NERSC Technical Report no. 233
Under contract 14992/01/NL/MM with
European Space Agency (ESA)

Study of the impact and relevance of ESAs missions in operational Oceanography and Climate research and monitoring

SIREOC
Draft Final Report

Nansen Environmental and Remote Sensing Center, Norway

Bergen, March 2003

EUROPEAN SPACE AGENCY
CONTRACT REPORT

The work described in this report was done under ESA contract. Responsibility for the content resides in the author or organisation that prepared it.

Nansen Environmental and Remote Sensing Center (NERSC)
Edvard Griegsvei 3a
N-5059 Bergen
Norway

phone +47 55297288
fax +47 55200050
email admin@nersc.no

TITLE	**REPORT IDENTIFICATION**
Study of the impact and relevance of ESAs missions in operational Oceanography and Climate research and monitoring	NERSC Technical report no.233
CLIENT	**CONTRACT NO.**
European Space Agency (ESA)	14992/01/NL/MM
CLIENT REFERENCE	**AVAILABILITY**
M. Drinkwater (EOP–SM0) technical matters M. Mezzadri (IMT–CAO) contractual matters	Open
INVESTIGATORS	**AUTHORIZATION**
M. Bentsen, L. Bertino, K. Brusdal, H. Drange, G. Evensen, T. Furevik, J.A. Johannessen, K.A. Lisæter, K. Lygre, L.J. Natvik, H. Sagen, A.B. Sandø	Bergen, September 23, 2003 Prof. Ola M. Johannessen Director NERSC

ESA STUDY CONTRACT REPORT

	SUBJECT: Study of the impact and relevance of ESAs mssions in operational Oceangraphy and climate research and montoring SIREOC Final Report		CONTRACTOR: Nansen Environmental and Remote Sensing Center
ESA CR( ) No	STAR CODE:	NO. OF VOLUMES: 1 THIS VOLUME: 1	CONTRACTOR's REFERENCE:
The main focus of this study contract is to assess and quantify the relative impact of different Earth Observation data types for climate research and monitoring and for operational ocean prediction systems. The Impact has been examined in light of availability of satellite observations of physical oceanographic variables, sea ice variables and marine ecosystem variables. Data from exising satellites including ERS–1 and ERS–2, TOPEX/POSEIDON, NOAA TIROS, DMSP and SeaSTAR has been used. In addition data from Envisat and JASON–1 has been explored, while the impact of Cryosat, SMOS and GOCE has been undertaken using simulated data.			
NAME OF AUTHORS: M. Bentsen, L. Bertino, K. Brusdal, H. Drange, G. Evensen, T. Furevik, J.A. Johannessen, K.A. Lisæter, K. Lygre, L.J. Natvik, H. Sagen, A.B. Sandø			
ESA STUDY MANAGERS: M. Drinkwater (EOP–SMO) technical matters M. Mezzadri (IMT–CAO) contractual matters			ESA BUDGET HEADING

EXECUTIVE SUMMARY

The provision of observational data, both satellite and in-situ based, and their integration into appropriate models of the Earth system are of paramount importance for monitoring the state of the environment and the climate, as required by national and international programmes, obligations and treaties. Well-documented examples are ozone depletion, atmospheric chemistry and transport of pollutants, El Nino Southern Oscillation, sea level and global ocean circulation, Arctic sea ice cover, and ice sheet elevation. In addition, the products and services delivered from numerical weather prediction centres as well as from the growing number of operational and pre-operational oceanographic monitoring and modelling systems all require access to integrated data systems. The EuroGOOS Space Panel report (*Font et al.*, 2001), the EuroGOOS operational oceanography data requirements survey, the IGOS Ocean Theme report (http://www.igospartners.org), and the implementation and operations of EUMETSAT Satellite Application Facilities (i.e. CM-SAF, OSI-SAF), the Argo profiling float program, the ESA Living Planet program (*SP-1227*, 1998) as well as the recent joint EU/ESA Global Monitoring for Environment and Security (GMES) initiative are fully compliant with this integrated view. Similarly are the requirements for integrated observations outlined in the plans for the Global Ocean Data Assimilation Experiment (GODAE; http://www.bom.gov.au/bmrc/ocean/GODAE) and the complementing strategy for Observing the Oceans in the 21st century.

The joint EU and ESA's GMES initiative, for instance, recognises that satellite-based multi-disciplinary Earth observation is a fundamental component of monitoring systems. In the context of the marine environment, integrated observing systems and numerical models are capable of producing a large range of information products (output), including physical (i.e. waves, currents, temperature, etc.), biological (i.e. algae concentration, primary production, etc.) and chemical (oil pollution, etc.) variables. Established users range from the research community, via operational applications to commercial customers and governmental entities. The reliability and utilization of these types of information products depend not only upon the performance of the models and assimilation tools, but also on the availability and quality of the observing systems, telecommunication networks, data processing and distribution, data access, and rapid information integration, flow and services.

In view of these outlooks and initiatives it is therefore clear that the Study of Impact and Relevance of ESAs earth observation for Operational oceanography and Climate research (SIREOC) is highly timely and relevant. The main objective of the SIREOC project (running from 2001 to 2003) has been to assess and quantify the relative impact of different Earth Observation data types of open ocean variables, Arctic sea ice cover and marine ecosystem for a) climate research and monitoring; and b) operational ocean prediction systems. This has been undertaken in two parallel and complementary workpackages, notably:

Climate research and monitoring using the MICOM model system, with particular attention on seasonal via inter-annual to decadal time scales, including trend analyses, such as those characteristic for the equatorial Atlantic, North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO).

Short-term (days to weeks) prediction capabilities including impact of state estimation and control of the evolving model state using the TOPAZ configuration of the NERSC model suite.

The Earth Observation data sources include mostly sea surface temperature, sea surface height, chlorophyll-a, and sea ice concentration and extent from currently operating satellites including ERS-1 and ERS-2, TOPEX/POSEIDON, NOAA TIROS, SeaWiFS and DMSP SSMI. In addition the investigation of the impact of simulated Earth Observation data has been performed to mimic the data flow from the three future ESA Earth Explorer opportunity and core missions, namely Cryosat, GOCE and SMOS. These are respectively planned for launch in 2004, 2006 and 2007 with the aim to obtain sea-ice thickness, marine geoid and ocean salinity measurements.

The NERSC model suite contains a physical ocean model (OGCM), an ecosystem model and a sea ice model. Three classes of input sources are interfaced to this model suite, notably atmospheric forcing fields, remote sensing data and in-situ data. The data assimilation module, which then combines the observation data fields (available in near real time or off-line) and the model data, forms the central element of the suite and can be applied to all three models, either individually or in combination. The system, in turn, is then delivering hindcast, nowcasts and forecasts products. Both global, regional and local model domains can be selected with accurate bathymetry and land boundaries, and based on a grid generation tool the model domain will also always be defined with best model resolution in the area of highest interest. Moreover, the system introduces a further downscaling capability by nesting of high resolution models for the regional and local areas of particular interest.

In the climate change monitoring and research the analysis has been focused on the role SSS play in the Subpolar region, especially on the deep water formation rate, and the subsequent link to the maximum strength of the Atlantic Thermohaline Circulation (THC). Four model integrations for the period 1948-2000 were analyzed giving the following main results:

- The THC shows a marked variability on decadal time scales, and also a significant (increasing) trend over the last three decades. Both the trend and the variability correlate, in general, to the NAO index.
- The SSS play a pivotal role in the intensity of deep convection, in particular in the Labrador and Irminger Seas, and thereby also on the actual strength of the THC. For instance, the freshwater anomalies in the early 70ies and the early to mid 80ies are followed by a rapid decline in the strength of the THC.

Direct validation of these simulation results are difficult as the existence of salinity observations are limited in time and space. However, by comparison of model and ADCP observations of Atlantic water inflow to the Nordic Seas we are confident that the above simulations give a realistic picture. It is therefore clearly anticipated that better knowledge of the SSS would lead to much better description and characterization of the THC and its variability and relation as well as influence on the deep convection rate of the waters in the Greenland and Labrador Seas.

The main results of the operational oceanography study lead us to conclude that the assimilation of SLA and SST data work well for HYCOM with the introduction of z-levels in the upper mixed layer. In fact, when compared to results obtained with the isopycnic coordinate model (MICOM) only, the better representation of the upper layers in HYCOM is clearly revealed in comparison to profiles from the GTSPP data base. It also appears that assimilation of SLA provides a better advection of signal vertically into the upper ocean, which is not the case if only SST data was available.

From hindcast and near real time experiments assimilation of SeaWiFS data demonstrate that the ocean colour data, and in particular the spatial distribution, were very useful for controlling the evolution of the model chlorophyll. As such the assimilation of these data also had a multivariate impact on other biological variables in the ecosystem model.

The assimilation experiments performed with sea-ice thickness and sea-ice concentration data showed improvement as compared to free-run experiments without assimilation. This improvement was most noticeable for the experiments with ice concentration, while the assimilation of ice thickness was somewhat hampered by model bias in the ensemble predictions.

In summary, the SIREOC results have demonstrated that in combination with in-situ data, Earth Observation are necessary both for climate change monitoring and for operational oceanography. In order to improve our understanding of the Earth system continuous access to such observations are needed to secure model development and validation as well as to allow for data assimilation. In the context of operational oceanography, the implementation of the Global Ocean Data Assimilation Experiment (GODAE) is now entering into an operational demonstration phase from 2003 to 2005. This global experiment is complemented and supported by several European regional operational oceanography systems including the French MERCATOR project for the North-East Atlantic, the UK FOAM system, the Mediterranean Forecasting System Pilot Project (MFSPP), and the TOPAZ system which has been extensively used in this SIREOC project.

The main findings of the SIREOC study is currently also extended with the execution of the MERSEA Strand-1 project funded in the initial phase of GMES under FP5 from 2003 to 2004. The main deliverables of this project provided in the context of monitoring for environment and security are to: Deliver reliable, high quality, information products; Report on the problems met and lessons learnt; and Contribute to longer-term improvement of knowledge, methods and tools.

More detailed information on the goals, approach and status of the MERSEA Strand-1 project can be found at http://www.nersc.no/~mersea. In addition it is noteworthy that several of the partners in MERSEA Strand-1 participate in the complementary ROSES project funded under ESA GMES initial phase. Moreover, the partners of MERSEA Strand-1 and ROSES are also members of the consortium that submitted the MERSEA Integrated Proposal targeting the GMES issues under the FP6 call 1.4 Aeronautics and Space. This proposal was selected in June 2003 and is currently undergoing the negotiation for a final approval and kick-off in January 2004.

List of Acronyms	
AMSR	Advanced Microwave Scanning Radiometer
AO	Arctic Oscillation
ARPEGE/IFS	An Atmospheric General Circulation model
ATHC	Atlantic Thermohaline Circulation
BCM	Bergen Climate Model
CERFACS	National Center for modelleing and global change
CLIC	Climate the Cryosphere (WCRP Project)
CLIVAR	Climate Variability and Predictability
CLIMAP	Climate Long-range Investigation and MApping and Prediction
CMIP	Coupled Model Intercomparison Project
DS	Denmark Strait
ECMWF	European Centre for Medium Range Weather Forecasting
EnKF	Ensemble Kalman Filtering
ENSO	El Niouthern Oscillationn
ETOPO-5	Global digital bathymetry maps generated by NOAA, National Geophysical Dataa Center, Boulder, Colorado
EO	Earth Observation
FBC	Faeroe-Bank Channel
FSC	Faeroe-Shetland Channel
GAIM	Global Analysis, Interpretation and Modelling
GCOS	Global Climate Observing System
GCM	General Circulation Model
GEWEX	Global Energy and Water Cycle Experiment
GIS	Greenland-Iceland-Scotland
GLOBEC	IGBP Project to identify the influence of climate change on the abundance, diversity and productivity of marine ecosystems
GLOSS	Global Sea Level Observing System
GOCE	Gravity Field and Steady-State Ocean Circulation Explorer (Earth Explorer Core Mission)
GODAE	Global Ocean Data Assimilation Experiment
GOOS	Global Ocean Observing System
GTSPP	Global Temperature and Salinity Profile Program
HYCOM	Hybride Coordinate Ocean Model
IFR	Iceland Faroe Ridge
IGBP	International Geosphere and Biosphere Programme
IOC	Intergovermental Oceanographic Committee
IPCC	Intergovermental Panel of Climate Change
JGOFS	Joint Global Ocean Flux Study
JRC	Joint Research Centre
MICOM	Miami Isopycnic Coordinate Ocean Model
ML	Mixed Layer
MPI	Max plank Institute
MRD	Mission Requirements Document
NAO	North Atlantic Oscillation
OASIS	Ocean Atmosphere Sea Ice Soil
OGCM	Ocean General Circulation Model
OS	Ocean Salinity
RD	Reference Document
SLA	Sea Level Anomaly
SMOS	Soil Moisture and Ocean Salinity (Earth Explorer Opportunity Mission)
SSH	Sea Surface Height
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
T_B	Brightness Temperature
THC	Thermohaline Circulation
WCRP	World Climate Research Programme

Contents

Chapter 1

Introduction

1.1 Motivation

At the onset of the new millennium the number of Earth Observing satellites are growing rapidly for scientific research and operational application within fields of marine meteorology and oceanography including sea ice covered regions.

As outlined in the Living Planet Program the key modules in an integrated Earth system model are based on combined use of in-situ and satellite observations with modelling and data assimilation tools. This is also a common view and strategy of the International Programme and Observing System which are defined and established for both climate research and monitoring and operational applications. Here we mention the World Climate Research Program (WCRP), International Geopshere and Biosphere Program (IGBP) and Intergovermental Oceanographic Committee (IOC) under which CLIVAR, GEWEX, CLIC, GCOS, GOOS, EuroGOOS, JGOFS, GAIM, GLOSS and GLOBEC all highlight the importance of continuous and regular access to Earth Observation data.

Thus, in combination with in-situ data, Earth Observation data have gradually become an important contribution to improve our understanding of the Earth system and the associated model development and validation. The skill of model representations of the marine climate and environmental systems has therefore improved dramatically over the last decade. The increased computational power and improved parameterizations of non-modeled processes have also contributed to this, and so has data assimilation. In the context of climate research global and regional predictions are now offered on seasonal , interannual and decadal scales, while the implementation of the Global Ocean Data Assimilation Experiment (GODAE) is under preparation with the aim to go into an operational demonstration phase from 2003 to 2005. European operational oceanography systems in support to GODAE include the French MERCATOR project for the Atlantic Ocean, the Mediterranean Forecasting System Pilot Project (MFSPP), the UK FOAM global system and the DIADEM/TOPAZ systems for the North Atlantic - Nordic and Arctic Seas and European coastal zones.

The time is therefore right to undertake a project, such as SIREOC, to investigate the impact and relevance of satellite data in both climate research and operational oceanography. The motivation behind this project is based on this plus the fact that models and assimilation methods (notably DIADEM and TOPAZ) required to carry out such a study have been implemented and undergone gradual validation at the Nansen Environmental and Remote Sensing Center during the past five years.

It is also clear that the conclusions and recommendations from SIREOC have relevance for ESA's Ocean Watch programme preparation as well as the Global Monitoring for Environment and Security (GMES). The latter is an initiative set up jointly by the European Commission and the European Space Agency to establish by 2008 a European capacity of Global Monitoring for Environment and Security (GMES). It is an integral part of the European Strategy for Sustainable Development approved at the Gothenborg Summit in June 2001 by the Heads of State and Government of the countries of the European Union. The commitment to sustainable development, now embedded in

all Community policies, requires a much enhanced information basis and one of a different nature. The need for policy decisions to address the environmental, social and economic issues simultaneously, in their interactions and in a long term perspective necessitates tailored indicators backed by wide-ranging high quality observations and validated models. The European Union needs independent information to play its part on the global scene, either as a party to international conventions or to implement and develop European policies. Rapidly developing policies, such as these related to environment and security, present new information requirements. The interplay between human activities and the environment needs to be assessed at different territorial levels, from the local to the global, which means that information can ”zoom” in and out. The increasing influence of human activities on the Earth System as well as the exposure to natural or technological hazards require rapid reactions.

Thus the mission and challenge of GMES is to contribute to the timely provision of such information necessary to enable all society agents, each in their own capacity, to take the decisions and actions which will make sustainable development become a reality. In so doing GMES needs to: (a) objectively identify the combination of causes that results in the rather low efficiency of the current European capacity to produce policy relevant information; and (b) specifically propose solutions for a European GMES to be operational and efficient by 2008.

1.2 Objective and workplan

The main objective of this project will be to assess and quantify the relative impact of different Earth Observation data types for a) climate research and monitoring; and b) operational ocean prediction systems. The impact will be examined in light of availability of satellite observations of physical oceanographic variables, sea ice variables and marine ecosystem variables. In so doing the workplan are conducted with specific focus on climate change in WP2000 and operational oceanography in WP3000.

WP 2000 - Climate Research and Monitoring

The overall goal of this workpackage which focus on climate research and monitoring, with particular attention on seasonal via inter-annual to decadal time scales, including trend analyses, such as those characteristic for the equatorial Atlantic, North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO), is to examine the effect of applying available and planned (simulated) remotely sensed data sets to simulate the mean state and variability in a state-of-the-art coupled ocean-sea ice model.

To meet this overall goal, three tasks shall be carried out including:

Task 2100: impact assessment,

Task 2200: identification of improvement,

Task 2300: skill assessment.

WP 3000 - Operational Oceanography

The main focus of this workpackage on operational oceanography associated with short-term (weeks to months) prediction capabilities including impact of state estimation and control of the evolving model state, where the DIADEM/TOPAZ configuration of the NERSC model suite is supplied with Earth Observation of open ocean variables, Arctic sea ice cover and marine ecosystem data sources, is to assess and quantify the relative impact of different EO data types in an operational prediction system for physical oceanographic variables, sea ice variables and the marine ecosystem.

To meet this goal the following tasks shall be carried out:

Task 3100: open ocean physical variables,

Task 3200: Arctic ice cover,

Task 3300: marine ecosystem.

1.3 Key input data sources

Based on available satellite data and other existing observation fields the following key input data sources will be established.

- Synoptic (daily) atmospheric forcing fields (surface humidity, pressure, vector wind including scatterometer observations, temperature and humidity, precipitation, short wave radiance, and cloudiness).
- Existing and analysed gridded satellite data fields of sea surface temperature (SST), sea level anomalies (SLA), ocean color and sea ice extent and concentration.
- Simulated (or constructed) satellite data fields of sea surface salinity (SSS), sea-ice thickness, and absolute dynamic topography (derived from combined use of altimetry and the marine geoid).
- Monthly mean SST fields from the Integrated Global Ocean Services System Products Bulletin which is accessible from (http://ingrid.ldeo.columbia.edu/SOURCES/IGOSS/).

In particular the following specifics are given for the various data sets:

Combined TOPEX/POSEIDON and ERS-2 data set produced by the CLS Space Oceanography Division as part of the European Union' Environment and Climate project DUACS (ENV4-CT96-0357). This data set is available on a 0.25° by 0.25° grid, at 10 days averages from October 1992 to December 2000. See *Le Traon et al.* (1995, 1998); *Ducet et al.* (2000) for references. CLS also makes available a joint satellite based SLA/SST data fields.

Ocean Color: SeaWiFS data were processed at Joint Research Centre (JRC), Ispra, Italy for the DIADEM project.

SSM/I brightness temperature for sea ice are obtained from National Snow and Ice Data Center (NSIDC), Boulder and converted to concentration and extent fields using NERSC based algorithms (NORSEX algorithm).

Simulated altimeter waveform (and ice thickness) data are obtained from S. Laxon at UCL, UK.

Forcing fields are taken from NCAR/NCEP for WP 2000 and ECMWF for WP 3000.

Levitus climatology are applied for temperature and salinity. In addition the Global Temperature and Salinity Pilot Project (GTSPP) data are used for validation purposes. This will soon be extended with Argo profile data.

