuous samples to the Jena CarboScope system *Rödenbeck* (2005). To explore the effect of an extended network providing continuous samples, we assume those were available at all 141 flask sampling locations of NOAA's Earth System Research Laboratory (www.esrl.noaa.gov/gmd/ccgg/trends/). For both networks we use an uncertainty of 1 ppm for the daily means. Next, we equip the 15 site network with the capability to also measure radiocarbon, a tracer that separates fossil fuel from natural carbon emissions (*Levin et al.*, 2003; *Rayner et al.*, 2010) To explore the sensitivity to the observational uncertainty, we also conduct an assessment for the network of 15 sites with an observational uncertainty of 0.5 ppm.

2 Approach

Simulated random and systematic errors for the Copernicus HPCM CO2M were provided by the University of Bremen. These were produced by applying error parameterisation formulas (EPFs) (*Buchwitz et al.*, 2013) for each individual satellite footprint (ground pixel) based on variables characterising the viewing and illumination geometry (e.g. solar zenith angle), surface (e.g. surface albedo) and atmospheric properties (e.g. cirrus and aerosol optical depth). A swath width of 250 km is assumed. Random and systematic errors were then gridded to 0.5° grid resolution for each day in 2008 (typically) by averaging the relevant values of all individual pixels of the image contained in a 0.5° grid cell for a given day.

Our observation operator for XCO_2 is the global atmospheric transport model TM3 (*Heimann and Körner*, 2003), operated with ≈4°× 5° horizontal resolution on 19 vertical levels. Similar to *Kaminski et al.* (2010), we use pre-computed response functions (observational Jacobians) that resolve monthly flux impacts in the same month as the observation and up to three months before an observation on the full model grid. Monthly flux impacts between four and 47 months before an observation are assumed to be uniform within a set of twelve latitudinal bands, each about 16° wide. For fluxes more than 47 months before an observation, we use a globally uniform response. The model is run with meteorological driving fields taken from the NCEP reanalysis (*Kalnay et al.*, 1996). To simulate observations by the station network we sample the daily mean at the grid cell and vertical model transport layer in which the respective site is located. XCO_2 corresponding to CO2M observations is calculated as the column-average of CO_2 with vertical weighting depending only on layer mass (corresponding to a uniform averaging kernel).

2.3 Target and observational Jacobians

The evaluation of Equation (2.1) and Equation (2.3) requires a target Jacobian $\mathbf{N}'$ for each target quantity and an observational Jacobian $\mathbf{M}'$ for each of the observational products we assess (Table 2.4). This subsection first describes the computation of these Jacobians and then discusses them. For a given product, the observational Jacobian is computed in two steps. The first step performs the following actions: a reference run is performed using the prior control vector x, the input variables to the observation operator are stored over the observational period, aggregated to the model grid, and the observation operator is applied to derive the observation equivalent $M(x)$ on the space-time grid of the observational product. In the second step, for each component of the control vector the following procedure is applied: A sensitivity run is performed with a control vector $x + p_i$ that is identical to the prior control vector but with the i-th component changed by a perturbation ϵ_i, and an observation equivalent $M(x+p_i)$ is computed in the same way as for the reference run. The Jacobian column is then computed as $\sigma_i(M(x + p_i) - M(x))/\epsilon_i$ where σ_i is the prior uncertainty of x_i. As a consequence of the normalisation by the prior uncertainty, each row in the Jacobian has the same unit as the respective observation. For a given product, column i of the corresponding observational Jacobian quantifies the sensitivity of the model-simulated equivalent to that product with respect to a change of the i component of the control vector x_i by one standard deviation (see Table 2.1 for the value).

For any given product the dimension of the observational Jacobian is the product of the dimension of the control space and the grid size of the observational product. As an example, Figure 2.17 displays the row of the Jacobians for April means of SIT, SIFB, RFB, LFB, and SND over a single point in space indicated by the black dot (and by the black cross on Figure 2.3).

The SIT sensitivity is dominated by the model's initial SIT in control region 6 (black bars in Figure 2.17) but shows also considerable sensitivities to the initial SIC, the initial SND, the initial ocean temperature (TEMP) and the zonal wind stress (WIX). The negative sensitivity to SIC in that region is caused by two mechanisms. The first mechanism is expressed by Equation (2.11): The observation $h_{i,eff}$ is the effective SIT (thickness averaged over the ice-covered grid cell) and is reduced if the initial SIC is increased (and vice versa) because the model conserves the total sea ice volume. The second mechanism is related to sea ice growth which depends on the open water fraction, i.e. more (less) sea ice can grow if the SIC is reduced (increased). The small negative sensitivity of SIT to SND is caused by the strong insulation effect of snow, which hampers the growth of sea ice (or fosters the growth if SND is reduced). The physical process behind the small negative sensitivities on the initial ocean temperature needs no further explanation; we recall, however, that, in the presence of sea ice, the control variable relates to a temperature change below the second model layer (at 17m depth). The negative sensitivity with respect to the zonal wind stress (WIX) mirrors less advection of thick sea ice originating from the Beaufort Gyre. WIX is positive for eastward wind stress. A positive sensitivity to WIX is most distinct in region 6 (but also evident in regions 7 and 8) and slows down the Beaufort Gyre which advects less

sea ice into the target region (sea ice behaves, at least in April and May, to a large extent like a rigid body, i.e. a change in regions 7 and 8 impacts almost instantaneously on the target regions) resulting in a negative sensitivity. The SIT sensitivities on model parameters (Figure 2.17 are very small compared to the sensitivities on the initial state or the atmospheric boundary conditions, as the short integration time (we sample the April mean of a model simulation starting on April 1) restricts the impact of the parameters.

The various freeboard products exhibit high sensitivity to initial SIT and SND (orange, red, and green bars in Figure 2.17 and enlarged in Figure 2.17). As SIT enters all freeboard observation operators in the same way (Section 2.2.5), the freeboard sensitivity to April mean SIT is equal for all products, which also renders their sensitivity to initial SIT almost equal. The LFB sensitivity on the initial SND is positive (LFB is the freeboard at the top of the snow layer) while the sensitivity of the RFB and SIFB is negative because an increased SND will reduce the RFB and SIFB through the increased weight on the ice floe (see Figure 2.12). Due to the definition of the observation operator for RFB (Equation (2.15)) its sensitivity to initial SND is larger than that of the SIFB (Section 2.2.5). The sensitivity of the freeboard products (yellow, red, and green bars) with respect to the parameters of the sea ice and ocean model is low. The impact of the sea ice density on the respective observation operators (Equation (2.13) to Equation (2.16)) is high, though, while sensitivity with respect to the snow density is much lower (because the sea ice thickness is much larger than the SND at the observational point). The SND shows only considerable sensitivity to the initial SND in control region 6 and some minor positive sensitivity with respect to the precipitation in the same region.

For the visualisation of the observational Jacobian for SST observations the point we have selected for the freeboard and SND observations is not suitable, because SST observations are only made over open water. Figure 2.18 shows the observational Jacobian for an SST observation (mean over the first seven days in April) west of Greenland. The constraint by this observation close to the ice edge is confined to the process parameters and the initial and boundary conditions over Region 9, in which the point is located. The process parameters with the highest impact are hiccpp, hibcc, and cw. We note that the wind direction is such that it pushes the ice away from the observational location close to the west coast of Greenland, and recall the formulation (*Hibler*, 1979) of the ice strength in Equation (2.6). A reduction in the ice strength parameter hicpp (pstar) increases the mobility of the ice. An increase of parameter hibcc (cstar) has the same effect. Higher ice mobility yields higher SST at the observational location. Increased drag cw also pushes the ice further west. A reduction in initial SIT, SIC, or SD reduces the insulating effect and yields to increased warming in region 9. We recall that initial temperature and salinity refer to the entire water column. While the impact of initial temperature is obvious, the change in initial salinity impacts the stratification of the water column, and the resulting geostrophic circulation changes SST. The impact of atmospheric temperature is obvious. An increase of the wind field components WIX and WIY increases the wind component from the Northeast, which hampers the inflow of warm

water from the north Atlantic and thus reduces SST.

Likewise we computed target Jacobians $\mathbf{N}'$ for each of the six target quantities (SIV and SNV each over 5 regions) described in Section 2.2.2. Each target quantity is a scalar and thus the Jacobian has one entry for each component of the control vector. As an example Figure 2.19 displays the Jacobians for SIV and SNV (on May 28) over the Outer New Siberian Islands (ONSI) region. The first point to note is that sensitivities of regional SIV and SNV to the control vector differ, so an observation must constrain different components of the control vector to perform well for one or the other target quantity.