An oceanographic database of the Norwegian and Greenland Seas, compiled at the Arctic and Antarctic Research Institute, Department of Ocean/Atmosphere Interaction in St. Petersburg have been used. The data set is mainly based on data from regular Russian cruises prior to the 1990s, and later on supplemented with data from the International Council for Exploration of the Sea (ICES) under the framework of the INTAS 97-1277 (2000) project. In total more than 127000 hydrographic stations are included in this database. In addition data from the ocean weather station Mike (OWSM) provided by the Geophysical Institute, University of Bergen, is included in this study. The station located at 66N, 2E, is the only surviving weather ship from the observational network starting in the late 1940s. Now we have more than 50 years of near-daily measurements of upper-ocean temperature and salinity, and weekly measurements of the deeper ocean hydrography, the longest existing deep ocean record in the World.

The impact analyses will take into account Earth Observation data such as sea surface temperature, sea surface height, chlorophyll-a, and sea ice concentration and extent from currently operating satellites including ERS–1 and ERS–2, TOPEX/POSEIDON, NOAA TIROS, DMSP, SeaSTAR, and from recently launched missions such as Jason–1 (December 2001) and Envisat (March 2002).

Investigation of the impact of simulated Earth Observation data shall also be performed for the three future ESA Earth Explorer missions, i.e. Cryosat (planned for launch in 2004 with the aim to obtain sea-ice thickness measurements), SMOS (planned for launch in 2005 with the aim to measure surface salinity) and GOCE (planned for launch in 2006 with the aim to measure the marine geoid) *Drange et al.* (1999); *J.A.Johannessen et al.* (2001); *SMOS* (2001); *Kerr* (November 1998); *GOCE group* (1998); *SP-1196(1)* (1996); *Wingham* (1999).

Chapter 2

WP 2000: Climate Research and Monitoring

This WP addresses both regional and global aspects of sea surface salinity (SSS) for both the present day climate and for a possible climate state in the second half of the 21st century. For this, two model systems are adopted. One is a global version of the Miami Isopycnic Coordinate Ocean Model (MICOM; (*Bleck et al.*, 1992)), forced with daily atmospheric re-analyses fields provided by NCAR/NCEP (*Kalnay et al.*, 1996). A dynamic-thermodynamic sea ice module is an integrated part of the model (*Drange* (1996); *Harder* (1996)). The second model system is the Bergen Climate Model (BCM; (*Furevik et al.*, 2003)), consisting of the atmospheric General Circulation Model (GCM) ARPEGE/IFS (*Déqué et al.*, 1994) coupled to the above mentioned version of MICOM.

The spatial-temporal characteristics of SSS have been explicitly examined for the following regions: The Indian-Indonesian-Pacific west of 150 °W; The Kuroshio Current; The North Pacific; The Gulf of Guinea; Equatorial Atlantic; Guyana Current (north of the South-Americal continent); The Caribbean Sea/the Gulf of Mexico; and The region off Newfoundland.

In addition, the coupling between SSS and the following processes have been studied: The North Atlantic Oscillation/Arctic Oscillation; and The Atlantic Thermohaline Circulation.

The objective of the analyses has been to explore and assess the effect of applying available and planned (simulated) remotely sensed data sets to simulate the mean state and variability in state-of-the-art coupled ocean-sea ice and in fully coupled atmosphere-sea ice-ocean GCMs.

The WP-description starts with a section summarising and synthesising the findings of the analyses. Thereafter, the applied model tools are described, followed by descriptions of the findings.

2.1 The OGCM - MICOM Modelling System

For the forced integrations, the MICOM OGCM was configured with a local horizontal orthogonal grid system with one pole over North America and one pole over western part of Asia (*Bentsen et al.*, 1999). The horizontal resolution is 60–80 km over the entire North Atlantic region between 30–60°N. Another integration with doubled horizontal resolution has also been conducted. The overall features from the latter integration are consistent with the 60–80 km resolution results presented here, although local differences exist. A third integration with four times the resolution is presently under way. These integrations, together with results from the coupled system with grid focus in both the atmosphere and the ocean, are the subject of another paper.

The OGCM has 24 layers in the vertical, of which the uppermost mixed layer (ML) has a temporal and spatial varying density. The specified potential densities of the sub-surface layers are chosen to ensure proper representation of the major water masses in the North Atlantic/Nordic Sea region. The densities of the isopycnic layers (in σ_0-units) are

24.12, 24.70, 25.28, 25.77, 26.18, 26.52, 26.80, 27.03, 27.22, 27.38, 27.52, 27.63, 27.71, 27.77, 27.82, 27.86, 27.90, 27.94, 27.98, 28.01, 28.04, 28.07, and 28.10.

The vertically homogeneous ML utilizes the *Gaspar et al.* (1990) bulk parameterization for the dissipation of turbulent kinetic energy, and has temperature, salinity and layer thickness as the prognostic variables. In the isopycnic layers below the ML, temperature and layer thickness are the prognostic variables, whereas salinity is diagnostically determined by means of the simplified equation of state of *Friedrich and Levitus* (1972). The bathymetry is computed as the arithmetic mean value based on the ETOPO-5 data base (Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth, NOAA, National Geophysical Data Center, Boulder, Colorado, 1988).

The thermodynamic module incorporates freezing and melting of sea-ice and snow covered sea-ice (*Drange and Simonsen*, 1996), and is based on the thermodynamics of *Semtner* (1976); *Parkinson and Washington* (1979); and *Fichefet and Gaspar* (1988). The dynamic part of the sea-ice module follows the viscous-plastic rheology of *Hibler* (1979), where sea-ice is considered as a two-dimensional continuum. The dynamic ice module has been further modified by *Harder* (1996) to include description of sea-ice roughness and the age of sea-ice, and utilizes the positive definite advection scheme of *Smolarkiewicz* (1984).

The continuity, momentum and tracer equations are discretised on an Arakawa C-grid configuration (*Arakawa and Lamb*, 1977). The diffusive velocities (diffusivities divided by the size of the grid cell) for layer interface diffusion, momentum dissipation and tracer dispersion are 0.02 m s^{-1}, 0.025 m s^{-1} and 0.015 m s^{-1}, respectively, yielding actual diffusivities of about 10^3 m^2 s^{-1}. A flux corrected transport scheme (*Zalesak* (1979); *Smolarkiewicz and Clark* (1986)) is used to advect the model layer thickness and the tracer quantities.

The diapycnal mixing coefficient K_d (m^2 s^{-1}) is parameterized according to the *Gargett* (1984) expression

$$K_d = \frac{3 \times 10^{-7}}{N},$$

where $N = \sqrt{\frac{g}{\rho}\frac{\partial \rho}{\partial z}}$ (s^{-1}) is the Brunt-Väisälä frequency (here g (m s^{-2}) is the gravity acceleration, ρ (kg m^{-3}) is the density and z (m) is the depth). The numerical implementation of the equation follows the scheme of *McDougall and Dewar* (1998).

2.1.1 OGCM spin-up and forcing

For the simulations presented here, the model was initialized by the January *Levitus and Boyer* (1994) and *Levitus et al.* (1994) climatological temperature and salinity fields, respectively, a 2 m thick sea-ice cover based on the climatological sea-ice extent (*Gloersen et al.*, 1992), and an ocean at rest. The model was then integrated for 150 years by applying the monthly mean NCEP/NCAR atmospheric forcing fields, and thereafter forced with daily NCEP/NCAR reanalysis fields (*Kalnay et al.*, 1996) twice for the five year period 1974–1978. From the NCEP/NCAR reanalysis, wind stress, short wave, long wave, latent and sensible heat fluxes, precipitation, runoff and sea-level pressure fields are used. The momentum, heat and fresh water fluxes are modified when the modelled surface state differs from the NCEP/NCAR reanalysis surface state by applying the *Fairall et al.* (1996) bulk parameterization scheme (*Bentsen and Drange*, 2000).

During the first 150 years of the integration, the ML temperature and salinity were linearly relaxed towards the monthly-mean climatological values of respectively *Levitus and Boyer* (1994) and *Levitus et al.* (1994). The e-folding relaxation time scale was set to 30 days for a 50 m thick ML, and it was reduced linearly with ML thicknesses exceeding 50 m. No relaxation was applied in waters where sea-ice is present in March in the Arctic and in September in the Antarctic to avoid relaxation towards temperature or salinity outliers in the poorly sampled polar waters. In addition, the relaxation was limited to a maximum model to observation mismatch of 0.5 psu and 0.5°C for salinity and temperature, respectively. The latter restriction is of special importance for the Atlantic Ocean as it tends to maintain the simulated hydrodynamic properties of the water masses associated with the Gulf Stream-Labrador Sea

system irrespective of the actual position of these current systems. Continental run-off is included by adding fresh water into the appropriate coastal grid cells (*Furevik et al.*, 2003).

For the integrations presented here, annually repeated heat and fresh water relaxation fluxes with weekly temporal resolution were added to the ML. In this way, temperature and salinity anomalies are free to evolve and propagate, whereas the mean thermodynamic state is kept fairly unchanged. It was found that such a procedure is of special importance for the SSS field, indicating problems with the forcing (either the prescribed NCEP/NCAR precipitation or run-off fields, or the computed evaporation field), or inherent model deficiencies linked to for instance advection or vertical mixing of the surface waters. The weekly heat and fresh water relaxation fluxes were diagnosed from the last 5 years of the spin-up integration. For this period, daily forcing fields for the period 1974–1978 were applied. The period was chosen as it represents a fairly neutral state of the North Atlantic Oscillation (NAO). The first NCEP/NCAR model integration was then initialized from the ocean end state of the 150 years spin-up integration.

2.2 The AOGCM - BCM Modelling System

2.2.1 The atmospheric component

The atmospheric model ARPEGE/IFS is a spectral model which was originally developed for weather prediction by METEO FRANCE and ECMWF - European Centre for Medium-Range Weather Forecasts (*Courtier et al.*, 1991), and later extended to a climate version by *Déqué et al.* (1994). Descriptions of later model improvements can be found in *Déqué and Piedelievre* (1995) and *Déqué et al.* (1998). We will here only give a brief outline of the model.

The ARPEGE/IFS is a spectral model with semi-Lagrangian 2-time level integration. This scheme provides a doubling of efficiency as compared with a 3-time level leapfrog scheme. Semi-Lagrangian formulation also gives the opportunity to use a linear grid for discrete computations. As the number of grid points are smaller in the linear grid than in the more common quadratic grid, this saves additional computational costs (*Hortal*, 1998). In this study we use a spectral truncation of wave number 63, linear grid and a time step of 1800 s. The linear T_L63 grid has the same number of points as the quadratic T42 grid.

The hydrostatic Navier-Stokes equations govern the evolution of the atmospheric flow by ensuring conservation of mass, energy and momentum. The acceleration of gravity is constant. The model atmosphere is a mixture of air, water vapour and an optional number of dynamically passive constituents. One of these is ozone, which is simulated using an Eulerian advection scheme. This option has been turned off in the present integration.

The energy sources and sinks in the equation system described above arise from discretisation, horizontal diffusion and the time filter. Sources and sinks due to small scale physical processes are parameterised. The grid-boxes are defined by the computational grid, consisting of the points at which the non-linear terms in the Navier Stokes equations are calculated (reduced Gaussian grid). This is a latitude/longitude grid, which in the T_L63 case has 64 nearly equidistant latitudes. The reduction of the Gaussian grid near the poles (*Hortal and Simmons*, 1991) gives approximately uniform horizontal resolution (on the target sphere). The horizontal grid distribution is shown in Figure 2.1. The vertical hybrid coordinate (*Simmons and Burridge*, 1981) follows the topography in the lower troposphere, but becomes gradually parallel to pressure surfaces with increasing height. For the experiments presented here, we apply 31 model levels, ranging from the surface to 0.01 hPa. Spurious reflection at the model top is avoided through strong horizontal diffusion in the top model layers.

The physical parameterisation is divided into several explicit schemes, which each calculates the flux of mass, energy and/or momentum due to a specific physical process. The physical parameterisation schemes in ARPEGE have originally been taken from the climatic version of Météo-France's EMERAUDE model, described in *Coiffer et al.* (1987). Different schemes employed in the present version of ARPEGE and relevant references are listed in Table 2.1.

Different from the model description in *Déqué et al.* (1998), the particular version used in BCM contains a convective

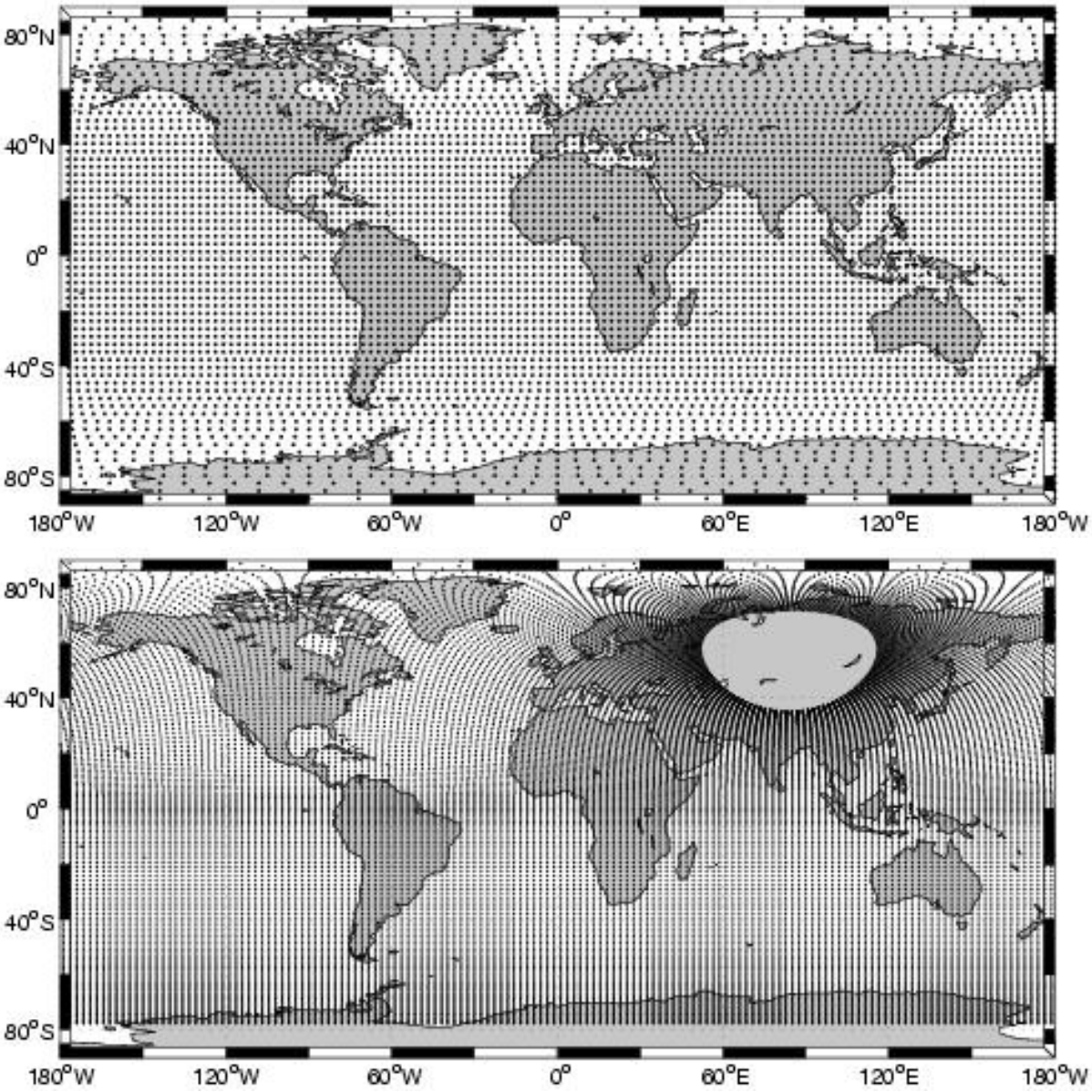


Figure 2.1: The horizontal grid distribution for the atmosphere (upper) and ocean (lower).

Parameterisation type	Reference
Surface	*Masutani et al.* (1995)
Turbulence	*Louis* (1979); *Geleyn* (1988); *Richard and Royer* (1993)
Shallow convection	*Geleyn* (1987)
Radiation	*Morcrette* (1991)
Convection	*Bougeault* (1985)
Ozone	*Cariolle and Déqué* (1986)
Gravity wave drag	*Déqué et al.* (1994); *Lott and Miller* (1997); *Lott* (1999)
Mesospheric drag	*Déqué et al.* (1994)
Soil	*Noilhan and Planton* (1989); *Mahfouf et al.* (1995)
Clouds and precipitation	*Richard and Royer* (1993); *Déqué et al.* (1994)

Table 2.1: Parameterisation schemes employed in ARPEGE.

Condition	Source
Orography	US Navy data
Vegetation roughness length	*Mahfouf et al.* (1995)
Land ice extent, vegetation, surface emissivity	CLIMAP 1981
Albedo	*Geleyn and Preuss* (1983)

Table 2.2: Boundary data employed in ARPEGE.

gravity-wave drag parameterisation (*Bossuet et al.*, 1998), a new snow scheme (*Douville et al.*, 1995), an increase of the orographic wave drag (*Lott*, 1999) and modifications in deep convection and soil vegetation schemes. Data for boundary conditions employed by the ARPEGE model are listed in Table 2.2.

2.2.2 The ocean-sea ice component

The main features of the ocean and sea ice modules in BCM are identical to the description given above. Therefore, only the major differences are described here.

To avoid grid singularities in the computational ocean domain (*Bentsen et al.*, 1999), one pole is located over central Siberia while the other is at the South Pole (Figure 2.1). The resulting ocean grid configuration is, for comparison, quite similar to the one presented in *Madec and Imbard* (1996). With the exception of the equatorial region, the ocean grid is almost regular, with horizontal grid spacing approximately $2.4^\circ \times 2.4^\circ$. In order to better resolve the dynamics near the equator, the horizontal spacing in the meridional direction is gradually decreased to 0.8° along the equator. In the vertical, the ocean model has 24 layers, with potential densities ranging from $\sigma_\theta = 23.54$ to $\sigma_\theta = 28.10$.

The ocean - sea ice module share the same grid, and the heat, salt and water fluxes among them are handled in an internally consistent way.

2.2.3 The coupler

The OASIS (Ocean Atmosphere Sea Ice Soil) coupler has been used to couple the atmosphere and ocean models. It was developed at the National Centre for climate modelling and global change (CERFACS), Toulouse, France (*Terray and Thual* (1995); *Terray et al.* (1995)), and is currently in use in many climate centres, among them the ECMWF, Max Plank Institute (MPI) and CERFACS (e.g. *Guilyardi and Madec* (1997); *Cassou et al.* (1998); *Baretta et al.* (1998)). In BCM, OASIS version 2.2 (*Terray et al.*, 1998) is used.

The main tasks of OASIS are to synchronise the models, so the fastest running model can wait for the other model until they are both integrated a prescribed time interval (1 day), and to read the exchange fields from the source model, apply weight coefficients for the interpolations, and finally write the new fields to the target model.

2.3 Model integrations and experiments

Three types of model integrations have been performed:

- One 300 years control integration with the BCM in the setting described above
- Four 80 years integrations in which the atmospheric concentration of CO_2 is increased by 1 % pr year (following the so-called CMIP-protocol). The only difference between the four integrations is the initial condition, taken as four arbitrary climate states from the 300 years control integration

- Six 54 years integrations with the uncoupled OGCM, forced with daily NCEP/NCAR reanalysis fields for the period 1948 to present. Here two types of experiments have been conducted: 1) The horizontal resolution in the central North Atlantic Ocean has been 80 km, 40 km and 20 km. 2) For the 80 km case, four members have been integrated in order to assess the role of different initial conditions on the solution.

The CO_2-integrations are representative of the *net* effect of the present day global warming, and the atmospheric CO_2 concentration doubles at about 70 years. The members are identified by the names E76–E79 in the figure panels below.

The reason for presenting the four-member CO_2-ensemble is to identify features that are *not* dependent on (random occurring) natural climate variability modes like El-Niño/Southern Oscillation (ENSO) and NAO/AO in the model system. Therefore, the differences between the four ensemble members can be considered to (at least partially) span out the natural variability characteristics in time and space w.r.t. SSS and SST fields.

The state of the natural variability modes is positively unknown prior to a satellite-mission, with the possibility that the ENSO-variability can be forecasted for up to 6 months. To cover the actual ocean state during a specific mission, the presented analyses should, ideally, take into consideration cases with the actual ENSO, NAO/AO, etc. modes. Such an analyses is - in principle - feasible with a coupled atmosphere-sea ice-ocean system. The analyses presented here have been limited to two 20-years time periods (simulation year 1-20 and year 61-80), and by that focussing on a possible present-day situation, and a scenario for the situation in the second half of the 21st century.

It should be mentioned that the presented results are based on one model system. The obtained variability should therefore be evaluated against, preferably, *in situ* observations. Unfortunately, *in situ* observations of the global SSS-field are not available. Therefore, the best evaluation of the results presented here need to be based on the relatively few available *in situ* SSS observations, on modelled hindcast experiments, and on comparable diagnostics from other OGCMs and AOGCMs.

Finally, the analyses have been extended to include the evolution and variability of SSS and SST at a doubling of the atmospheric concentration of CO_2. The latter fields yield an estimate of a possible change in the thermodynamic properties of the World Ocean surface waters during the second half of the 21st century. This comparison is of importance for remotely sensed SSS as it indicates the accuracy needed for present-day SSS-products to be of value to map (likely) changes in the SSS-field during the 21st century.

2.4 Results WP 2100 Impact Assessment

2.4.1 The atmosphere-sea ice-ocean system

The analyses presented below are based on the annual and seasonal temporal-spatial variability of the global SSS-field produced by the atmosphere-sea ice-ocean model BCM.

Four integrations of BCM are used in the analyses, and the only difference between the integrations are different initial states of the model system. There are basically two reasons for performing an ensemble-type of integrations with a coupled climate system. First of all, the 3-dimensional state of the global climate system for a specific time is unknown. Secondly, a climate model is not perfect (i.e., it does not reproduce the real climate system in all aspects), so even if the initial state of the climate system were known, the model evolution of the climate would depart from the real climate. The ensemble approach presented here gives a measure of the mean spatial-temporal state and the spreading about the mean state.

The four integrations, each one lasting for 80 years, are started out from four different states of a 300 years control integration with BCM (*Furevik et al.*, 2003). The concentration of atmospheric CO_2 is increased by 1 % per year in each of the four ensemble members. This increase corresponds roughly to the expected, net effect of greenhouse

gasses and aerosol particles in the atmosphere in the 21st century. For comparison, at the end of the ensemble integrations, the global surface temperature increases by about 2 °C, which is in the bulk-part of the climate change scenarios presented by the Intergovernmental Panel of Climate Change (IPCC).

In the analyses presented below, monthly mean fields of SSS have been used for the first 20 years of the integration (resembling the present-day climate), and for the last 20 years of the integration (resembling a possible climate state in the middle of the second half of the 21st century).

The obtained standard deviation (std) maps indicate where remotely sensed SSS can be used to detect changes in the annual to seasonal SSS-field. In addition, regions with low natural variability in SSS are candidates for long-term calibration of the SSS-sensor.

The SSS is given in psu, or the so-called practical salinity unit. This is the standard measure of the salinity in the ocean, and it essentially represents the weight in gram of dissolved minerals (or salts) per kg seawater. A typical salinity of the surface waters is 35 psu, or 35 parts per thousand.

The four ensember members are labelled E76 to E79.

2.4.2 Annual SSS

Figure 2.2 shows, in the upper row, the simulated global SSS (psu) and SST (°C) from ensemble member E77. In row 2–3 (4–5), the SSS (SST) sdv for each of the members are given based on year 1–20 of the integration (i.e., for the present day climate forcing).

Annual variability exceeding 0.1 psu std in SSS is mainly found in the tropics between 30 °S and 30 °N. All of the ensemble members show largest variability in the Central and Eastern Indian Ocean, in the Indonesian Region, and in the Pacific west of 150 °W. The variability in the tropical Atlantic Ocean is generally much weaker than in the Indian and Indonesian regions. One pronounced exception to this is the fairly large variability in SSS in the waters just north of the South-American continent.

There are essentially no variability in the SSS-field poleward of 30 °S. This does not necessarily imply that the Southern Ocean is a region of no variability. There are two reasons for this: Firstly, the Southern Ocean is a region without boundary currents like the Gulf Stream and Kuroshio Current. This means that the heat transport - and possibly the salt transport - is governed by eddies, and eddies are positively not present in the applied model system. Secondly, the Southern Ocean is probably the region where both ocean GCMs and climate GCMs differ most. This, in combination with very few SSS-observations in the Southern Ocean, makes it hard to assess the reliability of the simulated SSS-field and variability.

Poleward of 30 °N, variability at about 0.1 psu std in SSS is found over most of the the Pacific Ocean, and in the region off Newfoundland. This variability is linked to the southward flowing fresh water coming from the North Atlantic sub-polar gyre, and are in accordance with observations (Yashayaev *et al.*, available from http://www.mar.dfo-mpo.gc.ca/science/ocean/woce/labsea/labsea_poster.html).

2.4.3 Seasonal SSS

The seasonal variability is similar to the annual variability, although the variability is generally stronger. The following regions have been explicitly considered: 1) Indian-Indonesian-Pacific west of 150 °W, 2) Kuroshio Current, 3) North Pacific, 4) Gulf of Guinea, 5) Equatorial Atlantic, 6) Guyana Current (north of the South-Americal continent), 7) the Caribbean Sea/the Gulf of Mexico, and 8) the region off Newfoundland.

In the following, the SSS std is characterised *weak*, *intermediate* and *strong* based on comparison with the other seasons. Therefore, the characteristics ”Equatorial Atlantic: Strong, between 0.1-0.2 psu std” for season X means

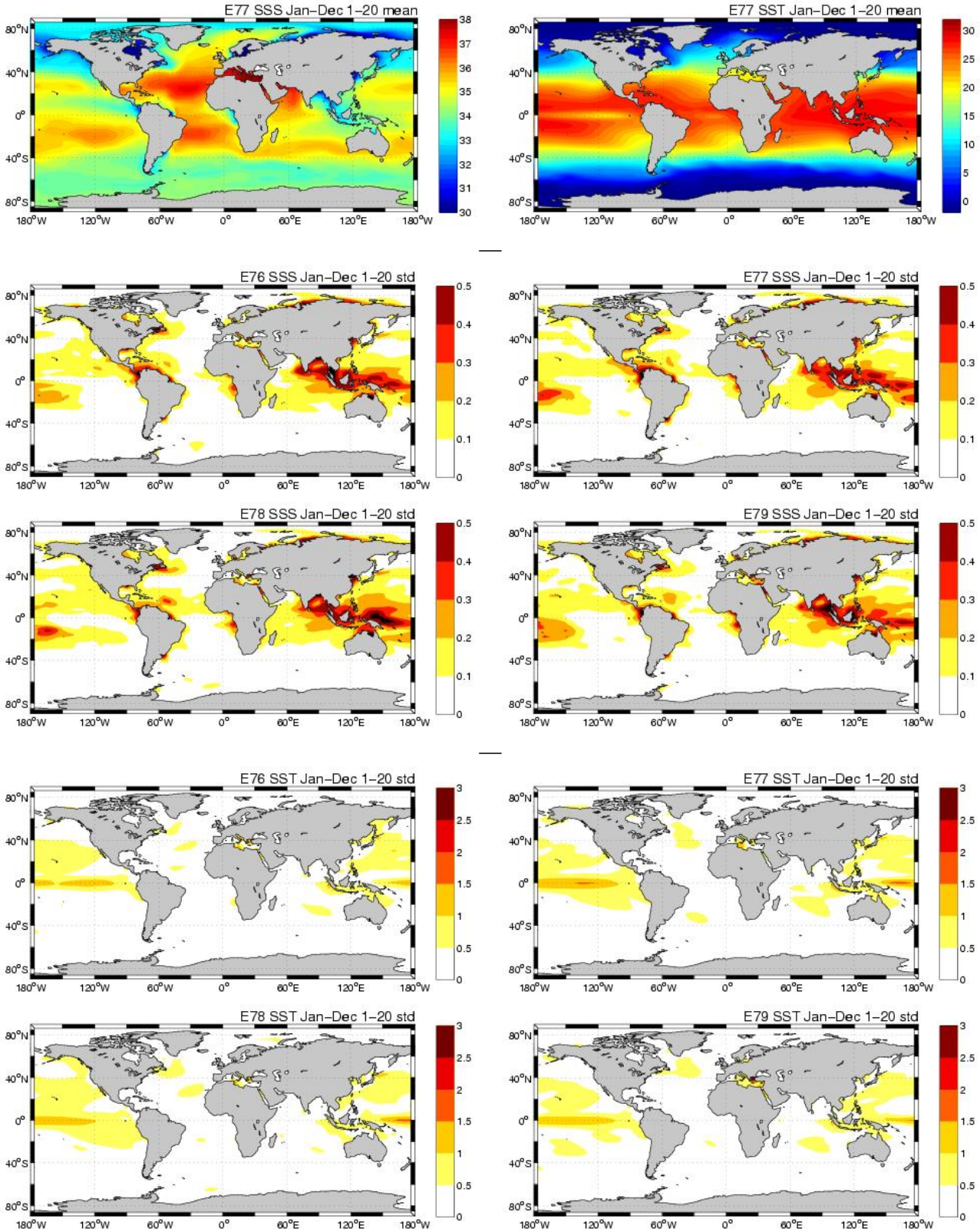


Figure 2.2: Uper row: Mean SSS and SST for year 1-20 of E77. Row 2–3: SSS sdv for E76–E79 for year 1-20. Row 4–5: SST std for E76–E79 for year 1-20.

that season X is the season with strongest variability in the Equatorial Atlantic, and that the variability is 0.1-0.2 psu std.

January–March

The variability in SSS as seen from Fig. 2.3 can be characterised as follows (based on the above mentioned regions):

1) Indian-Indonesian-Pacific west of 150 °W: Strong, $>$ 0.5 psu std over large regions bounded by 70 °E-180 °E
2) Kuroshio Current: Weak, at most 0.1 psu std
3) North Pacific: Weak, about 0.1 psu std
4) Gulf of Guinea: Strong, reaching 0.5 psu std towards the African coast
5) Equatorial Atlantic: Intermediate, between 0.1-0.2 psu std
6) Guyana Current (north of the South-Americal continent): Weak, about 0.1 psu std
7) The Caribbean Sea/the Gulf of Mexico: Intermediate, about 0.3 std psu south of the Hispaniolas
8) The region off Newfoundland: Weak, slightly above 0.1 std psu

April–June

The variability for the spring season April–June are given in Fig. 2.4, and can be characterized as .

1) Indian-Indonesian-Pacific west of 150 °W: Intermediate, generally $<$ 0.5 psu std
2) Kuroshio Current: Intermediate, 0.1-0.2 psu std
3) North Pacific: Weak, at most 0.1 psu std
4) Gulf of Guinea: Weak, only significant SSS-signal just off the African coast
5) Equatorial Atlantic: Intermediate, around 0.2 psu std
6) Guyana Current (north of the South-Americal continent): Strong, exceeding 0.3 psu std
7) The Caribbean Sea/the Gulf of Mexico: Strong, $>$ 0.5 std psu south of the Hispanolan islands
8) The region off Newfoundland: Intermediate, exceeding 0.2 std psu

July–September

The July-September SSS variability are provided in Fig. 2.5, and can be grouped as follows:

1) Indian-Indonesian-Pacific west of 150 °W: Weak, generally $<$ 0.4 psu std
2) Kuroshio Current: Strong, about 0.4 psu std
3) North Pacific: Strong, generally above 0.1 psu std over the entire region
4) Gulf of Guinea: Weak, 0.2 psu std
5) Equatorial Atlantic: Intermediate, at most 0.2 psu std
6) Guyana Current (north of the South-Americal continent): Strong, exceeding 0.3 psu std
7) The Caribbean Sea/the Gulf of Mexico: Weak, only significantly larger than 0.1 psu std just north of the South-American continent
8) The region off Newfoundland: Strong, 0.2-0.3 std psu

October–December

The October-December variability in SSS is given in Fig. 2.6, and has the following characteristics:

1) Indian-Indonesian-Pacific west of 150 °W: Strong, 0.4 psu std and more over a significant part of the region

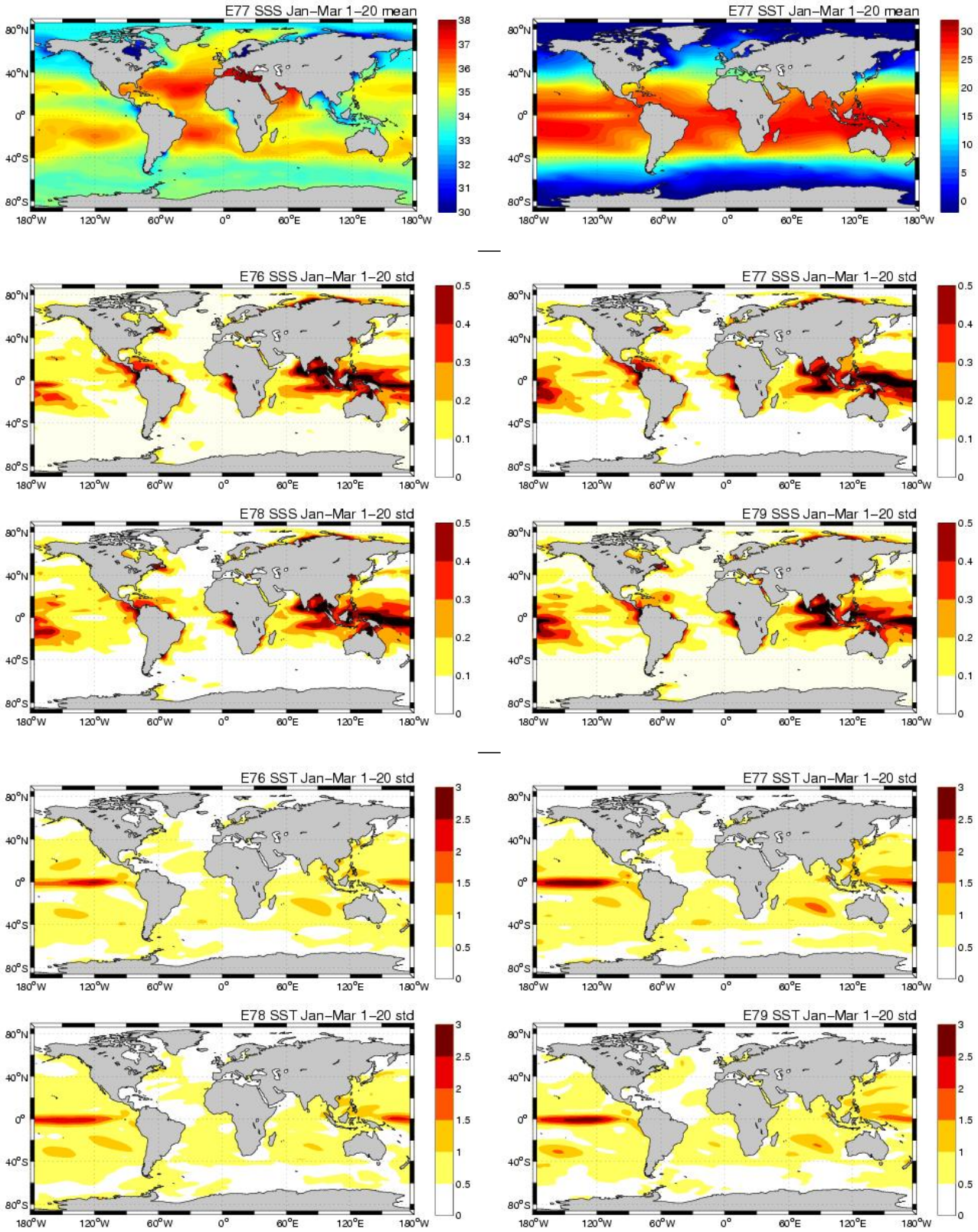


Figure 2.3: Uper row: Mean SSS and SST for Jan-Mar of year 1-20 of E77. Row 2–3: SSS std for E76–E79 for Jan-Mar of year 1-20. Row 4–5: SST std for E76–E79 for Jan-Mar of year 1-20.

2) Kuroshio Current: Weak, about 0.1 psu std

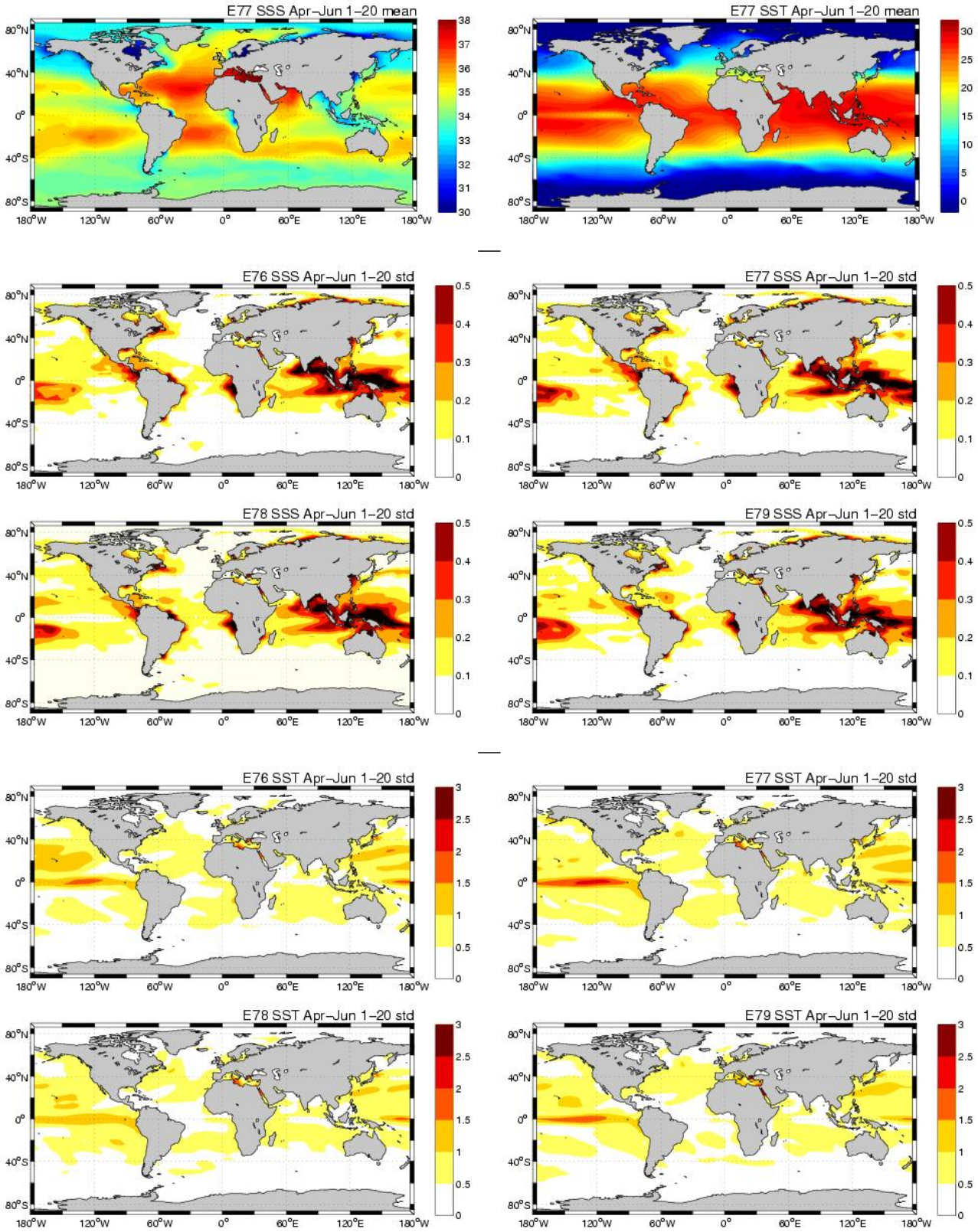


Figure 2.4: Uper row: Mean SSS and SST for Apr-Jun of year 1-20 of E77. Row 2–3: SSS std for E76–E79 for Apr-Jun of year 1-20. Row 4–5: SST std for E76–E79 for Apr-Jun of year 1-20.