SIV over the ONSI region is highly sensitive to initial SIT over control regions 5 and 6 (Figure 2.20) which at least partly overlap with the target area. As the SIT observation and due to the same mechanisms discussed above, the SIV target quantity also exhibits a negative sensitivity to the initial SIC, SND, and zonal wind stress. It is interesting to note that SIV is also sensitive to initial and boundary conditions over more remote control regions. For example, it exhibits a positive sensitivity to the initial SIT in the control regions 1 and 7 from which thick sea ice is advected into the target region during the period from April 1 to May 28. This also explains the negative sensitivity to the zonal wind stress in region 7 and the meridional wind stress in region 1: For sufficiently high SIC the sea ice almost behaves as an incompressible fluid allowing even for a sensitivity to wind stress changes in very remote control regions, e.g. the negative sensitivity to the zonal wind stress in region 8. The positive sensitivity to the zonal wind stress in region 1 (with thick ice) may be less obvious, as it follows the deflection of ice drift by about 20° to the right. The largest SIV sensitivity to model parameters (Figure 2.20) is found for the snow albedo of freezing conditions (albsn), but still that sensitivity is low compared to the sensitivity with respect to the initial state and atmospheric boundary conditions.

SNV shows particularly high sensitivity to the initial SND but also considerable sensitivity with respect to the precipitation and air temperature in region 6. The largest model parameter sensitivity is found for the snow albedo for melting conditions: Increasing the snow albedo will reduce the melting.

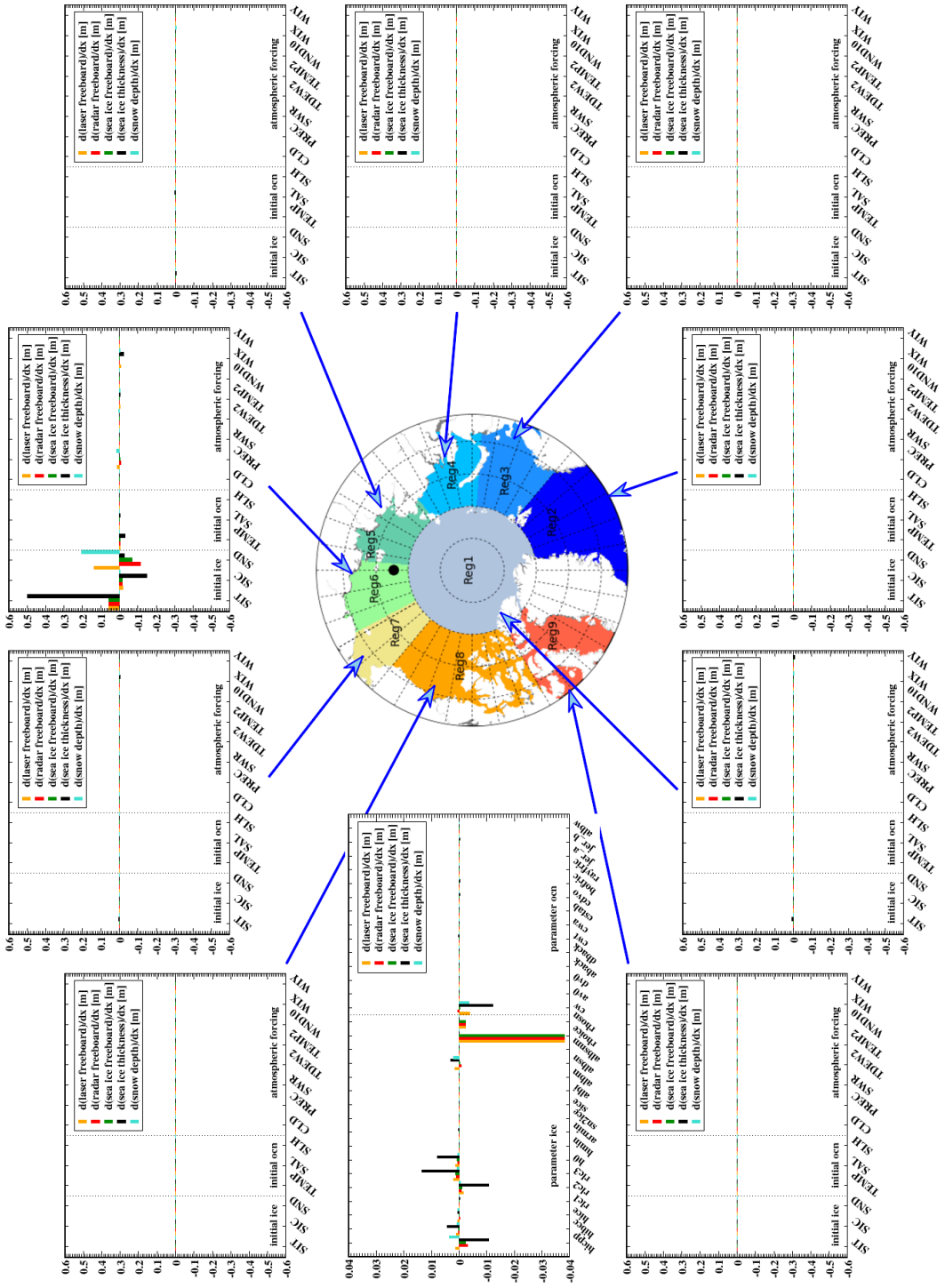


Figure 2.17: The sensitivities of the respective EO product to the control vector ("observational Jacobians") for April means of LFB (orange bars), RFB (red bars), SIFB (green bars), SIT (black bars) and SND (cyan bars) over a single point indicated by the black dot (and by black cross on Figure 2.3). The observational Jacobians with respect to the process parameters are shown in the left middle panel. The other panels show the observational Jacobians with respect to the initial and forcing fields for each control region (see Table 2.1 for an explanation of the abbreviations).

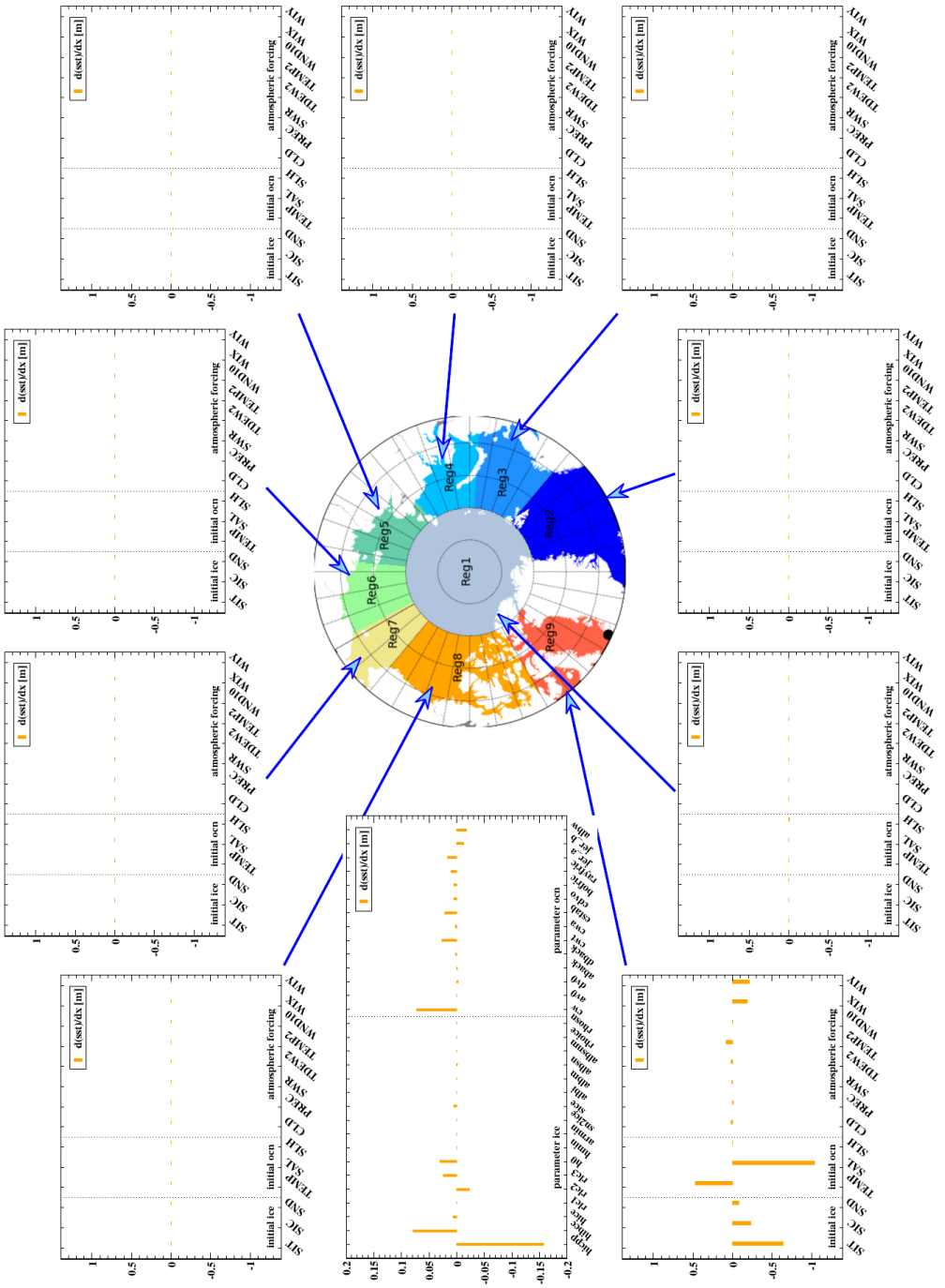


Figure 2.18: The sensitivities of SST to the control vector ("observational Jacobians") for a mean SST over the first seven days in April over a single point indicated by the black dot (and by black cross on Figure 2.3). The observational Jacobian with respect to the process parameters are shown in the left middle panel. The other panels show the observational Jacobian with respect to the initial and forcing fields for each control region (see Table 2.1 for an explanation of the abbreviations).

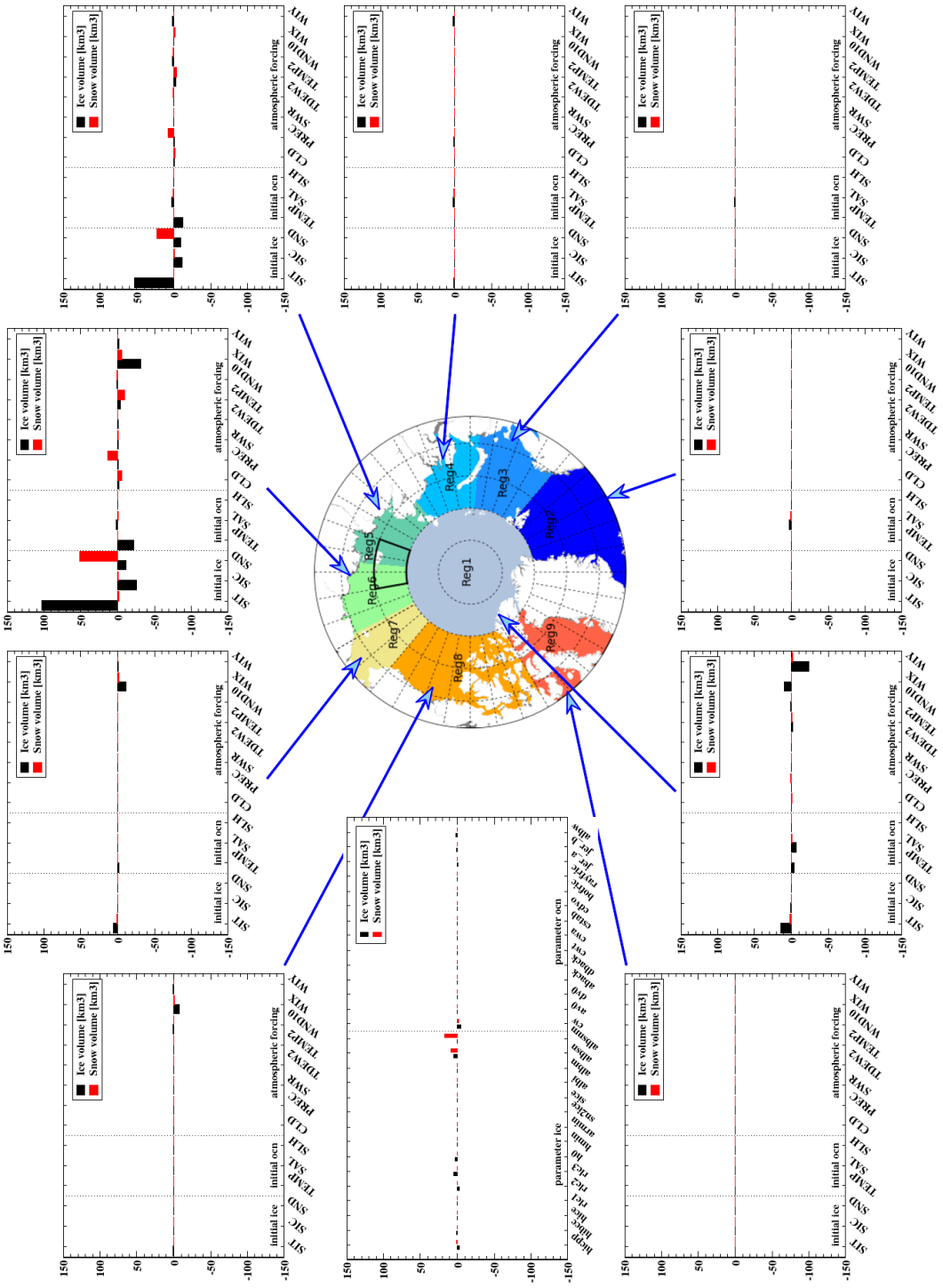


Figure 2.19: As Figure 2.17 but for the sensitivities of the sea ice (SIV) and snow (SNV) volume over the target region Outer New Siberian Islands (ONSI) on the control vector ("target Jacobians").

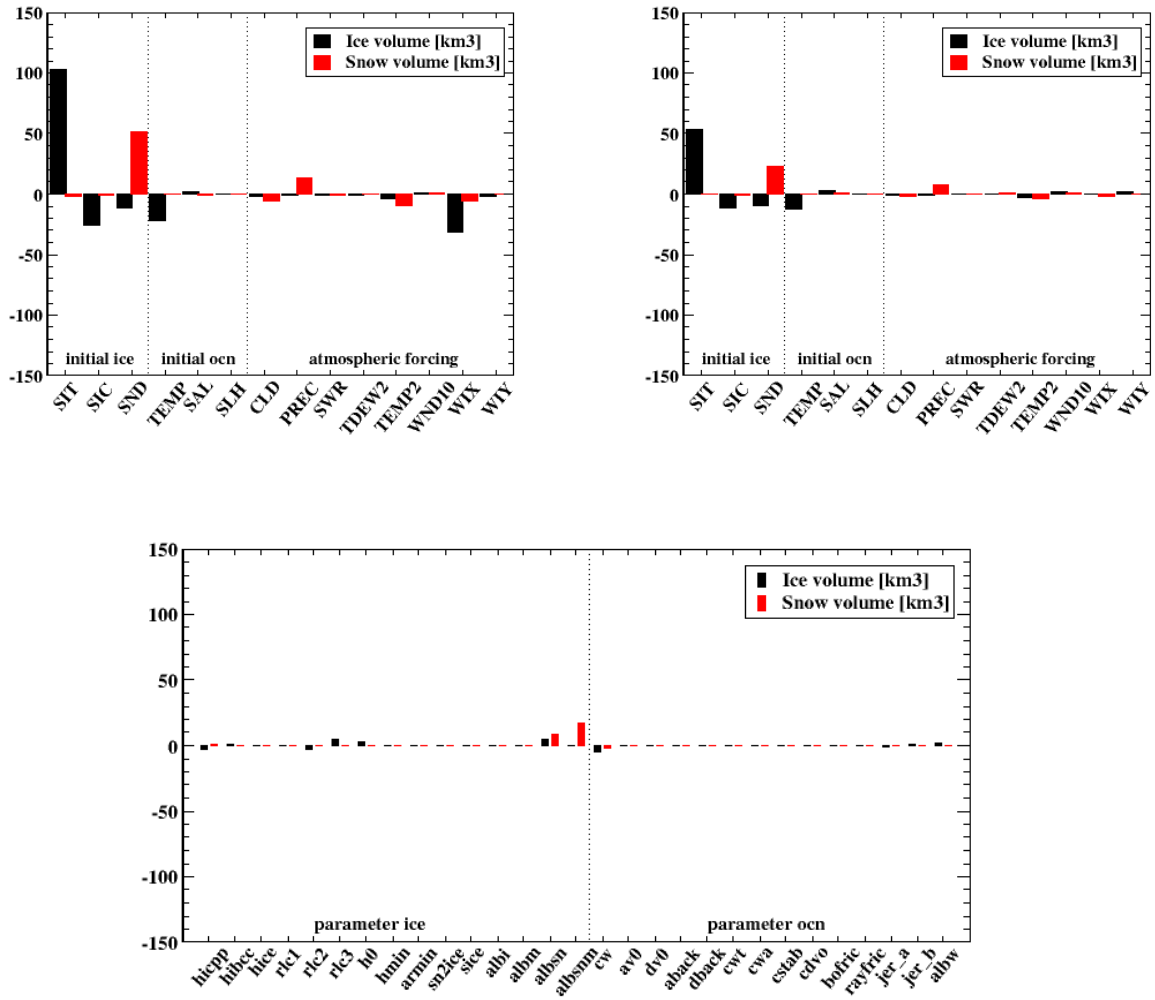


Figure 2.20: An excerpt of Figure 2.19 for the target Jacobian (top) in target region 6 ("East Siberian Sea" - left) and region 5 ("Laptev Sea" - right) and (bottom) for the model parameters.