3) North Pacific: Weak, hardly exceeding 0.1 psu std

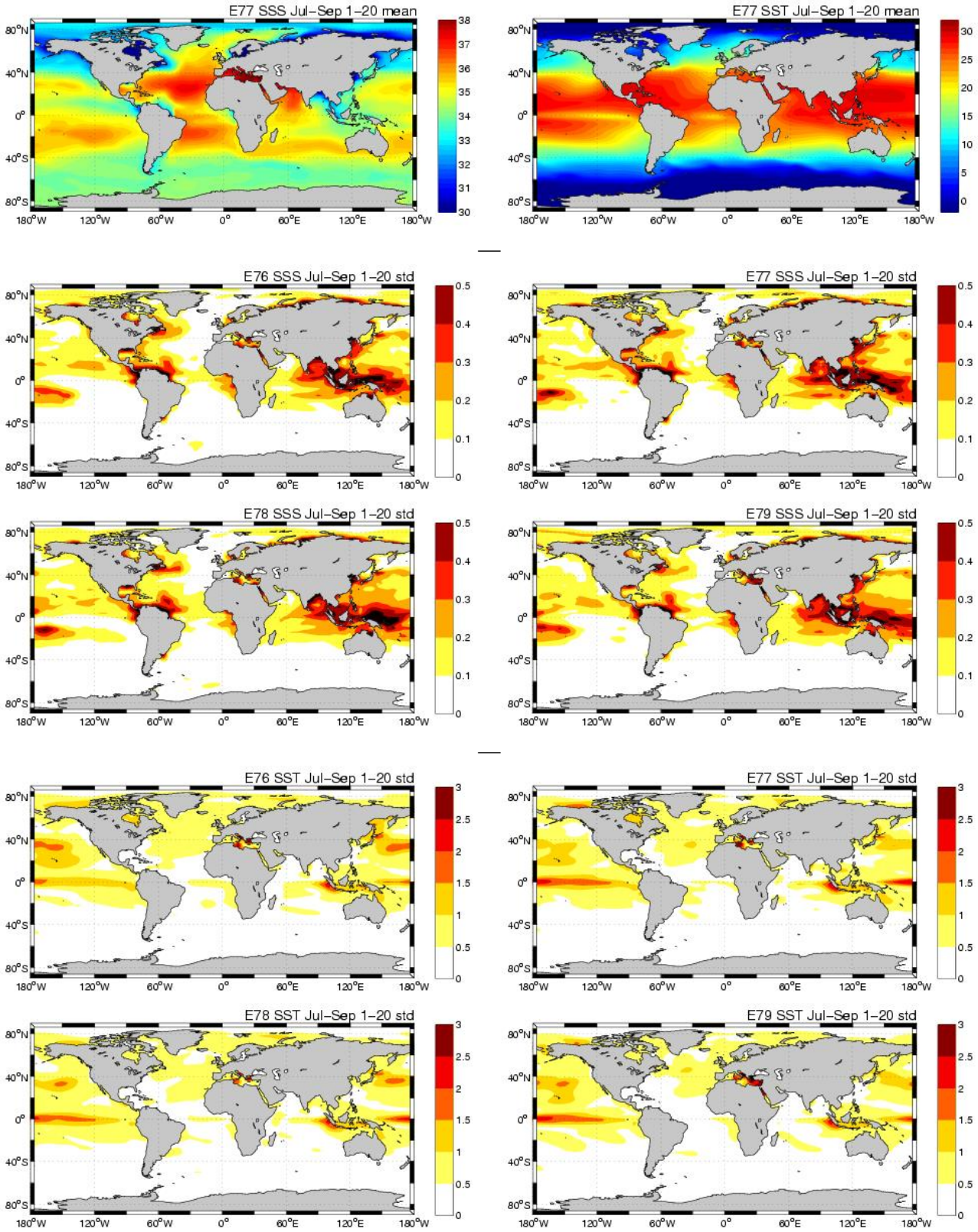


Figure 2.5: Uper row: Mean SSS and SST for Jul-Sep of year 1-20 of E77. Row 2–3: SSS std for E76–E79 for Jul-Sep of year 1-20. Row 4–5: SST std for E76–E79 for Jul-Sep of year 1-20.

4) Gulf of Guinea: Strong, $>$ 0.5 psu std

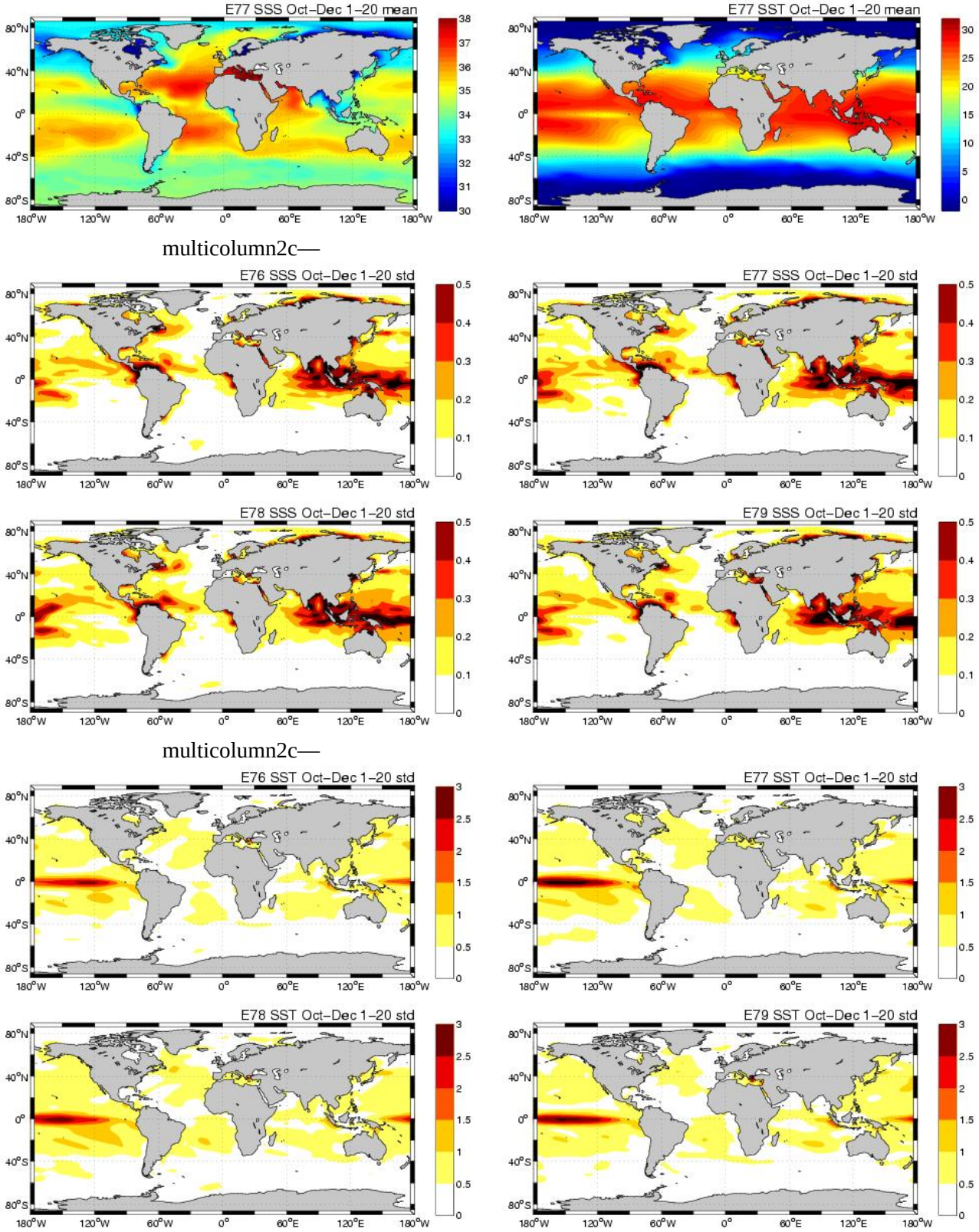


Figure 2.6: Uper row: Mean SSS and SST for Oct-Dec of year 1-20 of E77. Row 2–3: SSS std for E76–E79 for Oct-Dec of year 1-20. Row 4–5: SST std for E76–E79 for Oct-Dec of year 1-20.

5) Equatorial Atlantic: Weak, about 0.1 psu std

6) Guyana Current (north of the South-Americal continent): Weak, about 0.1 psu std
7) The Caribbean Sea/the Gulf of Mexico: Weak, 0.1-0.2 psu std
8) The region off Newfoundland: Intermediate, 0.1-0.2 std psu

2.5 WP 2200: Identification of improvement

2.5.1 The ocean-sea ice system

To examine the ocean response of the highly variable atmospheric forcing, years with high and low winter-NAO index forcings have been grouped, and the differences in the ocean state between these years and the mean ocean state were examined. In this exercise, deviations in the NAO index exceeding 1 std were considered as years with high or low NAO index. Over the time period 1948-1998, 9 years were identified as high NAO years (1957, 61, 73, 75, 76, 81, 83, 84, 89, 90, 92, 93, 94, 95, 99), and nine years where identified as low NAO years (1958, 60, 62, 63, 64, 65, 66, 68, 69, 70, 71, 77, 79, 87, 96).

2.5.2 SST and SSS variability linked to the NAO

The major findings are displayed in Fig. 2.7.

It follows from the figure (upper row) that there is a significant change in SST with changes in the NAO index. For years with high NAO index, the SST in the Nordic Seas, the Baltic and the sub-tropical Atlantic is up to 1 °C higher than on average, whereas the Irminger and Labrador Seas are up to 1 °C lower than on average. The opposite situation is the case for years with low NAO index. The extent and strength of the obtained three-pole pattern between the Nordic Seas, the Irminger and Labrador Seas, and the sub-tropics are in general accordance with observations (*Dickson*, 1997).

The most striking features seen in the SSS field for high NAO-years are the higher than normal SSS in the North Sea (above 0.3 psu), in the eastern part of the sub-tropical region (up to 0.1 psu), in the polar waters north of Newfoundland (up to 0.2 psu), in the Guyana Current region and south of the Hispaniola islands (about 0.1 psu), and the lower than normal SSS in the waters just north of the Gulf Stream (about 0.1 psu). Again, the situation is near reversed for low NAO-years. Therefore, the difference in SSS between high and low NAO-years is about twice of the above stated values.

There are also substantial differences in the thickness of the upper mixed layer between years with high and low NAO indexes (see Fig. 2.7, third row). The most prominent feature is the di-pole pattern between the Greenland Sea and the Labrador-Irminger Seas. This di-pole is caused by both changes in SST and SSS, and the pattern is consistent with observations (*Dickson*, 1997).

Another interesting feature is the change in the water transport as a consequence of variations in the atmospheric forcing. In Fig. 2.7, lower row, the anomalies in the vertically integrated stream function are displayed. It is seen that the largest differences in the mass transport (exceeding 5 Sv) are found in the sub-tropical and sub-polar gyres of the Atlantic Ocean and in the central Arctic Basin. There is also a 1-2 Sv difference in the strength of the circulation in the Norwegian Sea with 1 Sv higher (2 Sv weaker) than normal circulation for years with high (low) winter NAO indexes. There are no direct observations of the strength of the basin scale transport in the ocean. However, analyses of hydrographic data indicate a pattern similar to the pattern predicted by the model (*Curry et al.*, 1998).

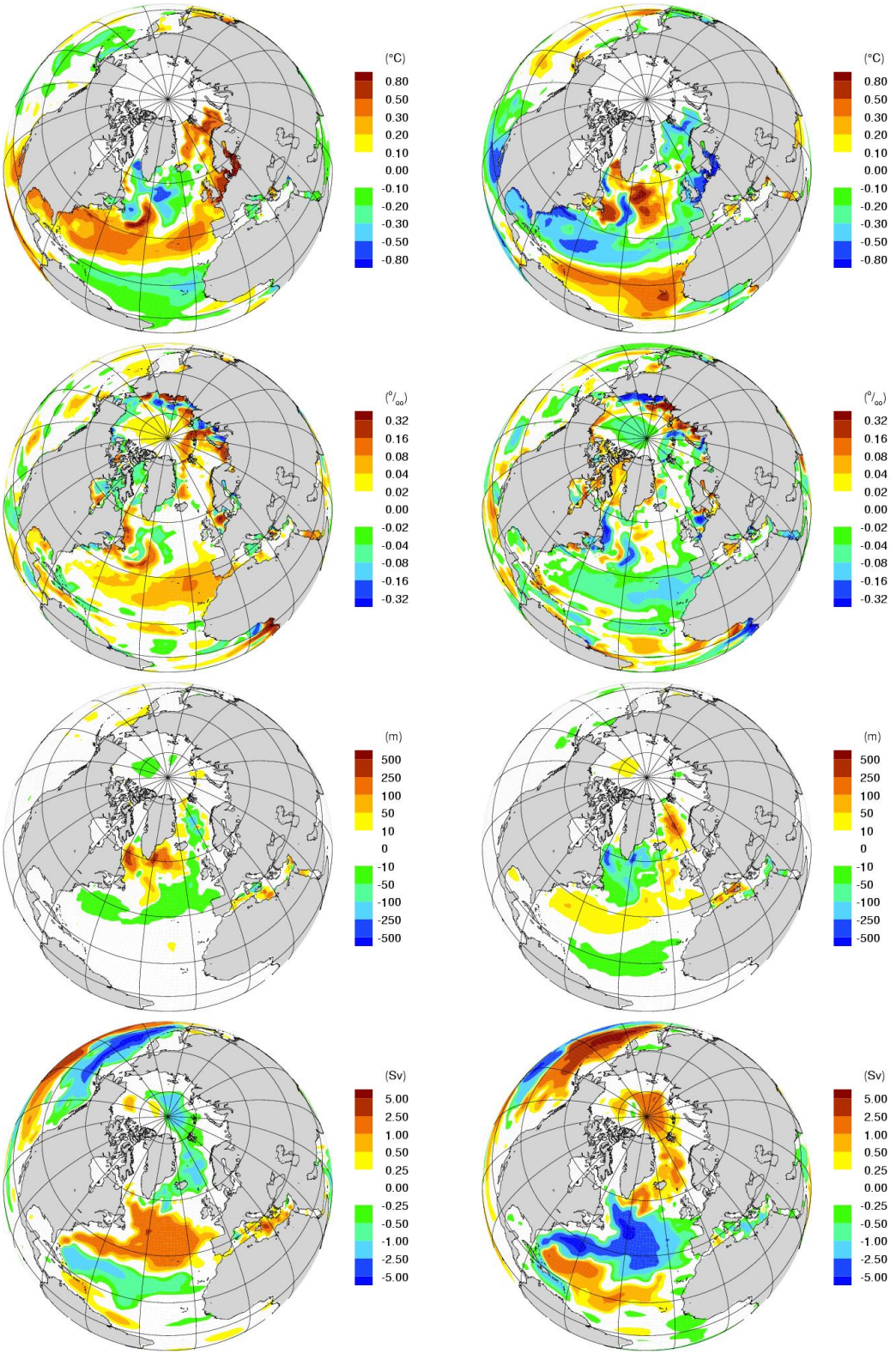


Figure 2.7: Simulated anomalies in March for winters with high (left panel) and low (right panel) of SST (°C) at the top row, SSS (psu) in second row, mixed layer thickness (m) in third row and stream function (Sv) at the bottom.

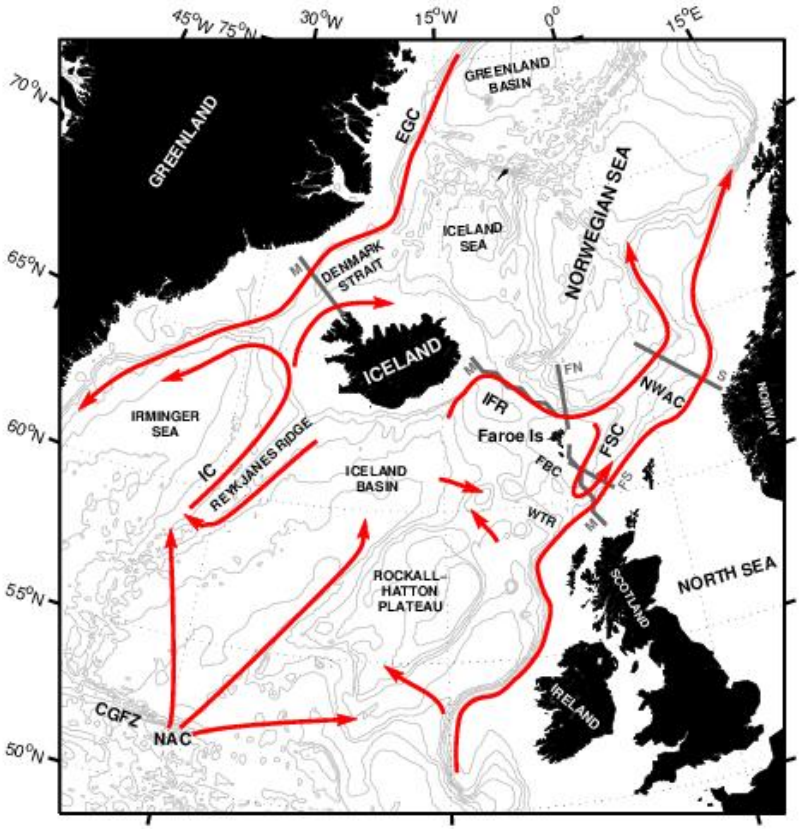


Figure 2.8: The GIS-Ridge and its surrounding waters. Isobaths are drawn for every 500 m. Schematic surface currents with key references are indicated. Abbreviations are explained in the text. Grey lines indicate model (M), Faroe north (FN) and south (FS), and Svinøy (S) sections. From *Nilsen et al.* (2003).

2.6 Importance of import of SSS to the Nordic Seas

The Atlantic Thermohaline Circulation (ATHC) is a dynamically active component of the climate system, in particular on multi-annual to decadal time scales (*Curry and McCartney*, 2001). The heat and salt carried northward across the Greenland-Iceland-Scotland (GIS) ridge are substantial, and both quantities are of importance for the water mass and ice distribution of the Nordic Seas and Arctic Ocean, and possibly also the deep mixing and water mass transformations taking place in the region (*Furevik et al.*, 2003). The gateways for the exchange of water masses between the two ocean basins are the 290 km wide and 200-620 m deep Denmark Strait (DS) between Greenland and Iceland, the 400 km wide and 300-500 m deep Iceland-Faroe Ridge (IFR), and the 200 km wide Faroe-Shetland Channel (FSC) with the 850 m deep Faroe-Bank Channel (FBC) at its entrance (Fig. 2.8).

Time series of simulated transports over the GIS ridge have been low-pass filtered using a Butterworth filter with cut-off period of 3 years (Fig. 2.9). All sections reveal substantial variability, with typical amplitudes for the transport anomalies of the order 1-2 Sv. The net flow through the FSC, for instance, was very low during the first 20 years (1.8 Sv), then for the next 17 years the average was 3.3 Sv, before decreasing to near 2 Sv for the remaining part of the simulation. The total AtI however has highest transport from the late 50's to early 70's.

As expected, also temperature and salinity show pronpunced multi-annual to decadal fluctuations in the inflow region. This is exemplified in Figure 2.10, showing time-series of temperature and salinity at three locations near or at the Faroe Islands.