3 Experimental setup

For the sea ice-ocean system, our assessments are based on the products shown in Table 2.4. We conducted assessments for the eight cases listed in Table 4.1 in terms of their impact on forecasts of SIV and SNV for the target regions listed in Section 2.2.2.

First we compare the performances of the RFB product from CryoSat-2 and of an S3-like RFB product (rows 4 and 5 of Tables 4.1 and 4.2)). Next, we assess the combination of the CryoSat-2 RFB product with three SND products described in Section 2.2.5:

1. CIMR-like,
2. CRISTAL-like,
3. Reanalysis-like

and contrast this with the combination of CryoSat-2 RFB with an ICESat-2 like LFB product. For this set of assessments we prefer the CryoSat-2 product, because (unlike our hypothetical S3 RFB product) the uncertainties are derived from real observations. Next, we compare two SST products, one CIMR like product and one product derived from infrared measurements. The reference for these assessments is a case without observations. Row three ("prior") shows the uncertainties in the target quantities that result from the prior uncertainty in the control vector. As mentioned in Section 2.2.2, our study year is 2015, and we have two different assimilation windows. For freeboard/ice thickness/snow depth products we assimilate an average field over April, followed by a four-week forecast. For SST we assimilate an average field over the first week of April, followed by a seven-week forecast.

Another set of assessments addresses the constraint of different combinations of atmospheric CO_2 observations (Table 2.5) on land-based fossil fuel emissions in selected Arctic countries. Our starting point is a network of 15 sites providing continuous in situ samples. We then explore three ways of extending/improving this network. The first is by equipping all sites of the NOAA ESRL network (currently providing flask samples) by continuous in situ analysers. The second is by reducing the observational uncertainty of the 15-site network from 1 ppm to 0.5 ppm. The third is by combining the 15-site network with one CO2M satellite. In summary we perform four assessments of atmospheric CO_2:

1. 15-site in situ surface continuous sampling network
2. 141-site in situ surface continuous sampling network
3. 15-site in situ surface continuous sampling network with lower uncertainty
4. 15-site in situ surface continuous sampling network + one CO2M satellite

We analyse the fossil fuel emissions for the period from June 1 to June 7, 2008 using observations for the same period.

4 Discussion of Assessment Results

4.1 Sea ice and Ocean Observations

This section presents the results of our assessments of sea ice and ocean observations. As explained in Section 2.2.1 the uncertainty component from the model error $\sigma(y_{mod})$ in Equation (2.3) covers the residual uncertainty that remains with an optimal control vector, i.e. it reflects uncertainty from uncertain aspects not included in the model error and structural uncertainty reflecting wrong or missing process formulations. $\sigma(y_{mod})$ is model-dependent and is probably the most subjective component in the prior and posterior uncertainties. $\sigma(y_{mod})$ acts as an offset (for the respective variance) for all cases, and reduces the contrast between the cases. As our assessments focus on the differences between the cases, we exclude it from the target uncertainties in Table 4.1 and Table 4.2 and provide an estimate in the bottom row. This estimate (denoted by $\sigma_{\text{mod, relative}}$) assumes a model that simulates the same SIV and SNV as our model with an uncertainty of 10% for SIV and 30% for SNV. We use a higher uncertainty for SNV because it has a stronger dependence on the surface forcing (mainly precipitation), for which the temporal and small-scale spatial structures are not resolved in the control vector.

We stress that, owing to the exclusion of model error (and biases in the observations) for a given observational scenario our results (Table 4.1 and Table 4.2) show the upper limit for the impact it can achieve. For example we assume that the penetration of radar and laser signals is treated correctly. Also, as our reference is a prior case without any observations, the impact of an observational data stream is much higher than in a setup where it is added to an assimilation system that already assimilates a variety of other data sets, as it is the case for the Copernicus systems described in Chapter 1.

Both RFB products (CryoSat 2 and Sentinel 3) provide strong constraints on forecasts of SIV and SNV (see also Figure 4.1). S3 RFB performs slightly better, at least for our target regions and variables. This means the reduced product uncertainty (owing to the higher sampling frequency) of S3 overcompensates for the larger pole hole. This does not only hold for the five target regions along the NSR and Northwestern Sea Route but also for the Arctic as a whole (last column in Table 4.1 and Table 4.2).

Adding SND products to the CS-2 RFB product achieves a significant performance gain. Among the SND products, the CIMR-like product performs better than the reanalysis-based product, i.e. the better precision overcompensates for the lower spatial coverage. The reanalysis-based product in turn performs slightly better than the CRISTAL-like product, owing to the better spatial coverage. Using instead of the SND products an ICESat-2 like LFB product (in combination with CS-2 RFB) yields the best performance. This can be attributed to two effects: The first is the assumed

Table 4.1: Prior and posterior uncertainties of sea ice volume (SIV, columns 3-8) respectively for five regions and the entire Arctic in km^3. Columns 1 and 2 indicate products. The bottom row gives an estimate for the uncertainty due to model error, i.e. the residual uncertainty with optimal control vector.

Product	Product	BAB	BFS	WLS	ONSI	ESS	ARCTIC
prior	-	113.7	327.9	136.5	131.6	289.6	2276.0
RFB CS-2	-	37.9	65.5	51.3	39.2	93.8	419.7
RFB S3	-	23.1	44.1	34.0	28.1	65.2	335.6
RFB CS-2	SND CIMR	11.8	27.4	4.3	5.9	19.7	111.3
RFB CS-2	SND CRISTAL	15.7	37.8	6.4	8.5	27.8	147.2
RFB CS-2	SND Reanalysis	14.6	36.1	6.2	8.0	25.9	128.6
RFB CS-2	LFB ICESat	5.7	13.1	2.1	2.7	9.1	63.3
SST CIMR	-	63.1	275.1	129.8	122.5	264.6	1414.8
SST IR	-	113.0	304.8	135.3	129.5	285.2	2171.7
$\sigma_{\text{mod, relative}}$	-	18.2	81.4	48.7	70.8	165.9	1896.6

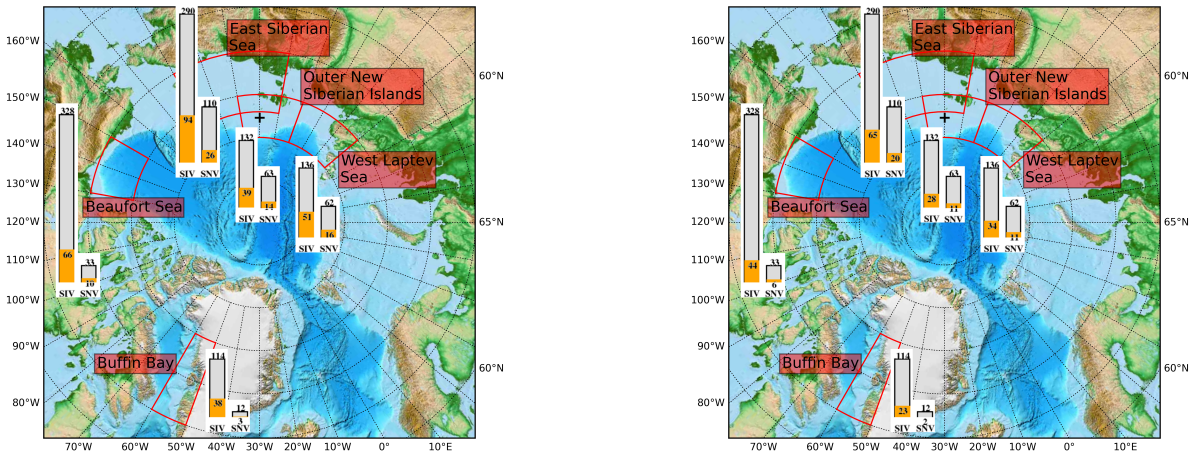


Figure 4.1: Assessment Results for RFB from CryoSat 2 and Sentinel 3. Prior and posterior uncertainty ranges [km^3]

higher precision (2 cm) compared to that in the SND products (4-8 cm). The other effect is that LFB has a different sensitivity to the control vector than SND (see Figure 2.17). To quantify the second effect, we compare with a previous study (*Kaminski et al.*, 2017a) that has evaluated the CS-2 RFB product in combination with an SND product (the same coverage and accuracy as our LFB product) for our target quantities along the NSR. For both SIV and SNV the performance was about 30% weaker than for our assessment here.