The following conclusions can be drawn based on the observation-model comparison: 1) The hind-cast model system

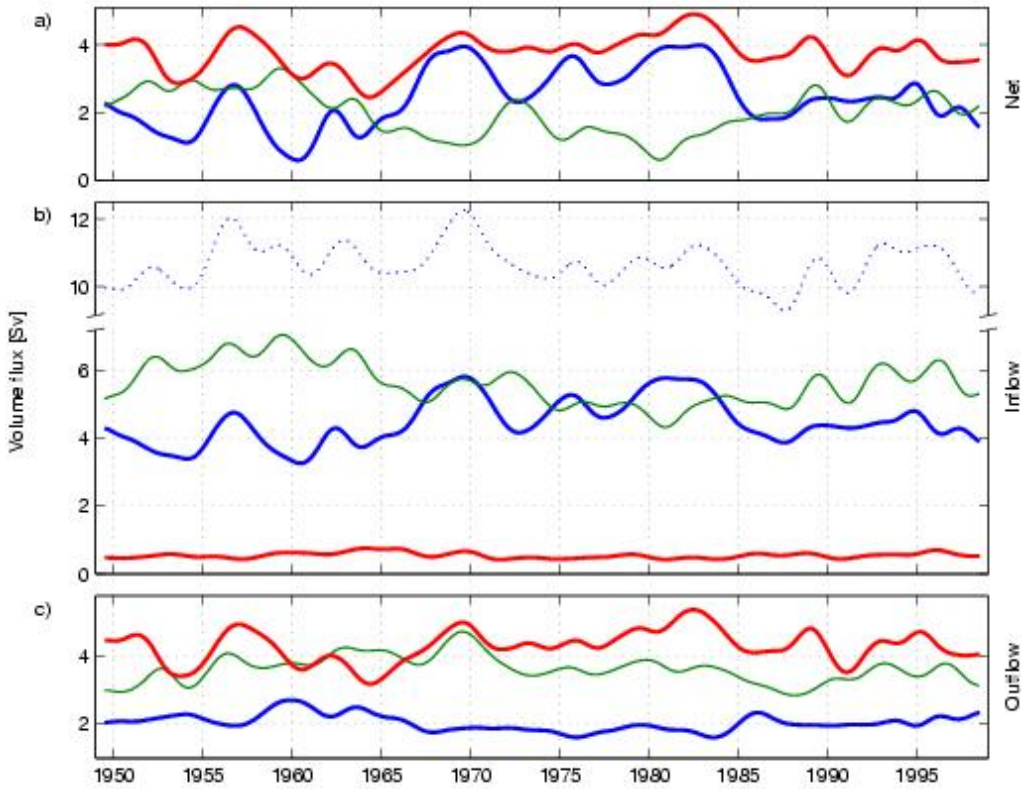


Figure 2.9: Temporal variation of the 3-years low pass filtered simulated net (a), inflow (b), and outflow (c) volume transports through the FSC (solid lines), IFR (dashed lines), DS (dot-dashed lines), and total inflow (dotted line). Note that the net flux through the DS is defined positive southwards. From *Nilsen et al.* (2003).

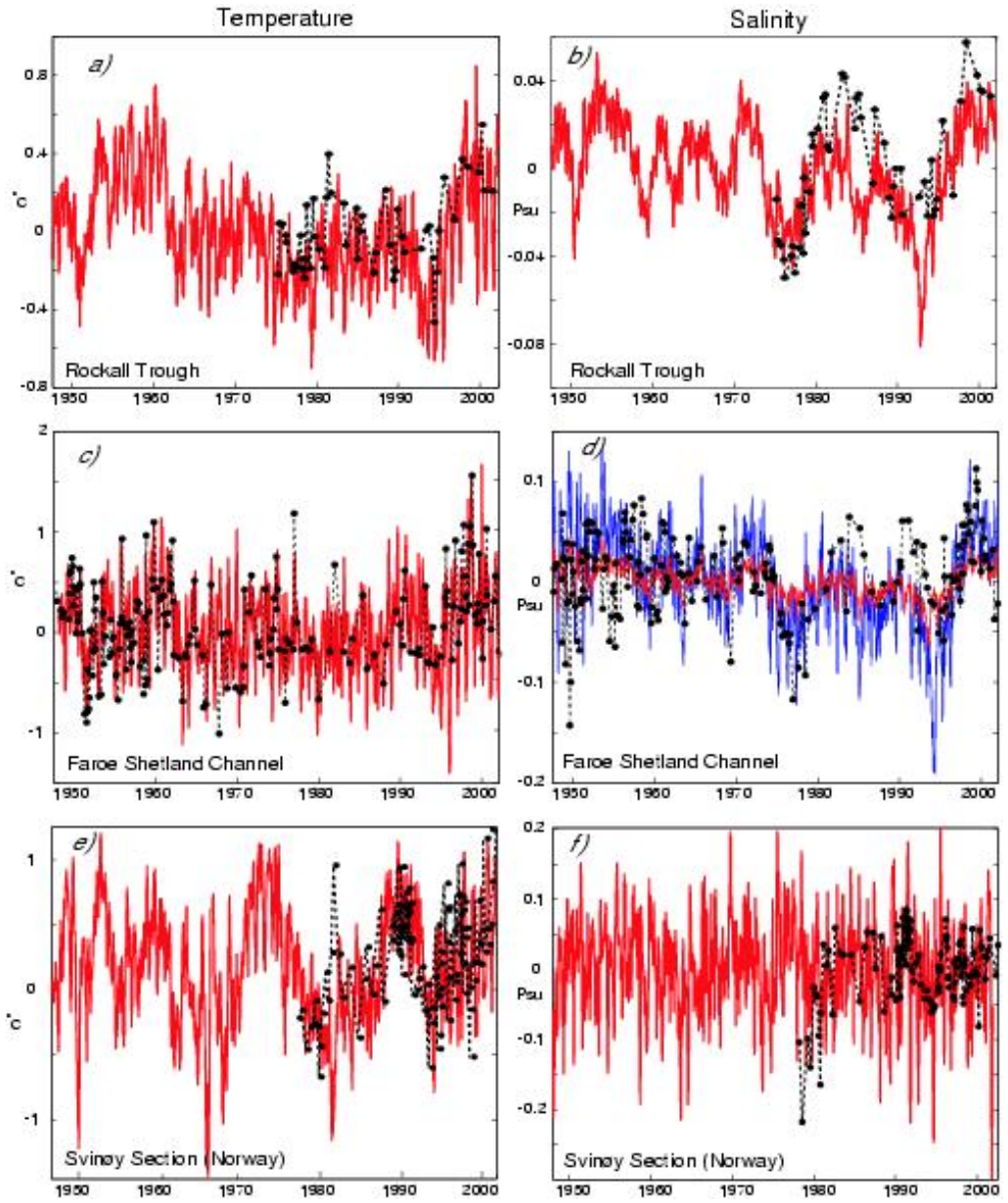


Figure 2.10: Temporal variation of the 3-years temperature and salinity *anomalies* south of the Greenland-Scotland Ridge (Rockall Trough), on the ridge (Faroe-Shetland Channel) and north-east of the ridge (Svinøy section). Simulated data from the regional model are shown in red and the observations are shown as filled circles. Note: Different vertical axes are used. The blue line in panel d) is three times the amplitude of the observed variability. Figure from Hátún *et al.*, in prep.

is able to qualitatively describe the inter-annual variability in temperature. 2) The fit to the observations are not so good for salinity – in general the simulated amplitude in the salinity variability is too low. 3) The observed, 3-years anomalies in salinity, range from 0.1–0.3 psu.

The reason for the relatively realistic simulation of temperature is believed to be linked to the tight coupleing between SST and the (prescribed) surface air temperature. The coupling between SSS and the atmopshere is much weaker, and takes place indirectly through evaopration−precipitation fluxes. Therefore, remotely sensed SSS have the potential to improve the model system in the region both in hind-case and in forcast (or predictability) modes.

2.6.1 Atlantic Thermohaline Circulation

Another way to display the basin scale response to the applied variability in the atmospheric forcing fields is to look at the Atlantic Thermohaline Circulation (ATHC).

A common diagnostics for the ATHC is the strength of the overturning, either extracted within a certain latitude-depth interval, or at a fixed latitude. In the upper panel of Fig. 2.11, the simulated annual mean ATHCs (in Sv=10^6 m^3 s^{-1}) from the forced integrations in the latitude band 20°N-50°N are shown for the period 1948–1999. Irrespective of the initial conditions, the ATHCs show the same trend and variability after an adjustment time of 5–10 years (this result also holds for the conducted integration with doubled horizontal resolution). In particular, the model realizations yield minima round year 1960 and 1980, maxima round 1975 and 1995, and a gradual increasing trend from 1960 to 1995.

Decadal scale variability is also found in the 300 years BCM simulation (lower panel of Fig. 2.11). From the BCM integration, variability modes on longer time scales are also present. Whether these modes are real or are caused by deficiencies of the coupled system, is hard to assess. However, the focus of this study is on decadal scale variability, so the long term variability has been filtered in most of the analyses presented here.

The locations of the convective mixing in the forced integrations and the BCM simulation are fairly similar (Fig. 2.12), and are located in the Labrador, Irminger and GIN Seas. The classical picture is that deep convective mixing takes place in the Labrador and the GIN Seas, but not in the Irminger Sea. However, as pointed by *Pickart and Lavender* (2000) and *Pickart et al.* (2002), historical and recent observations indicate that the Irminger Sea is ialso a region of deep mixing. This is also the case for the modelling results presented here.

Volume mixing indices are constructed as the normalized mean February to April (FMA) ML volume in the areas where the FMA ML depth exceeds 1500 m at least once during the actual integration (the shaded areas in Fig. 2.12). The maximum ML is between 100 and 200 m deeper in the BCM experiment compared to the forced integrations, but otherwise very similar. For both types of model experiments, the mixing penetrates deepest in the Irminger and GIN Seas (1000–1400 m on average), whereas there is mixing to 600–800 m in the Labrador Sea.

Lag correlations along vertical sections are a powerful diagnostics to examine the formation, propagation and decay of thermodynamic and dynamic anomalies. Here we present density, temperature, salinity and velocity anomaly correlations with respect to the leading PC of the ATHC from the BCM integration along two sections: One section follows the western continental slope of the Atlantic Ocean, and the other the 48°N parallel (Fig. 2.13). The corresponding fields from the forced integrations resemble those presented here, and are therefore not displayed.

For the section along the western continental slope (Fig. 2.14), density anomalies caused by temperature, and to some extent by salinity, form at lag of $+6$ to $+4$ years in the sub-polar gyre. These anomalies grow with time, and a deep southward propagating signal is seen as cold, fresh and dense water. From the velocity panels, it follows that the southward flow of dense water is compensated by a northward velocity anomaly extending over the uppermost 4000 m of the water column. It is also seen that the surface water density anomaly disappears at lags < -2 years, and that this is caused by warm and saline surface waters, with the former overcompensating the effect on density compared to the latter.

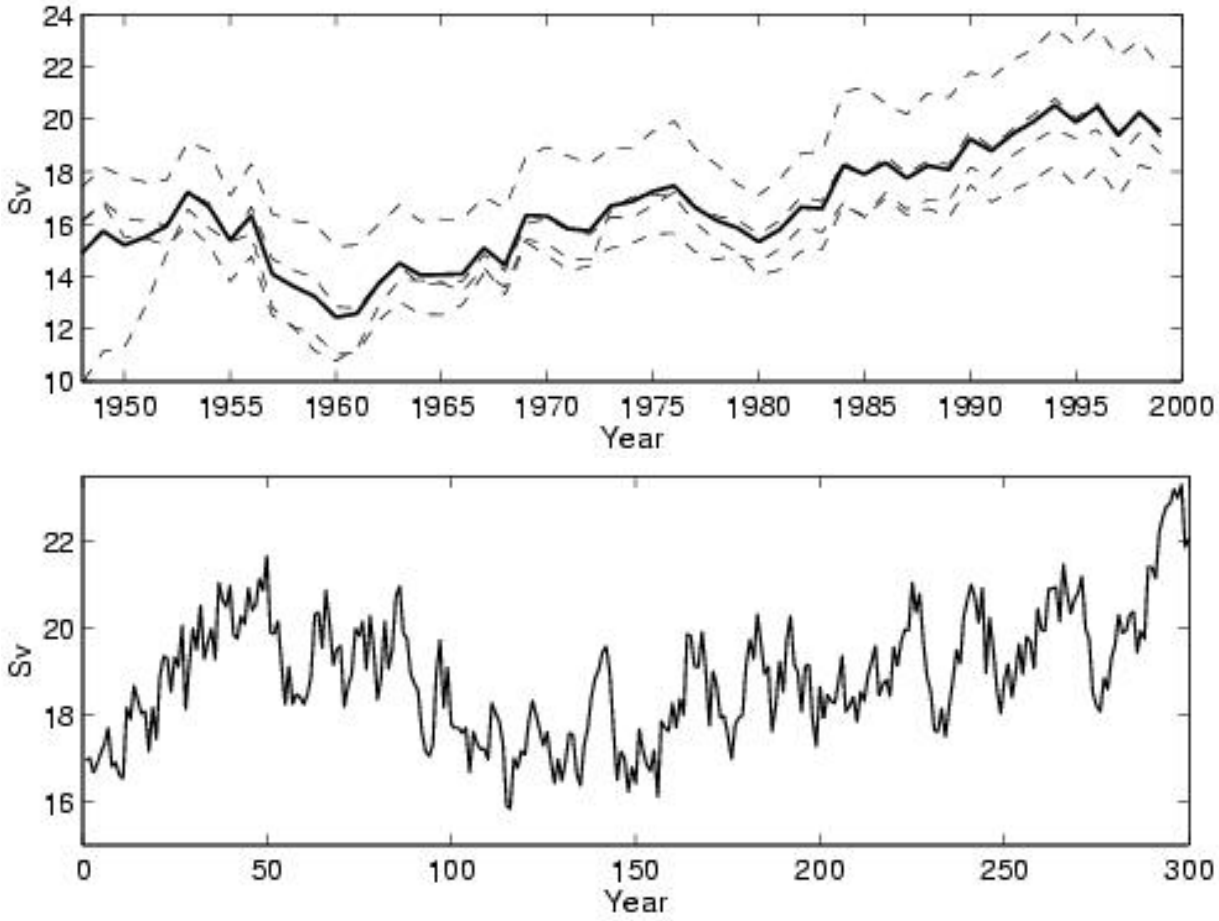


Figure 2.11: Annual mean ATHC from the forced integrations for the the period 1948–1999 (top) and from the 300 years BCM control integration (bottom). The four forced members are shown as dashed lines with the ensemble mean as solid thick line.

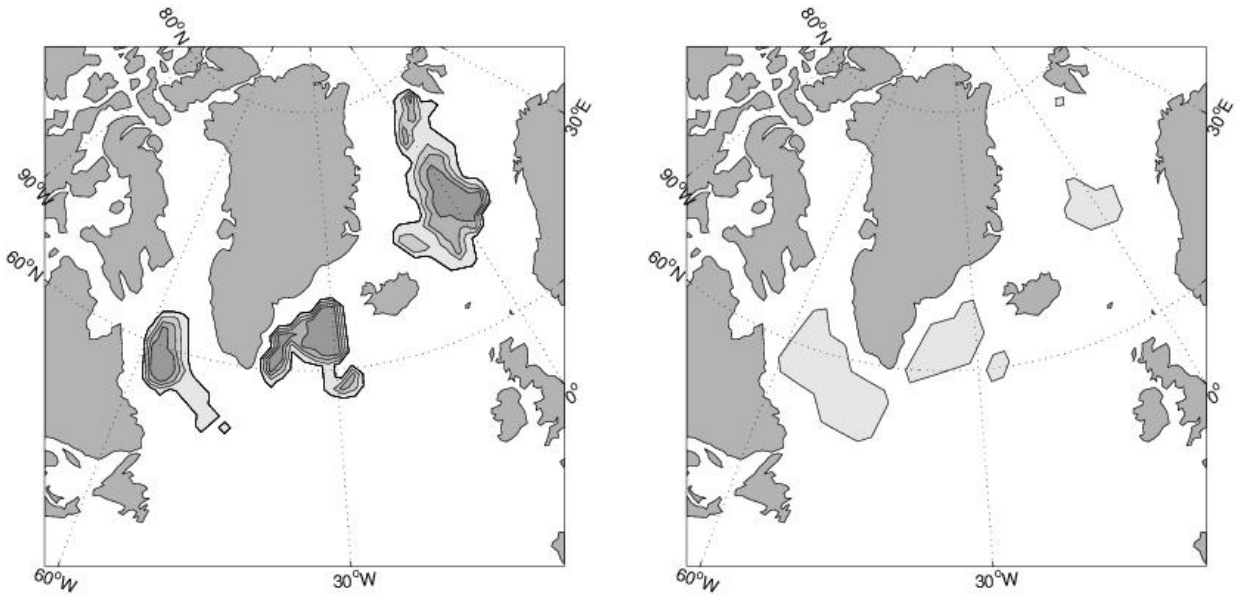


Figure 2.12: The shading shows the areas in the forced simulations (left) and the BCM simulation (right) where the mean February to April mixed-layer dept h exceeds 1500 m at least once during the simulations. In the left panel, the different shadings indicate how many of the realizations that experience deep mixing at each location, going from light grey (one realization) to dark grey (all four realizations).

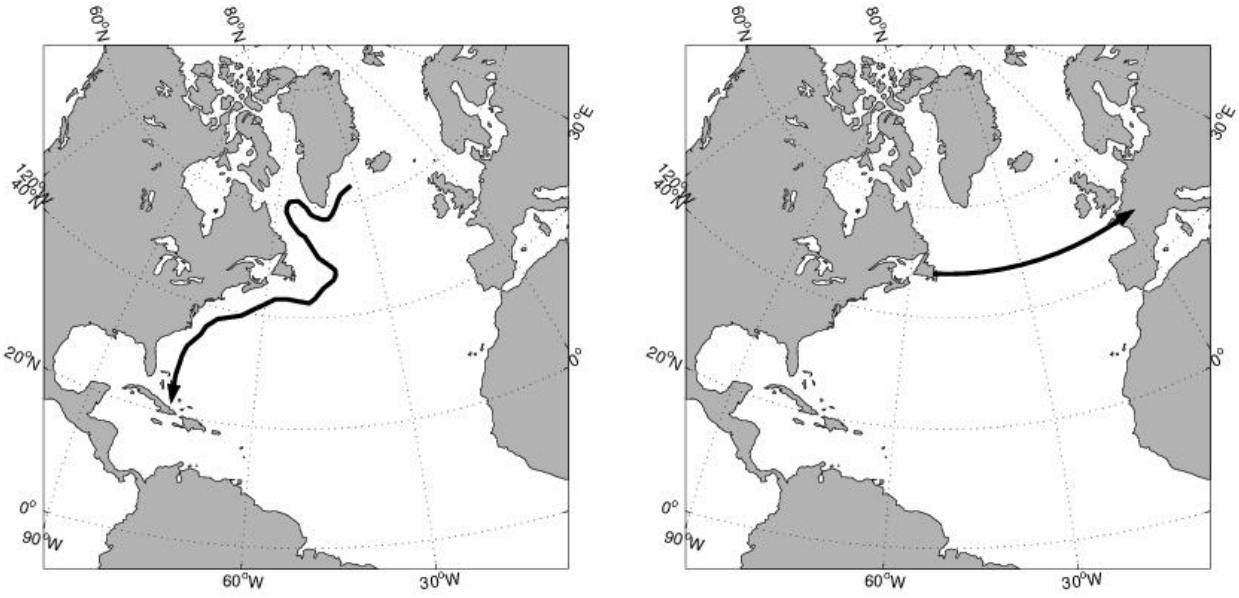


Figure 2.13: The location of the vertical sections used in Figs. 2.14 (left panel) and 2.15 (right panel). The sections end where the arrow is drawn. In Fig. 2.14, the velocity is directed along the section, with positive correlations indicating anomalous southward flows. In Fig. 2.15, the velocity is normal to the section, with positive correlations indicating anomalous northward flows.

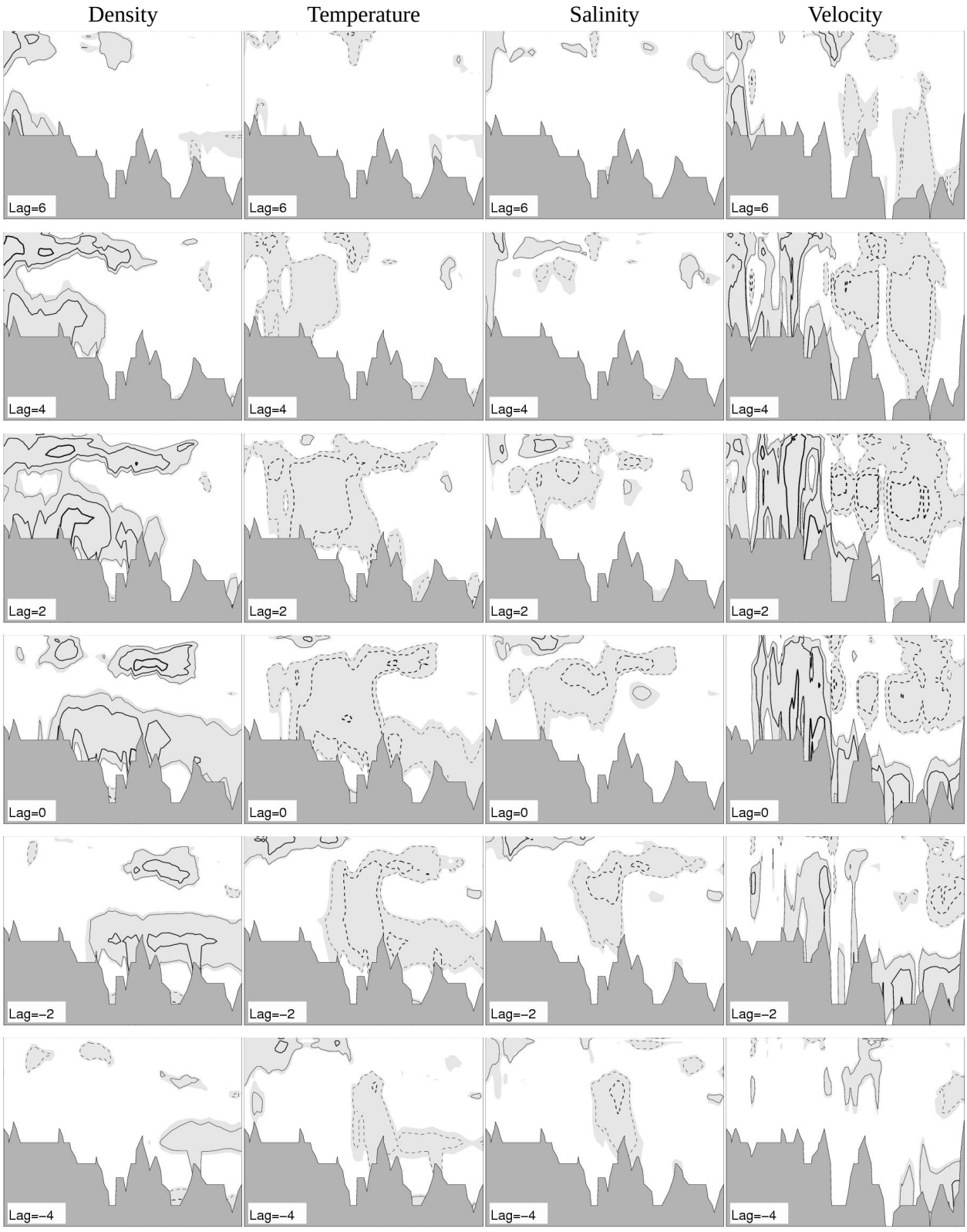


Figure 2.14: Lag correlations between potential density (first column), temperature (second column), salinity (third column), and velocity (forth column) along the section shown in the left panel of Fig. 2.13, and the leading PC of the ATHC from the BCM experiment. The time series have been band-pass filtered with a third order Butterworth filter with cut-off frequencies 1/100 and 1/5 yr^{-1}. The light shaded area indicates the 95% significance level. Positive (negative) correlations are given by solid (dashed) lines. Contour intervals are 0.25, 0.4 and 0.6 with increasing thickness of the contour lines. The lags (in years, positive when ATHC is leading) are indicated in the lower left corner of the panels. Section length (depth) is 7943km (4500m).

In the section along 48°N (Fig. 2.15), the deep southward propagating signal is clearly seen as a plume of cold and fresh water. Furthermore, the velocity anomaly is baroclinic with a northward component over the uppermost 2000 m of the water column in the western part of the basin. The northward component carries warm and saline waters, leading to a negative density anomaly there. The reduced density of the northward surface and near surface flow tends to inhibit convective mixing when these waters reach the sub-polar gyre. This is also shown in the density panels of Fig. 2.14 at lag -2 and -4 years. In this way, the previous mixing event can generally not be maintained over time due to the high temperature of the northward flowing surface waters. At the same time, the accompanying supply of saline surface waters acts as a preconditioning factor for the following mixing event when the surface water becomes sufficiently cooled.

A surface view of the zero time lag correlations between the ATHC and SSS is provieded in Fig. 2.16. Three main features can be identified from these panels: A region of positive correlation between in the sub-polar gyre and the band between 40–50 °N (in all integrations but E77); a region of negative correlation in the Gulf of Mexico; and a large region of negative correlation in the South Atlantic sub-tropical gyre (again in all integrations but E77).

Based on the presented correlation analyses we conclude that remotely sensed SSS have the potential to be used to detect decadal-scale changes in the ATHC.

2.7 WP 2300: Skill Assessment

2.7.1 Possible change in SSS during the 21st century

Figure 2.17.

The global mean change in surface temperature and net precipitation in BCM at a doubling of CO_2 coincides with the mean value of the 19 CMIP-models (Sorteberg, pers. com., 2002), indicating that the applied model system is consistent with other climate models.

At a doubling of CO_2, SST increases by more than 1.5 °C in the sub- tropical gyres, and by more than 0.5 °C over most of the other regions.

The situation for SSS is different. The most pronounced change is found in the Atlantic Ocean. Here the salinity increases by 0.1-0.5 psu, with the strongest increase in the two sub-tropical gyres, and smallest change at the high latitudes and in the equatorial region.

Changes in the SSS are small in the Pacific Ocean, with a tendency of increased SSS off the South-American continent of up to 0.2 psu. There is also a tendency for increased SSS in the Indian Ocean, notably in the south-western part, of 0.1-0.2 psu.

Interestingly, the variability in SSS is reduced in the Indian-Indonesian region in the second half of the century. The variability is otherwise similar to the variability of the current climate state.

2.8 WP 2000: Summary

2.8.1 SSS variability on annual to sesonal time scales for the present day climate system

An overview of the global annual and seasonal variability in the SSS is provided in Table 2.3. The table is based on the findings in Secs. 2.4.2 and 2.4.3.

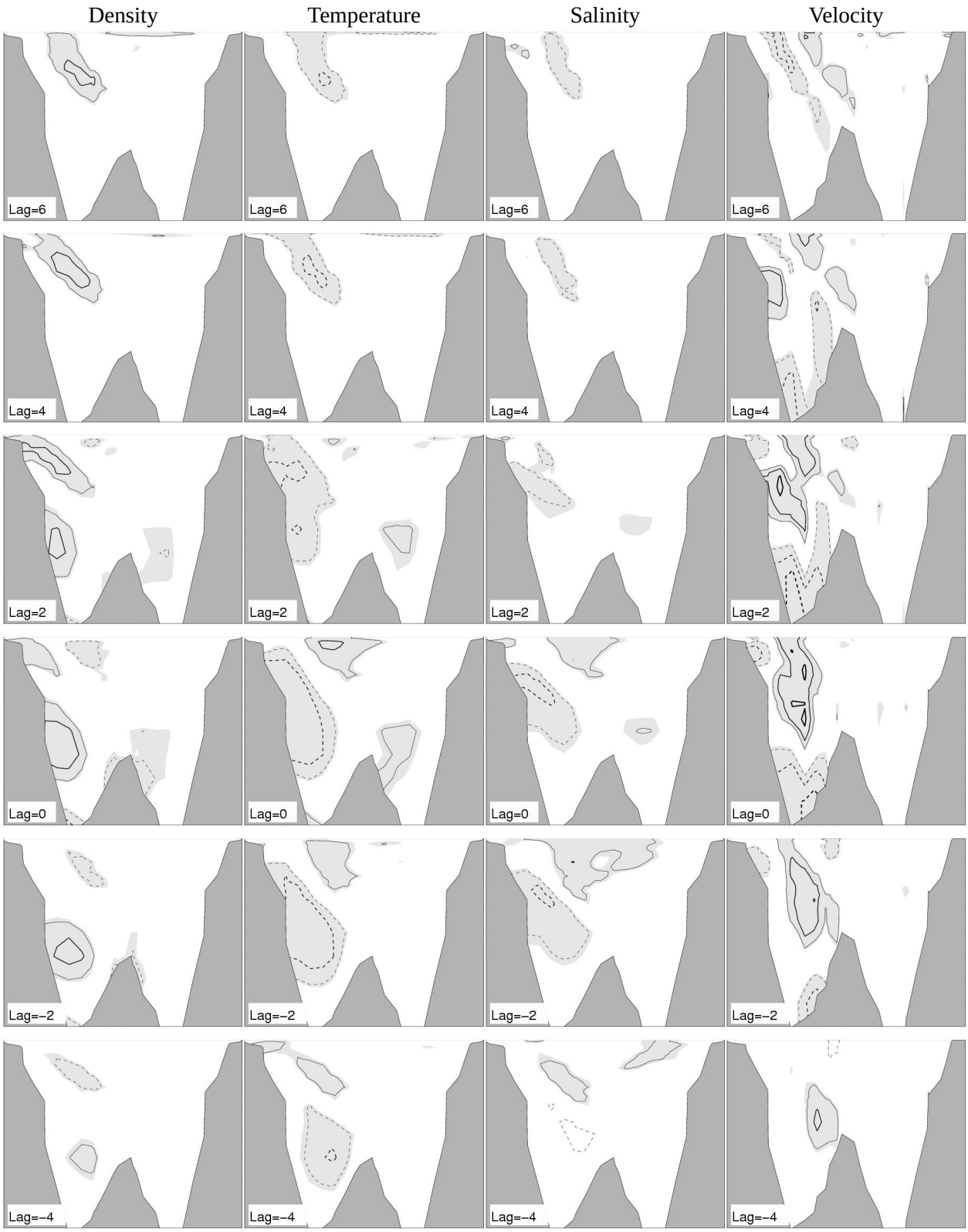


Figure 2.15: As Fig. 2.14, but for the section along 48°N, see Fig. 2.13. Section length (depth) is 3868km (4000m).

Region	Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
1	< 0.4	> 0.5	< 0.5	< 0.4	> 0.4
2	0.1	0.1	01.–0.2	0.4	0.1
3	< 0.1	0.1	$<$0.1	0.1	< 0.1
4	< 0.3	< 0.5	< 0.1	0.2	> 0.5
5	< 0.1	0.1–0.2	0.2	< 0.2	0.1
6	< 0.2	0.1	> 0.3	0.3	0.1
7	< 0.2	0.3	> 0.5	0.1	0.1–0.2
8	< 0.2	0.1	0.2	0.2–0.3	0.1–0.2

Table 2.3: Overview of the simulated variability in SSS for the present day climate expressed in terms of 1 std (psu) for the following regions: 1 Indian-Indonesian-Pacific west of 150 °W, 2 Kuroshio Current, 3 North Pacific, 4 Gulf of Guinea, 5 Equatorial Atlantic, 6 Guyana Current (north of the South-Americal continent, 7 the Caribbean Sea/the Gulf of Mexico, and 8 the region off Newfoundland.

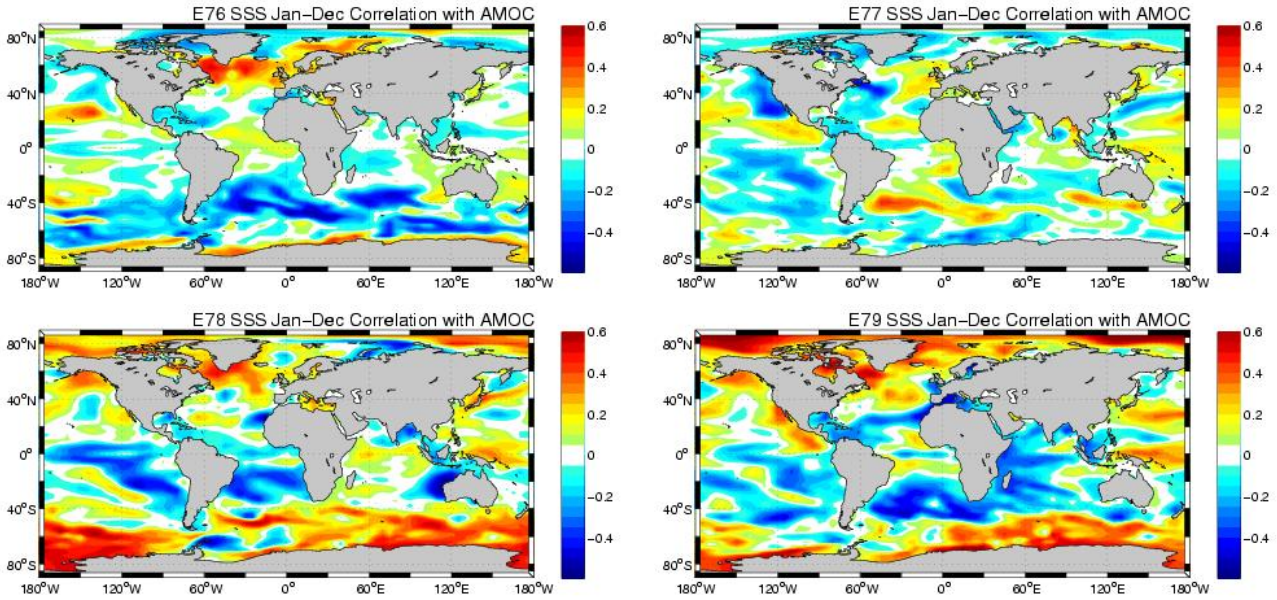


Figure 2.16: Correlation between Atlantic THC and SSS for year 1-20 of E76–E79.

2.8.2 Possible change in SSS at a doubling of CO_2

At a doubling of the atmospheric concentration of CO_2 (Sec. 2.7.1), the most pronounced change is found in the Atlantic Ocean. Here the salinity increases by 0.1–0.5 psu, with the strongest increase in the two sub-tropical gyres, and smallest change at the high (northern and southern) latitudes and in the equatorial region.

Furthermore, changes in the SSS are small in the Pacific Ocean, with a tendency of increased SSS off the South-American continent of up to 0.2 psu. There is also a tendency for increased SSS in the Indian Ocean, notably in the south-western part, of 0.1–0.2 psu.

2.8.3 SSS-variability linked to NAO-variability

It is found in Sec. 2.5.2 that during high NAO-years, the SSS field are the higher than normal SSS in the North Sea (exceeding 0.3 psu), in the eastern part of the sub-tropical region (up to 0.1 psu), in the polar waters north of Newfoundland (up to 0.2 psu), in the Guyana Current region and south of the Hispaniola islands (about 0.1 psu), and the lower than normal SSS in the waters just north of the Gulf Stream (about 0.1 psu). The situation is near reversed for low NAO-years, hence the difference in SSS between high and low NAO-years is about twice of the above stated values.

2.8.4 SSS-variability linked to the Atlantic Thermohaline Circulation

A potential relationship between the Atlantic SSS and the Atlantic THC is discussed in Sec. 2.6.1. It is concluded that there is a region of positive correlation between the SSS and the Atlantic THC in the sub-polar gyre and the band between 40–50 °N (in three out of four realisations). Furthermore, regions of negative correlations are found in the Gulf of Mexico and over a large region in the South Atlantic sub-tropical gyre (again in three out of four realisations). The correlations are large, typically up to 0.6. It is concluded that remotely sensed SSS have the potential to be used to detect decadal-scale changes in the ATHC.

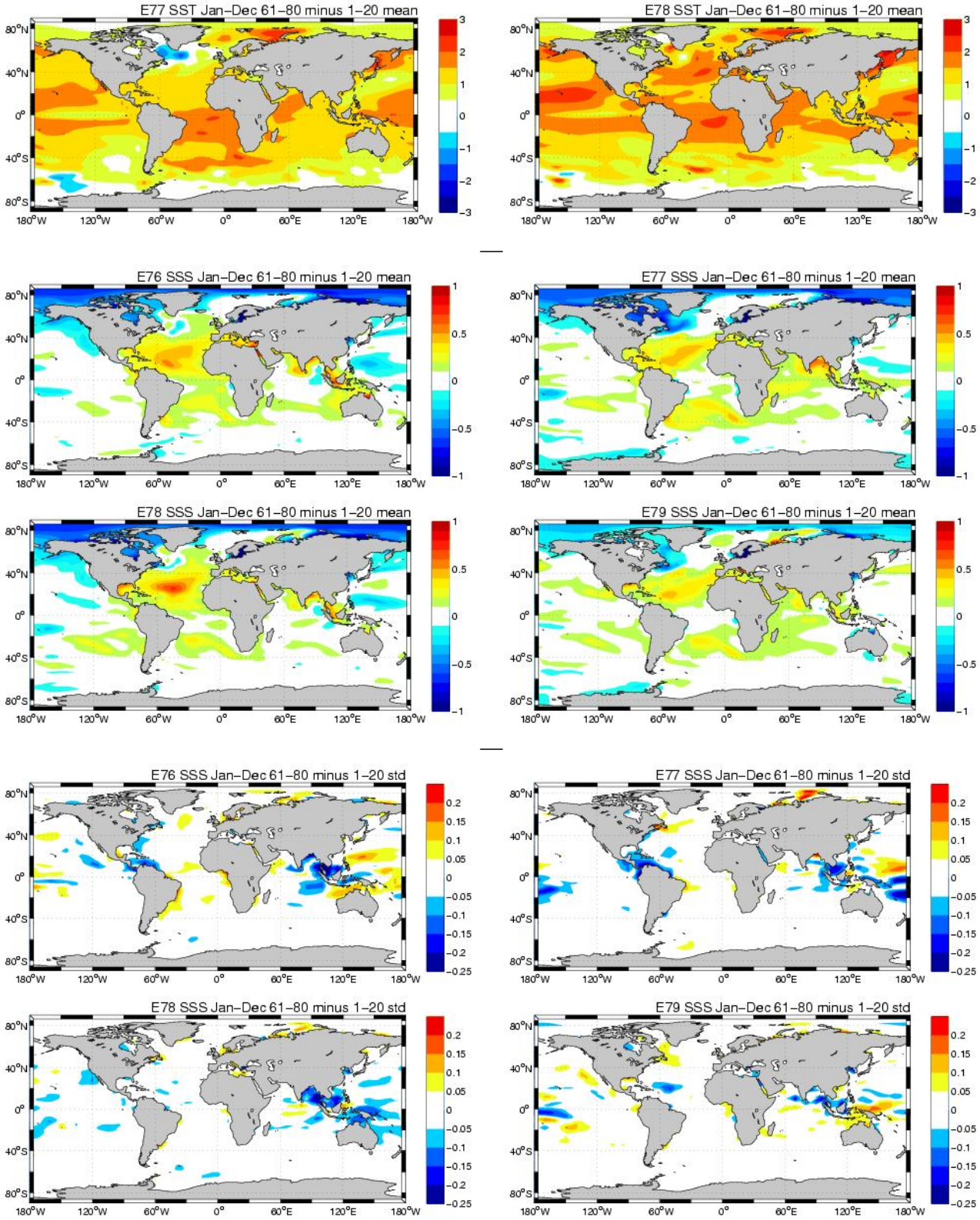


Figure 2.17: Row 1: Simulated change in annual mean SST at doubling of CO_2. Row 2–3: Simulated change in annual mean SSS at doubling of CO_2. Row 4–5: Corresponding change in the SSS std.

Chapter 3

WP3000- Operational Oceanography

The need for high quality predictions of marine parameters has been well identified. E.g., during recent years, offshore oil-exploration activites have expanded off the continental shelfs to deeper waters. Drilling and production of oil and gas at depths of 2000 meters or more are ongoing at several locations. This has introduced a need for real time forecasts of oceanic currents which in some cases may have severe impact on the safety related to drilling, production and critical operations. Sustainable exploitation of marine resources are becoming increasingly important, e.g. commercial fisheries and fish farming. In future fisheries management systems, information about marine parameters such as nutrient and plankton concentrations, and pollutants, will be increasingly important for accurate monitoring and prediction of fishstocks. Thus, there are needs for operational monitoring and prediction of both physical and biological marine parameters.