Table 4.2: Prior and posterior uncertainties of snow volume (SNV, columns 3-8) respectively for five regions and the entire Arctic in km^3. Columns 1 and 2 indicate products. The bottom row gives an estimate for the uncertainty due to model error, i.e. the residual uncertainty with optimal control vector.

Product	Product	BAB	BFS	WLS	ONSI	ESS	ARCTIC
prior	-	12.2	33.5	62.3	63.3	110.1	916.7
RFB CS-2	-	2.8	9.7	16.4	14.2	26.0	165.3
RFB S3	-	1.7	6.5	11.3	10.6	19.5	119.6
RFB CS-2	SND CIMR	0.6	2.5	4.0	4.1	7.9	45.6
RFB CS-2	SND CRISTAL	1.0	4.2	6.2	6.4	12.5	71.7
RFB CS-2	SND Reanalysis	0.9	3.9	6.0	6.2	11.9	67.2
RFB CS-2	LFB ICESat	0.4	1.5	1.6	1.7	3.3	23.1
SST CIMR	-	11.9	33.1	62.1	62.9	109.4	867.0
SST IR	-	12.2	33.4	62.3	63.2	110.0	911.5
$\sigma_{\text{mod, relative}}$	-	0.2	1.1	10.2	11.4	5.3	506.3

Table 4.3: Uncertainty reduction relative to prior (see Equation (2.5)) of sea ice volume respectively for five regions and the entire Arctic in %. Columns 1 and 2 indicate products.

Product	Product	BAB	BFS	WLS	ONSI	ESS	ARCTIC
RFB CS-2	-	66.7	80.0	62.4	70.2	67.6	81.6
RFB S3	-	79.7	86.5	75.1	78.6	77.5	85.3
RFB CS-2	SND CIMR	89.7	91.6	96.8	95.5	93.2	95.1
RFB CS-2	SND CRISTAL	86.2	88.5	95.3	93.5	90.4	93.5
RFB CS-2	SND Reanalysis	87.2	89.0	95.5	93.9	91.1	94.3
RFB CS-2	LFB ICESat	95.0	96.0	98.4	98.0	96.9	97.2
SST CIMR	-	44.5	16.1	4.9	6.9	8.7	37.8
SST IR	-	0.6	7.0	0.9	1.6	1.5	4.6

SST observations provide almost no constraint on SNV, because of the insulating effect from the sea ice below the snow. The SST constraint on SIV is small, except for the target region BAB, which is closest to the ice edge in the observation period. The slight constraint on the other target regions is due to constraints of the SST observation on the global process parameters (Figure 2.18). CIMR outperforms IR in BAB, because it provides much better coverage close to the region.

Differences between target regions in the performance of the same product are the result of a complex interplay of the Jacobians $\mathbf{N}'$ for the target regions and the product's constraint on the control vector quantified by $\mathbf{C}(x)$ (see Equation (2.3)). For each of the target regions a different (combination) of control regions is most relevant: For WLS this is control region 5 (not shown), for ONSI these are control regions 5 and 6 (Figure 2.19 and enlarged in Figure 2.20) and for ESS these are control regions 6 and 7 (not shown). BFS is different because, in contrast to the NSR regions, it contains a

Table 4.4: Uncertainty reduction relative to prior (see Equation (2.5)) of snow volume respectively for five regions and the entire Arctic in %. Columns 1 and 2 indicate products.

Product	Product	BAB	BFS	WLS	ONSI	ESS	ARCTIC
RFB CS-2	-	76.6	71.1	73.6	77.6	76.4	82.0
RFB S3	-	86.0	80.6	81.9	83.3	82.2	86.9
RFB CS-2	SND CIMR	94.8	92.5	93.6	93.5	92.9	95.0
RFB CS-2	SND CRISTAL	91.5	87.4	90.1	89.9	88.7	92.2
RFB CS-2	SND Reanalysis	92.2	88.5	90.3	90.2	89.1	92.7
RFB CS-2	LFB ICESat	96.9	95.6	97.4	97.4	97.0	97.5
SST CIMR	-	1.9	1.2	0.3	0.6	0.6	5.4
SST IR	-	0.0	0.0	0.1	0.1	0.1	0.6

large fraction of multi-year ice. The particularity of BFS is the large fraction of open water. The ability of a product to constrain a particular control region is determined by the combination of the observational Jacobian of the product and the product uncertainty (see Equation (2.1)).

4.2 Atmospheric CO_2

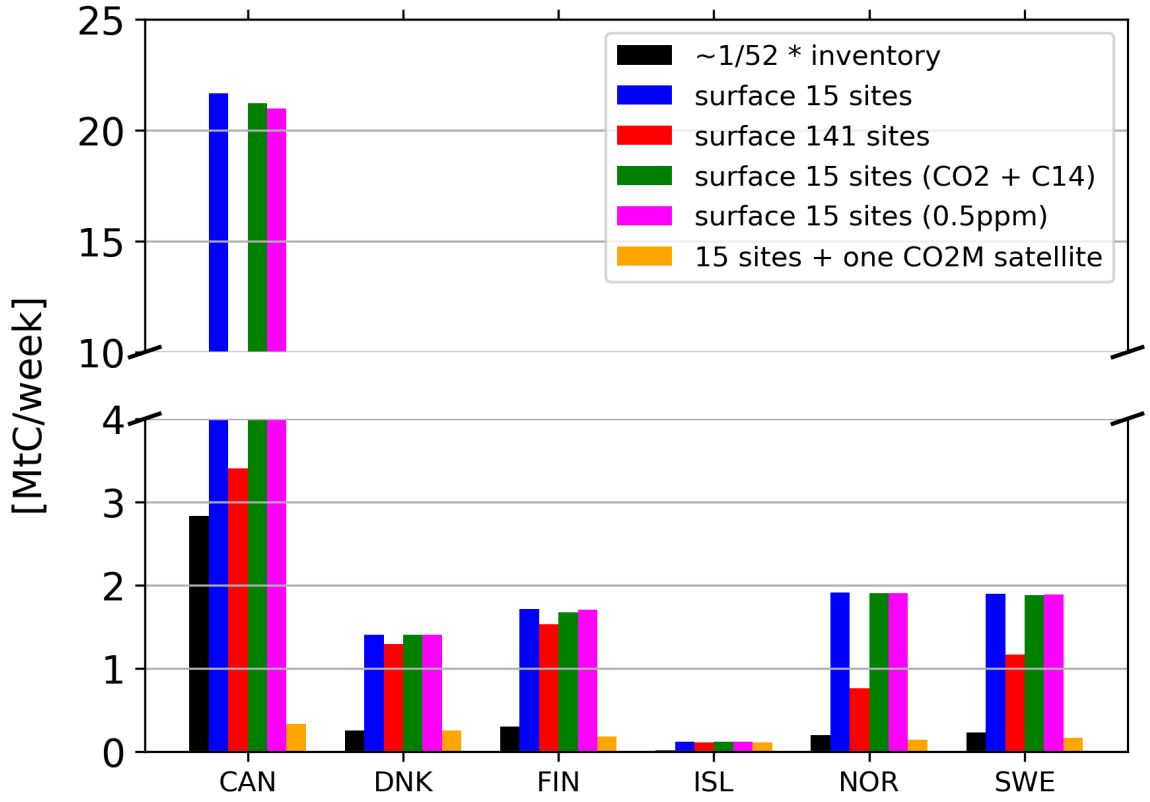


Figure 4.2: Posterior uncertainty (including atmospheric and other observational constraints) of fossil fuel emission rates during first week of June 2008 for selected countries shown by their ISO codes in MtC/week (blue, red, green, violet, and orange bars). 1/52 of the 2008 fossil fuel emission from inventories (back bars).