The optimal approach for an ocean forecasting system is usually based on integrated use of both high resolution ocean- and ecosystem models and observations of physical, chemical and biological variables. This integration can best be done using data assimilation techniques. The DIADEM/TOPAZ configuration of the NERSC model suite is supplied with Earth Observations (EO) of open ocean variables, Arctic sea ice cover and the marine biology. From the assimilation we, in turn, assess and quantify the relative impact of these EO in a pre-operational prediction system. The main focus is associated with short-term (weeks) prediction capabilities including impact of state estimation and control of the evolving model state. Further, the real time processing and flow of observational data must be developed and maintained.

The data assimilation system uses the Ensemble Kalman Filter (EnKF) developed by *Evensen* (1994). This is a filter method which means that the model is integrated forward in time and every time there are new measurements available these are used to reinitialise the model before the integration continues. This reinitialisation is determined as a weighted linear combination of the model prediction and the measurements. The weights are the inverses of the error covariances for the model prediction and the measurements. The EnKF uses ensemble integrations to predict model error statistics which is used to update the model solution whenever observations are available. The ensemble consists of model states that are integrated forward in time in parallell. The EnKF is a fully multivariate assimilation method that gives dynamically consistent model estimates both for linear and nonlinear models.

The impact analyses, which is further addressed and reported under the two main workpackages, is undertaken using the NERSC model suite which is composed of a physical ocean model (OGCM), an ecosystem model and a sea ice model. Three classes of input sources are interfaced to this model suite, notably atmospheric forcing fields, remote sensing data and in-situ data. The data assimilation module, which then combines the observation data fields and the model data, forms the central element of the suite and can be applied to all three models, either individually or in combination.

Both global, regional and local model domains can be selected with accurate bathymetry and land boundaries, and based on a grid generation tool the model domain will also always be defined with best model resolution in the area

of largest interest. Moreover, the system introduces a further downscaling capability by nesting of high resolution models for the regional areas of particular interest.

In this workpackage the focus is to assess the impact of existing and planned Earth Observations. This includes the impact of assimilating satellite derived sea surface temperature (SST), sea level anomalies (SLA), marine geoid information, sea surface salinity (SSS), ice concentration and thickness and ocean colour.

In the first part of the project we used the DIADEM system and worked on examining the effect of assimilating SST and SLA into the Miami Isopycnic Coordinate Ocean Model (MICOM). Following the transition to HYCOM in the TOPAZ project the remaining work was done using the HYCOM based TOPAZ system.

In Section 3.1 we will give an overview of the MICOM and HYCOM model systems as used at NERSC. Next, in Section 3.2 some background information is given on the DIADEM and TOPAZ projects which form the basis for the work in SIREOC. In Section 3.3 we provide some background on the EnKF assimilation method. Thereafter, the assimilation of different EO are discussed in individual sections before a final synthesis is given.

3.1 The Ocean Modelling System

During the ongoing TOPAZ project the new Hybrid Coordinate Ocean Model (HYCOM) has been adapted in the model and data assimilation system developed in the DIADEM project. The model is an extension of the previous MICOM model where a more general vertical coordinate is used. All interfaces between the HYCOM dynamical module and the modules for thermodynamics, ice dynamics, atmospheric forcing fields, and the assimilation schemes have been recoded. Since this was a major upgrade of the model system we also restructured the coding of the overall system to make it more flexible and robust.

In this section we will present a general discussion of the HYCOM model and its setup for the TOPAZ and SIREOC projects, and explain the major differences and advantages compared to the original MICOM model.

3.1.1 Model domain

The model domain used in the real time forecasting system is shown in Figure 3.1. This is based on a grid configuration where the North and South Poles are mapped to locations near the Equator in the central Pacific Ocean. This results in a grid with fairly uniform grid resolution in the Atlantic and Arctic Oceans. As an example, the grid spacing reduces with a factor 0.5 from the central Atlantic to the North Pole. In the standard spherical coordinate, the same reduction is obtained at 60 °N, and then there is also a singularity at the North Pole. This new grid is very different from the one used in the DIADEM project which had enhanced resolution in the Gulf Streem extension and in the Nordic Seas. The access to significantly larger computers has now made it possible to aim for relatively high resolution in all of the Atlantic, and this makes the TOPAZ model system generally applicable for the whole Atlantic and Arctic Oceans. The final grid will maintain higher resolution in the Gulf Stream region and have overall much higher resolution than in the DIADEM project. However, we will sacrifice some resolution in the Nordic Seas since this area will be modelled using a nested high resolution regional model.

3.1.2 Vertical mixing processes in MICOM

The main difference between HYCOM and MICOM is related to a change of the vertical discretization. MICOM used isopycnic layers, i.e. layers of constant density, for all but the surface mixed layer which was treated as a bulk mixed layer with variable temperature and salinity. The mixed layer in MICOM interacts with the atmosphere through the specification or computation of fluxes for freshwater, heat and momentum. Further, it interacts with the deeper layers through vertical mixing processes. These are parameterized as convection, mixed layer entrainment and detrainment, and diapycnal mixing. The convection is a process that mixes the mixed layer water with one or several of the layers below whenever the mixed layer becomes denser than the lower layers. This can happen due to, e.g., a strong cooling of often saline mixed layer water. Thus, convection acts to stabelize the water column and results in a deepening of the mixed layer. In the Greenland and Labrador Seas the mixed layer may, due to convective processes, exceed 1000 meter during the winter time deep water formation. The mixed layer entrainment is a process where the mixed layer deepens due to, e.g., increased turbulence in the mixed layer. Water from the deeper layers are then entrained into the mixed layer where it is mixed with the water previously held by the mixed layer. Detrainment is the reverse process where the mixed layer depth is reduced, and thus the mixed layer water must be disposed in the layers below the mixed layer. Finally the diapycnal mixing is a process accounting for the vertical mixing which is always present in the ocean, although below the surface mixed-layer it is several orders of magnitude less than the horizontal mixing.

3.1.3 Vertical coordinate system in HYCOM

Before we can discuss the parameterization of the vertical mixing processes in HYCOM we need to explain the new hybrid vertical coordinate used. Note first that the vertical movement of water masses can be divided into a

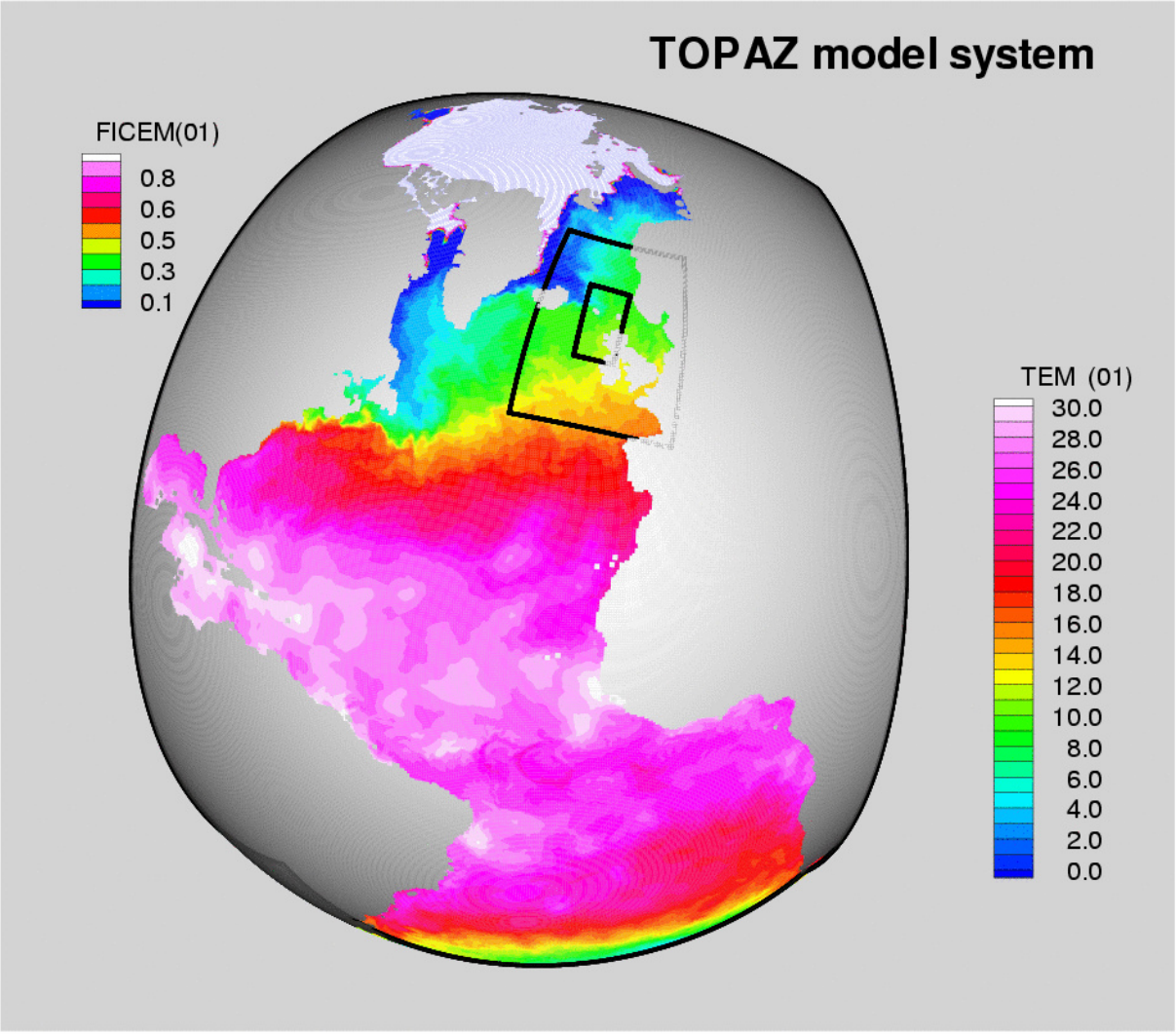


Figure 3.1: The model domain used in the TOPAZ real time experiment. The plot shows surface temperature and ice concentration.

Lagrangian movement where a coordinate surface is moving with the water in the vertical, as in MICOM, and the movement of water through the coordinate surface as is done in all models with a fixed vertical coordinate system, e.g. z-level models and σ-coordinate models. HYCOM includes both representations of vertical movement of water masses. This allows for a combination of material coordinate surfaces as in MICOM with fixed surfaces as in z-level models and σ-coordinate models. The current algorithm exploits that all layers have an assigned reference density as in MICOM. However, whenever a layer thickness becomes zero because this light water does not exist in the particular vertical column, this layer is used as a vertical level coordinate within the mixed layer. Further, this level coordinate is located in depth according to a predefined rule. The algorithm results in a stack of levels located from the surface and downwards with a specified resolution. Thus, the model allows for arbitrary high vertical resolution near the surface by adding a sufficient number of light (and therefore always massless) layers to the model. An example is given in Figure 3.2 showing a vertical section from South to North across the Faroe–Iceland ridge. Here it is seen how the isopycnal layers are reverting to level coordinates in the upper part of the ocean. Note in particular how the lightest isopycnal layers are transferred into z-levels when moving north towards cooler and denser upper ocean water masses.

To summerize, based on the number of layers defined and their chosen reference densities, the layers will distribute themselves in the vertical, starting with isopycnal layers from the sea floor and upwards towards the surface. The

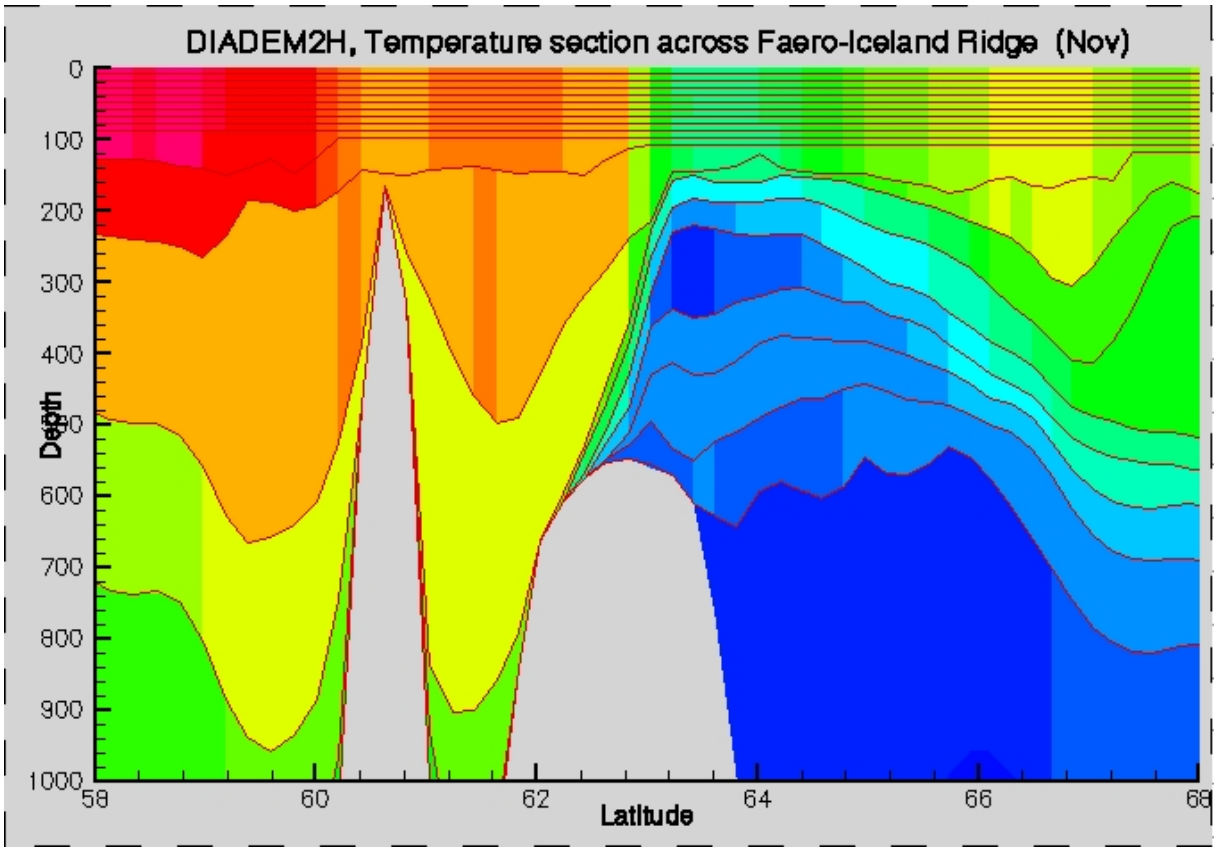


Figure 3.2: A vertical section from the HYCOM model across the Faroe-Iceland ridge. Note the transition from isopycnal layers to levels when going northward, and the high resolution on top of the ridge where there are strong gradients in the density of the water masses. The ventilation of isopycnal layers with the mixed layer is also clearly illustrated.

layers with reference densities lighter than the existing water masses in the present water-column will be stacked from the surface and downwards with a specified vertical resolution, and used as level coordinates.

3.1.4 Vertical mixing processes in HYCOM

Vertical mixing in HYCOM is a combination of cabbeling and restoration processes and the explicitly prescribed mixing. Normally, the prescribed mixing exceeds other mixing processes by several orders of magnitude.

The horizontal advection of layer thicknesses, tracers and momentum is computed using the same algorithms as in MICOM. However, the advection of layer thicknesses in the continuity equation will introduce a vertical movement of the layer interfaces, also among the level coordinates near the surface. Further, horizontal diffusion of temperature and salinity in an isopycnic layer may lead to a deviation from the reference density. Thus, a new routine called “hybgen” which is short for hybrid coordinate generator is used every time step to restore the correct location of the coordinate surfaces. Among the isopycnal layers in the deep ocean there is a restoration towards reference densities, an effect called cabbeling, where a slight amount of water is mixed between adjacent layers to restore the reference densities. This is normally a correction needed to correct for a small deviation in density resulting from the diffusion of temperature and salinity in a layer. (The nonlinearity of the equation of state, implies that the mixing of two water masses with different T–S properties but the same density, may result in a new water mass with a different density). For the level coordinates near the surface, water is moved/mixed between layers to restore the layer interfaces to their predefined locations in depth. Thus, vertical advection in the level coordinates is parameterized by a horizontal

advection of tracers in a layer, and the layer thickness, followed by the restoration of the level coordinates. This process is designed to conserve temperature, salinity and momentum when water is moved between layers.

The prescribed vertical mixing in HYCOM is vastly different from the algorithm used in MICOM. While MICOM has separate routines/algorithms for computing convection, diapycnal mixing, and mixed layer entrainment/detrainment, HYCOM solves for all these processes using a vertical turbulence closure scheme developed by *Large et al.* (1994). The scheme computes the vertical mixing coefficient over a vertical column in the model, and takes into account the effect of wind stirring inducing mixed layer turbulence and additional mixing parameterization for processes, such as, internal wave breaking, Richardson dependent vertical current shear, salt fingering and double diffusion. A background vertical mixing coefficient ensures the presence of a low diapycnal diffusion in the deep ocean. The scheme is using an algorithm to compute the vertical diffusivity in a water column, and thereafter a one-dimensional diffusion equation is solved for temperature, salinity and momentum. Every time the vertical mixing has been applied the "hybgen" routine is called to restore all layers/levels to their reference values.

3.1.5 Time stepping in HYCOM and MICOM

The time stepping algorithm can be summarized as follows:

```
do n=1,...
   call forcing_fields      ! reads/computes atmospheric forcing fields
   call continuity_eq       ! advects the layer thicknesses
   call temp_saln_advection ! advects temperature and salinity in each layer
M  call convection          ! explicit convection algorithm
M  call diapycnal_mixing    ! explicit computation of diapycnal mixing
   call momentum_eq         ! solves the momentum equation
   call barotropic_eq       ! solves the barotropic equation
   call thermodynamics      ! computes all theromdynamic surface fluxes
H  call vertical_mixing     ! computes vertical mixing from KPP scheme
H  call hybgen              ! restore layer thicknesses using the hybrid grid generator
   call icemodel            ! dynamic ice model and ice advection
   call nesting_conditions  ! apply nesting boundary conditions
M  call mixed_layer         ! update mixed layer thickness
enddo
```

The routines specific for HYCOM and MICOM are marked with an H or an M in the first column. Thus, the major change when going from MICOM to HYCOM is that the `convection`, `diapycnal_mixing`, and `mixed_layer` routines are replaced by `vertical_mixing` and a call to `hybgen`. The two models are using many of the same subroutines and a next release of the model is planned where an option can be set to chose between MICOM and HYCOM.

3.1.6 Nesting of regional models

Open boundary conditions and nesting in ocean circulation models are considered more as an art than real science. The main problem is that for a model with open boundaries the number of boundary conditions is dependent on the structure of the flow field penetrating the boundary. There are actually four cases which must be considered, i.e., inflow or outflow and for each of these one can have supersonic or subsonic velocities. To avoid dealing with the problem of exactly specifying the boundary conditions in a "proper" nesting scheme, most approaches use some kind of boundary relaxation towards the outer model solution. This results in what one normally would call the one way nesting schemes where the boundary conditions of the regional model are relaxed towards the output from a coarser large scale model. For the slowly varying variables, i.e., baroclinic velocities, temperature, salinity and

layer interfaces, this is a fully appropriate way to include the boundary conditions. For the barotropic variables the relaxation approach requires careful tuning to avoid reflection of waves at the open model boundaries. In HYCOM the barotropic model is a hyperbolic wave equation for pressure and vertically integrated velocities. Following an approach outlined by *Browning and Kreiss* (1982, 1986), it is possible to compute the barotropic boundary conditions exactly while taking into consideration both the waves propagating into the regional model from the external solution and the waves propagating out through the boundary from the regional model. The scheme has been tested extensively and has shown no problematic behavior yet. In addition, it also made it fairly simple to include the tidal forcing on the barotropic mode.

The practical implementation of the nesting scheme is based on communication through files stored on disk. The outer model dumps the solution interpolated to grid points in the relaxation zone of the regional model every 6 hours. For the baroclinic mode which only changes slowly this is considered to be high resolution in time. It should also be sufficient for the barotropic mode as long as the outer model does not contain tides. The regional model reads the files every six hours and uses interpolation in time to specify the relaxation boundary conditions at every time step. The communication between the grids is general and there is no restriction on the relative orientation or resolution of the grids. An example from the nested regional model is given in Figure 3.3.