Figure 4.2 quantifies the impact of the four observational scenarios defined in Section 3 on the posterior uncertainty estimates of total land-based fossil fuel emissions:

1. 15-site in situ surface continuous sampling network (blue bars)
2. 141-site in situ surface continuous sampling network (red bars)
3. 15-site in situ surface continuous sampling network also providing radiocarbon samples (green bars)
4. 15-site in situ surface continuous sampling network with lower uncertainty (violet bars)

5. 15-site in situ surface continuous sampling network + one CO2M satellite (orange bars)

The impact is quantified through the posterior uncertainties for the fossil fuel emissions of selected Arctic countries over the period from June 1 to June 7, 2008. There is typically no information from inventories on this time scale, which we could use as a reference. Instead, as a reference, we show (black bars) the average 7-day fraction (i.e. 7/366) of the respective annual fossil fuel emissions provided by the IEA (*International Energy Agency*, 2011) for the (leap) year 2008.

For the 15-site network scenario, the posterior uncertainty is well above the reference for all countries. Compared to this scenario, the extension to the 141-site scenario achieves a substantial uncertainty reduction over Canada, Norway and Sweden, and a modest uncertainty reduction over Denmark and Finland, and a small uncertainty reduction (about 3%) over Island. Adding radiocarbon observations to the 15-site network or reducing the observational uncertainty for the 15-site network by a factor of two has very little effect. These scenarios are inferior to the 141-site scenario. The scenario that combines the 15-site network with one CO2M satellite achieves a much better performance than the 141-site network, except over Island, where CO2M performs only about 1% better than the 141-site surface network. Posterior uncertainties are well below the reference (average 7-day fraction of annual total) for Canada, and below the reference for Finland, Norway, and Sweden. They are just below the reference for Denmark, and well above the reference for Island. Small emissions are generally more difficult to detect with atmospheric measurements than large emissions. Also, with our relatively coarse (global scale) atmospheric transport simulation we tend to underestimate the performance over spatially small countries such as Island. In addition, small isolated countries may also suffer more from cloud contamination. We note, however, that the measurements by CO2M rely on reflected sunlight, i.e., unlike the in situ surface network, it will provide much lower coverage in winter.

Bibliography

Asefi-Najafabady, S., P. J. Rayner, K. R. Gurney, A. McRobert, Y. Song, K. Coltin, J. Huang, C. Elvidge, and K. Baugh (2014), A multiyear, global gridded fossil fuel co2 emission data product: Evaluation and analysis of results, *Journal of Geophysical Research: Atmospheres*, *119*(17), 10,213–10,231, doi:10.1002/2013JD021296. 21, 22

Bhaduri, B., E. Bright, P. Coleman, and M. Urban (2007), LandScan USA: A high resolution geospatial and temporal modeling approach for population distribution and dynamics, *GeoJournal*, *69*, 103–117. 21

Bovensmann, H., H. Bösch, D. Brunner, P. Ciais, D. Crisp, H. Dolman, G. Hayman, S. Houweling, and L. Lichtenberg (2015), Report for mission selection: Carbonsat - an earth explorer to observe greenhouse gases, *Tech. rep.*, European Space Agency, Noordwijk, The Netherlands, Noordwijk, The Netherlands, freely available via Official URL link. 3

Braakmann-Folgmann, A., and C. Donlon (2019), Estimating snow depth on arctic sea ice using satellite microwave radiometry and a neural network, *The Cryosphere*, *13*(9), 2421–2438, doi:10.5194/tc-13-2421-2019. 30

Buchwitz, M., M. Reuter, H. Bovensmann, D. Pillai, J. Heymann, O. Schneising, V. Rozanov, T. Krings, J. P. Burrows, H. Boesch, C. Gerbig, Y. Meijer, and A. Löscher (2013), Carbon monitoring satellite (carbonsat): assessment of atmospheric co_2 and ch_4 retrieval errors by error parameterization, *Atmospheric Measurement Techniques*, *6*(12), 3477–3500, doi:10.5194/amt-6-3477-2013. 33

Buitenhuis, E. T., R. B. Rivkin, S. Sailley, and C. Le Quéré (2010), Biogeochemical fluxes through microzooplankton, *Global Biogeochemical Cycles*, *24*(4), n/a–n/a, doi:10.1029/2009GB003601, gB4015. 24

Chevallier, M., G. C. Smith, F. Dupont, J.-F. Lemieux, G. Forget, Y. Fujii, F. Hernandez, R. Msadek, K. A. Peterson, A. Storto, T. Toyoda, M. Valdivieso, G. Vernieres, H. Zuo, M. Balmaseda, Y.-S. Chang, N. Ferry, G. Garric, K. Haines, S. Keeley, R. M. Kovach, T. Kuragano, S. Masina, Y. Tang, H. Tsujino, and X. Wang (2017), Intercomparison of the arctic sea ice cover in global ocean–sea ice reanalyses from the ora-ip project, *Climate Dynamics*, *49*(3), 1107–1136, doi:10.1007/s00382-016-2985-y. 16

Chuvieco, E., C. Yue, A. Heil, F. Mouillot, I. Alonso-Canas, M. Padilla, J. M. Pereira, D. Oom, and K. Tansey (2016), A new global burned area product for climate assessment of fire impacts, *Global Ecology and Biogeography*, *25*(5), 619–629, doi:10.1111/geb.12440. 24

Bibliography

Crisp, D., R. Atlas, F.-M. Breon, L. Brown, J. Burrows, P. Ciais, B. Connor, S. Doney, I. Fung, D. Jacob, C. Miller, D. O'Brien, S. Pawson, J. Randerson, P. Rayner, R. Salawitch, S. Sander, B. Sen, G. Stephens, P. Tans, G. Toon, P. Wennberg, S. Wofsy, Y. Yung, Z. Kuang, B. Chudasama, G. Sprague, B. Weiss, R. Pollock, D. Kenyon, and S. Schroll (2004), The orbiting carbon observatory (oco) mission. trace constituents in the troposphere and lower stratosphere, *Advances in Space Research*, *34*(4), 700 – 709, doi:10.1016/j.asr.2003.08.062. 3

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The era-interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597, doi:10.1002/qj.828. 12

Eastwood, S. (2011a), Atlantic high latitude l3 sea surface temperature product user manual version 2.1, *Technical Report*, Norwegian Meteorological Institutes, Oslo, Norway. 30

Eastwood, S., M. Jenssen, T. Lavergne, A. Sorensen, and Tonboe (2015), Eumetsat ocean and sea ice satellite application facility. global sea ice concentration reprocessing product (v1.2), *Product user manual, Technical Report, Norwegian and Danish Meteorological Institutes, Oslo, Norway and Copenhagen, Denmark*, *14*, 2079–2087. 13

Farrell, S. L., S. W. Laxon, D. C. McAdoo, D. Yi, and H. J. Zwally (2009), Five years of arctic sea ice freeboard measurements from the ice, cloud and land elevation satellite, *Journal of Geophysical Research: Oceans*, *114*(C4), doi:10.1029/2008JC005074. 30

Feng, L., P. I. Palmer, H. Bösch, and S. Dance (2009), Estimating surface co_2 fluxes from space-borne co_2 dry air mole fraction observations using an ensemble kalman filter, *Atmospheric Chemistry and Physics*, *9*(8), 2619–2633, doi:10.5194/acp-9-2619-2009. 3

Friedl, M., A. Strahler, and J. Hodges (2010), ISLSCP II MODIS (Collection 4) IGBP Land Cover, 2000-2001, doi:10.3334/ORNLDAAC/968. 22

Gent, P., and J. McWilliams (1990), Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, *20*, 150–155, doi:10.1175/1520-0485(1990)020⟨0150:IMIOCM⟩2.0.CO;2. 11

Gent, P., J. Willebrand, T. McDougall, and J. McWilliams (1995), Parameterizing eddy-induced tracer transport in ocean circulation models, *J. Phys. Oceanogr.*, *25*, 463–474, doi:10.1016/j.ocemod.2010.08.002. 11

Giglio, L., J. T. Randerson, and G. R. van der Werf (2013), Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (gfed4), *Journal of Geophysical Research: Biogeosciences*, *118*(1), 317–328, doi:10.1002/jgrg.20042. 24

Giorgetta, M. A., J. Jungclaus, C. H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, K. Glushak, V. Gayler, H. Haak, H.-D. Hollweg, T. Ilyina, S. Kinne, L. Kornblueh, D. Matei, T. Mauritsen, U. Mikolajewicz, W. Mueller, D. Notz, F. Pithan, T. Raddatz, S. Rast, R. Redler, E. Roeckner, H. Schmidt, R. Schnur, J. Segschneider, K. D. Six, M. Stockhause, C. Timmreck, J. Wegner, H. Widmann, K.-H. Wieners, M. Claussen, J. Marotzke, and B. Stevens (2013), Climate and carbon cycle changes from 1850 to 2100 in mpi-esm simulations for the coupled model intercomparison project phase 5, *Journal of Advances in Modeling Earth Systems*, *5*(3), 572–597, doi:10.1002/jame.20038. 11

Griffies, S. M. (1998), The gent-mcwilliams skew flux, *J. Phys. Oceanogr.*, *28*, 831–841, doi:10.1175/1520-0485(1998)028⟨0831:TGMSF⟩2.0.CO;2. 11

Guerrier, D., and F. Horley (1970), *Archimedes: Archimedes' Principle and the Law of Flotation*, Discovering with the Scientists Series, Blond & Briggs. 28

Haak, H., J. Jungclaus, U. Mikolajewicz, and M. Latif (2003), Formation and propagation of great salinity anomalies, *Geophysical Research Letters*, *30*(9), n/a–n/a, doi:10.1029/2003GL017065, 1473. 11

Haas, C., J. Beckers, J. King, A. Silis, J. Stroeve, J. Wilkinson, B. Notenboom, A. Schweiger, and S. Hendricks (2017), Ice and snow thickness variability and change in the high arctic ocean observed by in situ measurements, *Geophysical Research Letters*, *44*(20), 10,462–10,469, doi:10.1002/2017GL075434. 15

Hardt, M., and F. Scherbaum (1994), The design of optimum networks for aftershock recordings, *Geophys. J. Int.*, *117*, 716–726, doi:10.1111/j.1365-246X.1994.tb02464.x. 3

Harris, I., P. Jones, T. Osborn, and D. Lister (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, *International Journal of Climatology*, *34*(3), 623–642, doi:10.1002/joc.3711. 24

Heimann, M., and S. Körner (2003), The global atmospheric tracer model TM3, *Tech. Rep. 5*, Max-Planck-Institut für Biogeochemie, Jena, Germany. 33

Hendricks, S., R. Ricker, and V. Helm (2016), User guide - AWI CryoSat-2 sea ice thickness data product (v1.2). 28, 29, 30

Hibler, W. (1979), A dynamic thermodynamic sea ice model, *Journal Geophysical Research*, *9*, 815–846, doi:10.1175/1520-0485(1979)009⟨0815:ADTSIM⟩2.0.CO;2. 11, 35

Houweling, S., F.-M. Breon, I. Aben, C. Rödenbeck, M. Gloor, M. Heimann, and P. Ciais (2004), Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, *Atmos. Chem. Phys.*, *4*, 523–538, doi:10.5194/acp-4-523-2004. 3

Hungershoefer, K., F.-M. Breon, P. Peylin, F. Chevallier, P. Rayner, A. Klonecki, S. Houweling, and J. Marshall (2010), Evaluation of various observing systems for the global monitoring of co_2 surface fluxes, *Atmospheric Chemistry and Physics*, *10*(21), 10,503–10,520, doi:10.5194/acp-10-10503-2010. 3

International Energy Agency (2011), CO_2 emissions from fuel combustion, OECD/IEA, Paris, 1997, available at http://www.iea.org/. 22, 48

Jungclaus, J., M. Giorgetta, C. Reick, S. Legutke, V. Brovkin, T. Crueger, M. Esch, K. Fieg, N. Fischer, K. Glushak, V. Gayler, H. Haak, H. Hollweg, S. Kinne, L. Kornblueh, D. Matei, T. Mauritsen, U. Mikolajewicz, W. Müller, D. Notz, T. Pohlmann, T. Raddatz, S. Rast, E. Roeckner, M. Salzmann, H. Schmidt, R. Schnur, J. Segschneider, K. Six, M. Stockhause, J. Wegner, H. Widmann, K. Wieners, M. Claussen, J. Marotzke, and B. Stevens (2012), CMIP5 simulations of the Max Planck Institute for Meteorology (MPI-M) based on the MPI-ESM-P model: The 1pctCO2 experiment, *served by ESGF, WDCC at DKRZ*, doi:10.1594/WDCC/CMIP5.MXEPc1. 11

Jungclaus, J. H., N. Keenlyside, M. Botzet, H. Haak, J.-J. Luo, M. Latif, J. Marotzke, U. Mikolajewicz, and E. Roeckner (2006), Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, *Journal of Climate*, *19*(16), 3952–3972, doi:10.1175/JCLI3827.1. 11

Jungclaus, J. H., N. Fischer, H. Haak, K. Lohmann, J. Marotzke, D. Matei, U. Mikolajewicz, D. Notz, and J. von Storch (2013), Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth system model, *J. Adv. Model. Earth Syst.*, *5*, 422–446, doi:10.1002/jame.20023. 11

Kadygrov, N., S. Maksyutov, N. Eguchi, T. Aoki, T. Nakazawa, T. Yokota, and G. Inoue (2009), Role of simulated gosat total column co2 observations in surface co2 flux uncertainty reduction, *Journal of Geophysical Research: Atmospheres*, *114*(D21), n/a–n/a, doi:10.1029/2008JD011597. 3

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, et al. (1996), The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, *77*(3), 437–471. 33

Kaminski, T., and P.-P. Mathieu (2017), Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the earth system and the carbon cycle, *Biogeosciences*, *14*(9), 2343–2357, doi:10.5194/bg-14-2343-2017. 5, 26

Kaminski, T., and P. J. Rayner (2017), Reviews and syntheses: guiding the evolution of the observing system for the carbon cycle through quantitative network design, *Biogeosciences*, *14*(20), 4755–4766, doi:10.5194/bg-14-4755-2017. 3, 4

Kaminski, T., W. Knorr, P. Rayner, and M. Heimann (2002), Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, *Global Biogeochemical Cycles*, *16*(4), 14–1–14–16. 22, 24

Kaminski, T., M. Scholze, and S. Houweling (2010), Quantifying the Benefit of A-SCOPE Data for Reducing Uncertainties in Terrestrial Carbon Fluxes in CCDAS, *Tellus B*, doi:10.1111/j.1600-0889.2010.00483.x. 3, 33

Kaminski, T., P. J. Rayner, M. Voßbeck, M. Scholze, and E. Koffi (2012), Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, *Atmospheric Chemistry and Physics*, *12*(16), 7867–7879, doi:10.5194/acp-12-7867-2012. 3, 4

Kaminski, T., F. Kauker, H. Eicken, and M. Karcher (2015), Exploring the utility of quantitative network design in evaluating arctic sea ice thickness sampling strategies, *The Cryosphere*, *9*(4), 1721–1733, doi:10.5194/tc-9-1721-2015. 3, 16

Kaminski, T., M. Scholze, M. Voßbeck, W. Knorr, M. Buchwitz, and M. Reuter (2017a), Constraining a terrestrial biosphere model with remotely sensed atmospheric carbon dioxide, *Remote Sensing of Environment*, pp. 109 – 124, doi:http://dx.doi.org/10.1016/j.rse.2017.08.017. 22, 24, 25, 44

Kaminski, T., B. Pinty, M. Voßbeck, M. Lopatka, N. Gobron, and M. Robustelli (2017b), Consistent retrieval of land surface radiation products from eo, including traceable uncertainty estimates, *Biogeosciences*, *14*(9), 2527–2541, doi:10.5194/bg-14-2527-2017. 24

Kaminski, T., F. Kauker, L. Toudal Pedersen, M. Voßbeck, H. Haak, L. Niederdrenk, S. Hendricks, R. Ricker, M. Karcher, H. Eicken, and O. Gråbak (2018), Arctic mission benefit analysis: impact of sea ice thickness, freeboard, and snow depth products on sea ice forecast performance, *The Cryosphere*, *12*(8), 2569–2594, doi:10.5194/tc-12-2569-2018. 2, 3, 26, 29, 30

Knorr, W. (2000), Annual and interannual CO_2 exchanges of the terrestrial biosphere: Process-based simulations and uncertainties, *Glob. Ecol. Biogeogr.*, *9*(3), 225–252. 24

Knorr, W., and M. Heimann (1995), Impact of drought stress and other factors on seasonal land biosphere CO_2 exchange studied through an atmospheric tracer transport model, *Tellus, Ser. B*, *47*(4), 471–489. 22

Kurtz, N., and S. Farrell (2011), Large-scale surveys of snow depth on arctic sea ice from operation icebridge, *Geophysical Research Letters*, *38*(L20505), 10,462–10,469, doi:10.1029/2011GL049216. 15

Kwok, R., and G. F. Cunningham (2008), Icesat over arctic sea ice: Estimation of snow depth and ice thickness, *Journal of Geophysical Research: Oceans*, *113*(C8), n/a–n/a, doi:10.1029/2008JC004753, c08010. 15

Lawrence, I. R., M. C. Tsamados, J. C. Stroeve, T. W. K. Armitage, and A. L. Ridout (2018), Estimating snow depth over arctic sea ice from calibrated dual-frequency radar freeboards, *The Cryosphere*, *12*(11), 3551–3564, doi:10.5194/tc-12-3551-2018. 30

Levin, I., B. Kromer, M. Schmidt, and H. Sartorius (2003), A novel approach for independent budgeting of fossil fuel co_2 over europe by $^{14}co_2$ observations, *Geophysical Research Letters*, *30*(23), doi:10.1029/2003GL018477. 32

Lindsay, R., and A. Schweiger (2015), Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, *The Cryosphere*, *9*(1), 269–283, doi:10.5194/tc-9-269-2015. 13, 14, 15

Mikolajewicz, U., D. Sein, D. Jacob, T. Königk, R. Podzun, and T. Semmler (2015), Simulating arctic sea ice variability with a coupled regional atmosphere-ocean-sea ice model, *Meteorol. Zeitschrift*, *14*, 793–800, doi:10.1127/0941-2948/2005/0083. 12

Müller, W. A., J. Baehr, H. Haak, J. H. Jungclaus, J. Kröger, D. Matei, D. Notz, H. Pohlmann, J. S. von Storch, and J. Marotzke (2012), Forecast skill of multi-year seasonal means in the decadal prediction system of the max planck institute for meteorology, *Geophysical Research Letters*, *39*(22), n/a–n/a, doi:10.1029/2012GL053326, l22707. 11

Nakicenovic, N. (2004), Socioeconomic driving forces of emissions scenarios, in *The global carbon cycle: Integrating humans, climate, and the natural world*, edited by C. B. Field and M. R. Raupach, pp. 225–239, Island Press, Washington, D.C. 21

Niederdrenk, A. (2013), The arctic hydrologic cycle and its variability in a regional coupled climate model, *PhD Thesis, Unversity Hamburg*, pp. 1–186. 11, 12

Niederdrenk, A. L., D. V. Sein, and U. Mikolajewicz (2016), Interannual variability of the arctic freshwater cycle in the second half of the twentieth century in a regionally coupled climate model, *Climate Dynamics*, *47*(12), 3883–3900, doi:10.1007/s00382-016-3047-1. 12

NOAA (1988), Data announcement 88-mgg-02, digital relief of the surface of the earth, *NOAA, National Geophysical Data Center, Boulder, Colorado*. 12

Notz, D., F. A. Haumann, H. Haak, J. H. Jungclaus, and J. Marotzke (2013), Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's earth system model, *Journal of Advances in Modeling Earth Systems*, *5*(2), 173–194, doi:10.1002/jame.20016. 13

Pacanowski, R., and S. Philander (1981), Parameterization of vertical mixing in numerical-models of tropical oceans, *J. Phys. Oceanogr.*, *11*, 1443–1451, doi:10.1175/1520-0485(1981)011⟨1443:POVMIN⟩2.0.CO;2. 11

Patra, P. K., S. Maksyutov, Y. Sasano, H. Nakajima, G. Inoue, and T. Nakazawa (2003), An evaluation of co2 observations with solar occultation fts for inclined-orbit satellite sensor for surface source inversion, *Journal of Geophysical Research: Atmospheres*, *108*(D24), n/a–n/a, doi:10.1029/2003JD003661. 3

Petty, A. A., M. Webster, L. Boisvert, and T. Markus (2018), The nasa eulerian snow on sea ice model (nesosim) v1.0: initial model development and analysis, *Geoscientific Model Development*, *11*(11), 4577–4602, doi:10.5194/gmd-11-4577-2018. 30

Pinty, B., T. Lavergne, M. Voßbeck, T. Kaminski, O. Aussedat, R. Giering, N. Gobron, M. Taberner, M. M. Verstraete, and J.-L. Widlowski (2007), Retrieving surface parameters for climate models from MODIS-MISR albedo products, *J. Geophys. Res.*, *112*, doi:http://doi.acm.org/10.1029/2006JD008105. 24

Pinty, B., M. Clerici, I. Andredakis, T. Kaminski, M. Taberner, M. M. Verstraete, N. Gobron, S. Plummer, and J.-L. Widlowski (2011), Exploiting the modis albedos with the two-stream inversion package (jrc-tip): 2. fractions of transmitted and absorbed fluxes in the vegetation and soil layers, *Journal of Geophysical Research – Atmospheres*, *116*, D09,106, doi:10.1029/2010JD015373. 24

Rayner, P., M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann (2005), Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS), *Global Biogeochemical Cycles*, *19*, doi:10.1029/2004GB002,254. 3

Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO_2 concentration data in surface source inversions, *Geophys. Res. Let.*, *28*, 175–178, doi:10.1029/2000GL011912. 3

Rayner, P. J., I. G. Enting, and C. M. Trudinger (1996), Optimizing the CO_2 observing network for constraining sources and sinks, *Tellus*, *48B*, 433–444, doi:1680-7324/acp/2004-4-413. 3

Rayner, P. J., M. R. Raupach, M. Paget, P. Peylin, and E. Koffi (2010), A new global gridded data set of CO_2 emissions from fossil fuel combustion: Methodology and evaluation, *Journal of Geophysical Research: Atmospheres*, *115*(D19), n/a–n/a, doi:10.1029/2009JD013439, d19306. 32

Rayner, P. J., S. R. Utembe, and S. Crowell (2014), Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, *Atmospheric Measurement Techniques*, *7*(10), 3285–3293, doi:10.5194/amt-7-3285-2014. 3

Rayner, P. J., A. M. Michalak, and F. Chevallier (2019), Fundamentals of data assimilation applied to biogeochemistry, *Atmospheric Chemistry and Physics*, *19*(22), 13,911–13,932, doi:10.5194/acp-19-13911-2019. 3, 4

Redi, M. H. (1982), Oceanic isopycnal mixing by coordinate rotation, *J. Phys. Oceanogr.*, *12*, 1154–1158, doi:10.1175/1520-0485(1982)012⟨1154:OIMBCR⟩2.0.CO;2. 11

Ricker, R., S. Hendricks, V. Helm, H. Skourup, and M. Davidson (2014), Sensitivity of cryosat-2 arctic sea-ice freeboard and thickness on radar-waveform interpretation, *The Cryosphere*, *8*, 1607–1622, doi:10.5194/tc-8-1607-2014. 26, 28

Rödenbeck, C. (2005), Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, *Technical Report 6*, Max Planck Institute for Biogeochemistry, Jena, Germany. 8, 32

Röske, F. (2001), An atlas of surface fluxes based on the ecmwf reanalysis—a climatological data set to force global ocean general circulation models, *Tech. Rep. 323, Max-Planck-Inst. für Meteorol., Hamburg, Germany.* 12

Semtner, A. (1976), A model for the thermodynamic growth of sea ice in numerical investigastions of climate, *Journal pf Physical Oceanography*, *6*, 379–389, doi:10.1175/1520-0485(1976)006⟨0379:AMFTTG⟩2.0.CO;2. 11

Steele, M., R. Morley, and W. W. Ermold (2001), Phc: A global ocean hydrography with a high quality arctic ocean, *J. Climate*, *14*, 2079–2087, doi:10.1175/1520-0442(2001)014⟨2079:PAGOHW⟩2.0.CO;2. 12

Stocker, T., D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley (2013), Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change [stocker, t.f., qin, d., plattner, g.-k., tignor, m., allen, s.k., boschung, j., nauels, a., xia, y., bex, v., and midgley, p.m. (eds.)], *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp*, doi:10.1017/CBO9781107415324. 11

Takahashi, T., R. H. Wanninkhof, R. A. Feely, R. F. Weiss, D. W. Chipman, N. Bates, J. Olafsson, C. Sabine, and S. C. Sutherland (1999), Net sea-air CO_2 flux over the global oceans: An improved estimate based on the sea-air pCO_2 difference, in *Extended abstracts of the 2nd International CO_2 in the Oceans Symposium, Tsukuba, Japan, January 18-22, 1999*, pp. 18–01. 24

Tarantola, A. (2005), *Inverse Problem Theory and methods for model parameter estimation*, SIAM, Philadelphia, doi:10.1137/1.9780898717921. 3

Warren, S., I. Rigor, N. Untersteiner, V. Radionov, N. Bryaz-gin, Y. Aleksandrov, and R. Colony (1999), Snow depth on arctic sea ice, *J. Clim.*, *12*, 1814–1829, doi: 10.1175/1520-0442(1999)012⟨1814:SDOASI⟩2.0.CO;2. 15