3.1.7 Tidal boundary conditions

It is also possible to include tides in the regional nested models. The tides are then specified as a barotropic forcing on the open boundaries for the regional model. This is fairly easy using the nesting boundary conditions explained in the previous section. The data set used has just been released by the University of Texas and is based on several years of altimeter data collected by the TOPEX satellite.

3.1.8 Coupling of the ecosystem model

The implementation of the ecosystem model takes a completely new approch with HYCOM. In MICOM the ecosystem variables were solved for on a separate vertical grid which was required to properly resolve the MICOM bulk mixed layer. This required a lot of additional code and complicated the implementation. In HYCOM, the generalized vertical coordinate allows for a much simpler approach to be used. Since HYCOM already has high vertical resolution in the upper part of the ocean the ecosystem can be solved for on the same grid as is used by the physical model. Thus, the ecosystem model has been recoded and stripped for a large part of the code which is not needed. The approach taken has been to discretize the biology on the same grid as the physics, and then adapting the physical model routines for all transport and mixing processes, e.g. advection, horizontal diffusion and KPP vertical mixing. After the hybrid regridding biological tracers are moved accordingly between the model layers. This means that the only parts left of the original biological model are the subroutines and functions needed to compute the interactions between the various biological compartments. This is now set up as a function call done for each model grid point, an approach which makes it simple to replace one biological model with another.

The new NERSC biochemical model system is illustrated in Figure 3.4; if a new ecosystem formulation is to be added to HYCOM, additional code must be added as indicated by the dark shaded boxes in the figure. The light shaded boxes are model independent; note in particular that the routines for the advection, mixing and regridding are model independent and that these need not be changed for a new ecosystem model. The figure indicates three different ecosystem models, one three component model (EVA85) based on *Evans and Parslow* (1985), one 11 compartment model (FDM02) based on *Drange* (1994a) and a 7 compartment model (SCH02) which is the model developed by M. Schartau and coworkers at AWI for use in TOPAZ. Note that it is easy to switch between these models (just define a flag in MODEL.CPP, compile and run), and it is also a relatively simple task to add new ecosystem formulations at a later stage.

To test the above HYCOM-ecosystem setup, we implemented a simple 3-component biochemical model consisting

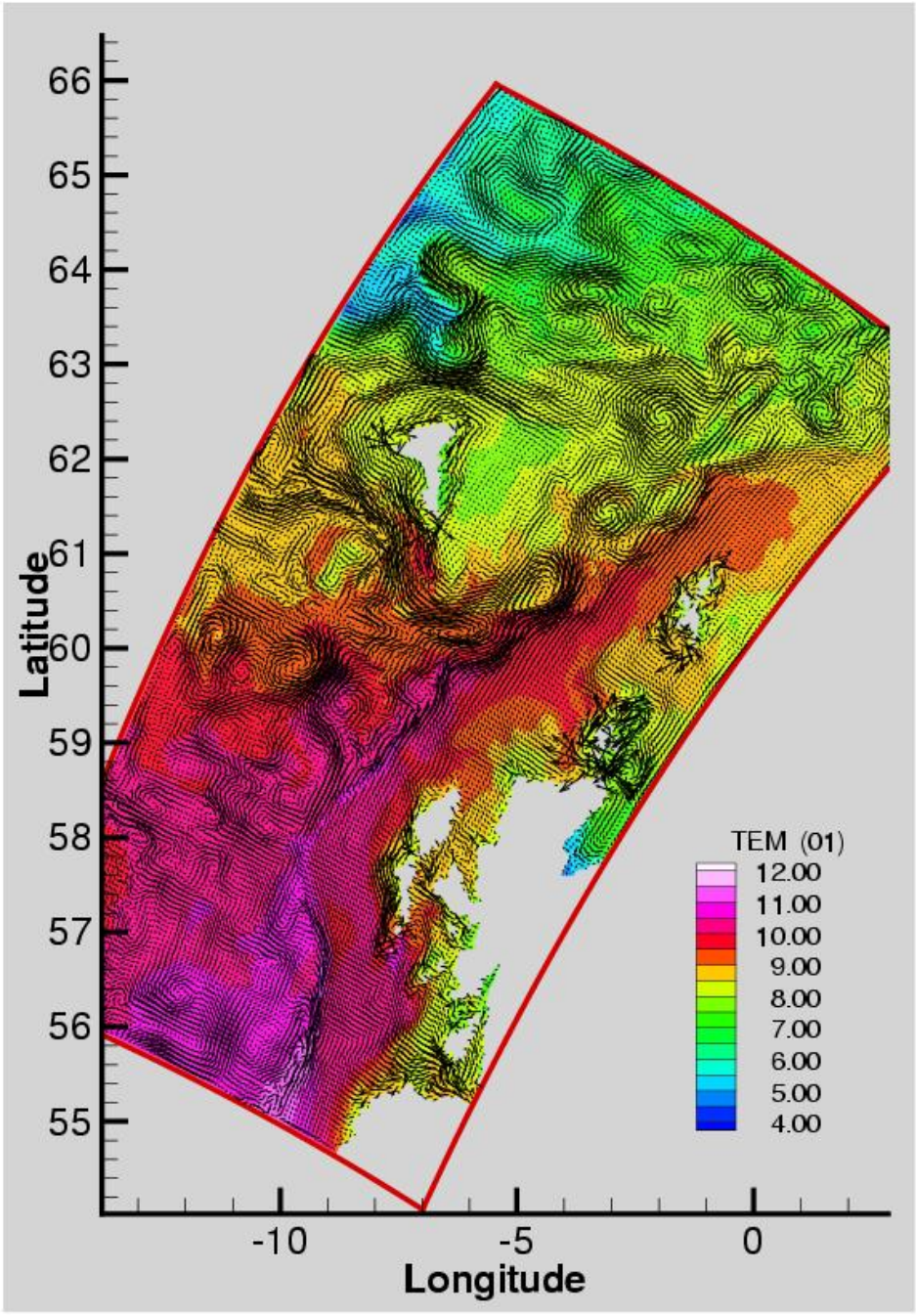


Figure 3.3: An example from a regional high resolution (2 km) model for the Faroe Shetland Channel, nested into the Atlantic system. The plot shows sea surface temperature and surface current velocities (every third vector shown).

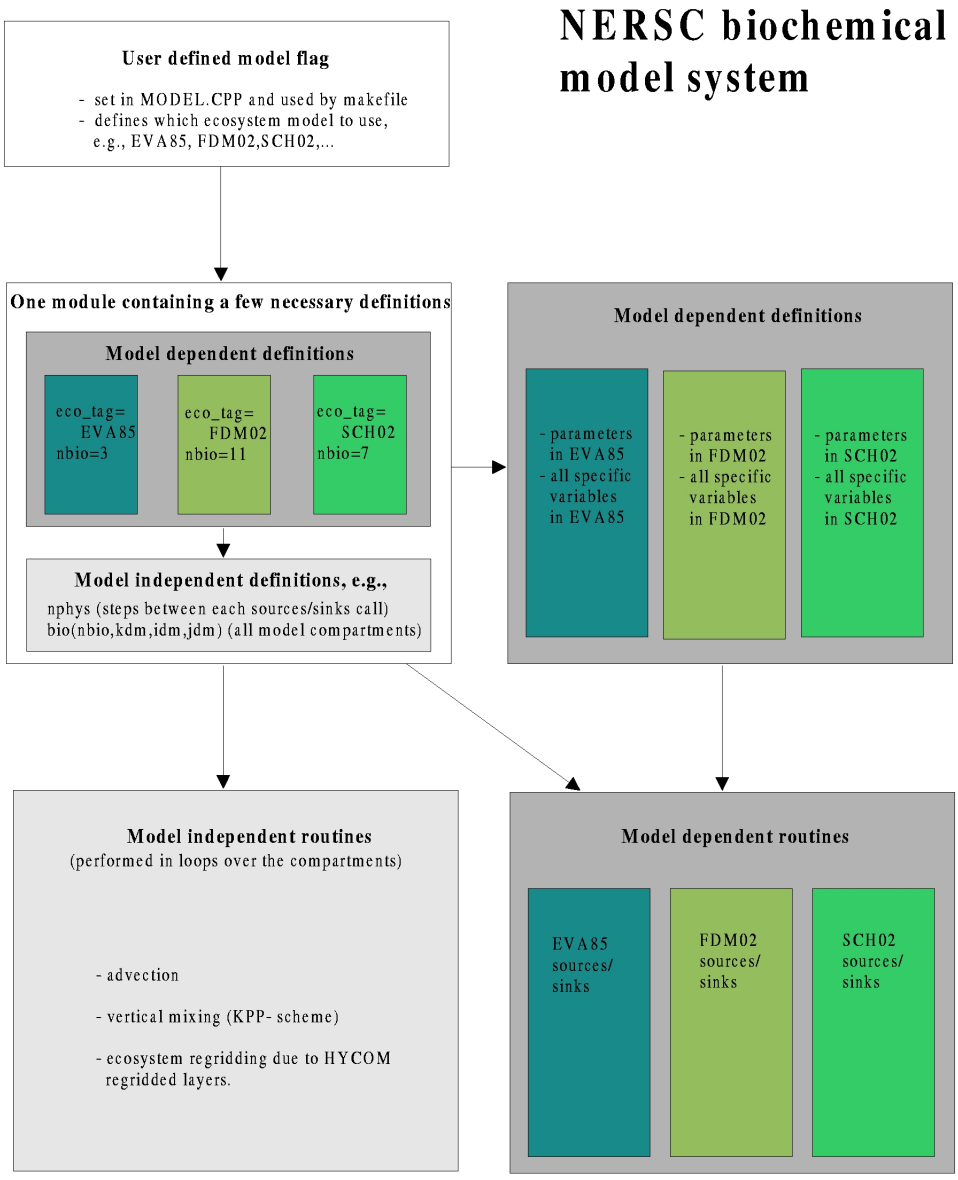


Figure 3.4: The NERSC biochemical model system is programmed such that one easily can switch between different formulations for the ecosystem, and new formulations can easily be added. The user defines which model to compile and run through a flag in the file MODEL.CPP, which is used by the makefile. If a new model is to be implemented in the HYCOM-ecosystem system, the number of compartments and a short character model flag is defined in the module containing the necessary definitions. Then, all the compartments will be stored in the variable 'bio'. All additional specific variables and model parameters are gathered in separate modules for the particular model. The most important property of the system, is that the routines for advection, mixing and the ecosystem regridding are model indepentent, just defining loops over the number of tracer variables. Note also that these routines are not dependent on the model dependent definitions in the top-right box. The only model specific routine, is the call for the sources and the sinks. Three existing biochemical formulations are shown in the figure (EVA85, FDM02, SCH02), but any number of formulations can be added. When a user has made a choice for a model flag in MODEL.CPP, the correct formulation will be compiled and run.

of nutrients, phytoplankton and zooplankton (see *Evans and Parslow* (1985) for a description of the model). First, we run the model as a pure tracer model, i.e., without the sources and sinks, to test the general routines for the advection, mixing, and ecosystem regridding. Note again that these routines are model independent, and that they will work for any biochemical formulation coupled to HYCOM. After the debugging/validation of these routines, we included the sources and sinks in a short 3 month spinup experiment. The simple ecosystem behaved as expected: when light increases during spring, phytoplankton starts to bloom while nutrients decrease, and then a zooplankton bloom follows. Note that this model is too simple to be realistic, and its relevance is only academic when coupled to a fully 3-dimensional ocean circulation model. However, it provides a good tool to understand "typical" ecosystem dynamics, and it was valuable to implement a simple formulation during the debugging/validation process. The 11 compartment FDM model (*Drange*, 1994a) has also been implemented and made as simple as possible in the new system, and the coupling with SCH02 is ongoing.

The MICOM-FDM coupled model system was used in the hindcast experiment described in Task 3300. The FDM model is described in detail in *Drange* (1994a, 1996). It was originally based on the chemical *Peng et al.* (1987) model, the seven compartment ecosystem model by *Fasham et al.* (1990) and *Fasham* (1993) and the 3-dimensional extensions as described in *Sarmiento et al.* (1993). Some further improvements of the ecosystem model was done by *Drange* (1994a, 1996), e.g., the bacteria equation was slightly reformulated to ensure a non-accumulating flow of nitrogen and carbon through the bacterial compartment. Also, equations for the evolution of total dissolved inorganic carbon and alkalinity were included, in addition to the seven compartments explicitly modelled in *Fasham et al.* (1990). Further, carbon based compartments of organic dissolved material and particulate matter were included explicitly in the model. In this formulation, the Redfield carbon to nitrogen ratios of DOC/DON and POC/PON are free floating and dependent on the flow of carbon and nitrogen through these compartments, as described by the model equations.

The following eleven compartments are explicitly modelled; phytoplankton (P), zooplankton (Z), bacteria (B), nitrate (N_n), ammonium (N_r), dissolved organic nitrogen - DON (N_d), dissolved organic carbon - DOC (C_d), detritus as particulate organic nitrogen - PON (D_N), detritus as particulate organic carbon - POC (D_C), dissolved inorganic carbon - DIC (C_T) and alkalinity (A_T).

Since nitrogen is generally regarded as the limiting nutrient for phytoplankton growth, nitrogen is used as the basis model currency. Also, this allows us to partition the total primary production into new production fuelled by nitrate, and regenerated production fuelled by ammonium. In the photic zone, the exchanges of nitrogen and carbon can be described by simple dynamical relations between the biochemical model compartments. A schematic illustration of the photic zone ecosystem is shown in Figure 3.5. A nutrient regeneration model is used below the photic zone.

3.1.9 Coupling of the sea ice model

The dynamical sea ice model is fairly simple to couple to the system since it only requires the atmospheric and upper ocean inputs to compute the internal ice-stress which is used to compute the ice drift velocities. The dynamic ice model determines how the ice moves in response to stress from the atmosphere and ocean as well as internal ice stress. The internal ice stress is determined by the rheology model. The dynamic ice model uses the Elastic–Viscous–Plastic (EVP) rheology of *Hunke and Dukowicz* (1997). The EVP model gives similar results as the Viscous–Plastic rheology of *Hibler* (1979), but the EVP formulation has some numerical advantages. For instance, the EVP formulation allows for practical use of an explicit numerical scheme.

The ice thermodynamic model determines the atmospheric and oceanic heat fluxes and calculates the ice growth rate. There are two different ice models used at NERSC. The first type is relatively simple, using a single ice category to describe the ice in a grid cell. The treatment of heat conduction in the ice is also relatively simple, where the heat capacity of the ice is ignored, this is also known as the 0-layer formulation of *Semtner* (1976). This thermodynamic model is described in detail by *Drange and Bleck* (1996).

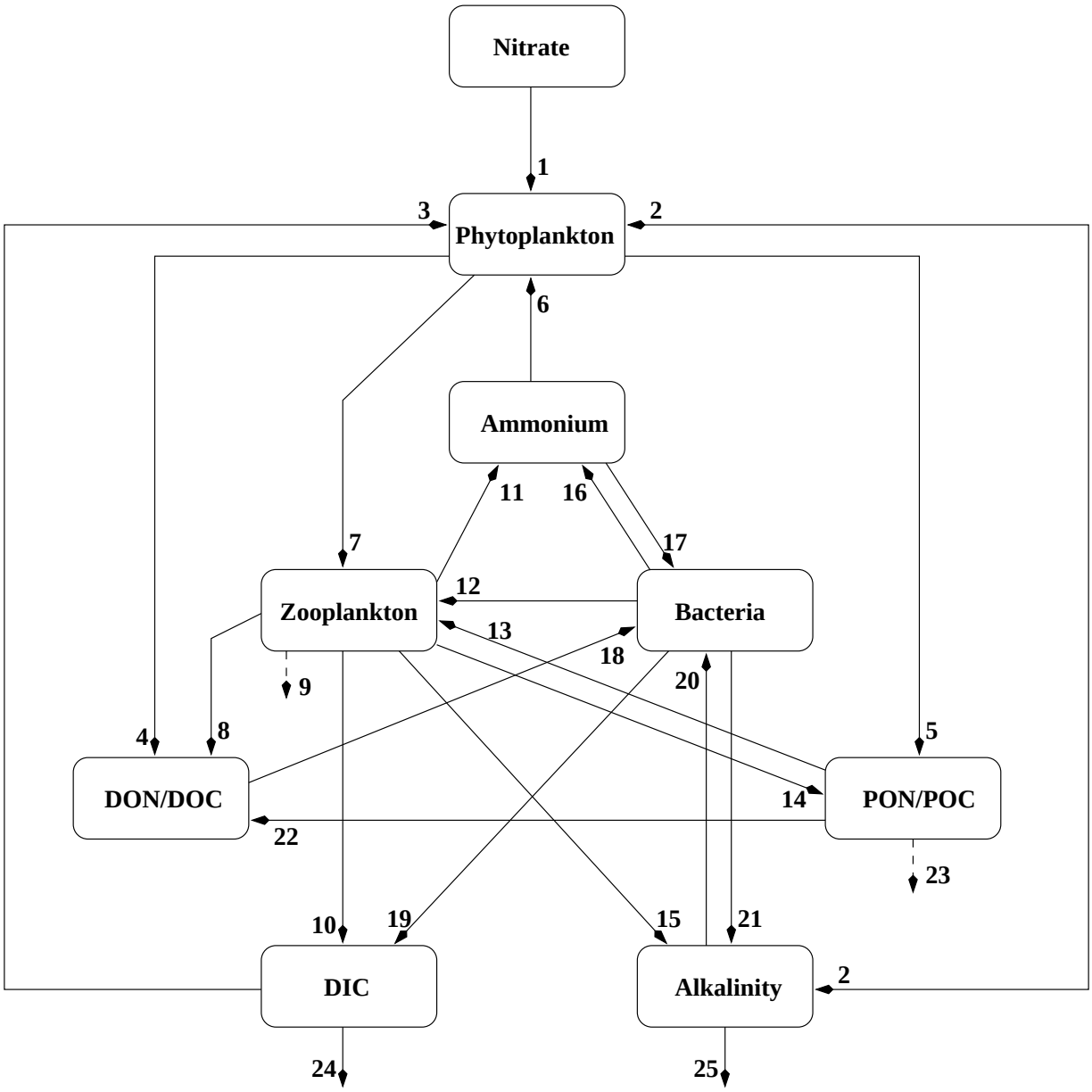


Figure 3.5: A schematic representation of the photic zone biochemical system. Primary production is fuelled by nitrate (1) and ammonium (6), and reduced by phytoplankton exudation (4) and mortality (5). Zooplankton graze on phytoplankton (7), bacteria (12) and detritus (13) with prescribed efficiencies, and the fecal pellets represent a sink to detritus (14). Further, zooplankton losses (excretion, mortality, messy ingestion, Redfield balance term, grazing by higher predators, etc.) are split up as sinks to ammonium (11), dissolved organic matter (8) and an export term representing large sinking particles (9). The bacteria take up ammonium (17) and dissolved organic matter (18), and excrete ammonium (16). Particulate matter may sink (23) or may be broken down to dissolved matter (22). The concentration of dissolved inorganic carbon decreases during photosynthesis (3) and during formation of $\mathrm{CaCO_3}$ (24), and increases during zooplankton and bacterial excretion (10, 19). Alkalinity increases during nitrate fuelled and decreases during ammonium fuelled primary production (2). Further, zooplankton and bacterial excretion increase the alkalinity (15, 21), while bacterial uptake of ammonium and the formation of $\mathrm{CaCO_3}$ decrease A_T (20, 25).

Another ice thermodynamic model named ICESTATE (developed in the MAST–II ICESTATE project) is also implemented at NERSC, and is presently being tested in longer model integrations. It gives improvement in many areas compared to the first type. The major improvements are the ability to describe the distribution of ice thickness in a model grid cell, and a more realistic treatment of the vertical heat conduction in the ice. The description of the ice thickness distribution also allows a kinematic ice model, describing the ridging and rafting of ice (*Thorndike et al.*, 1975).

3.2 The TOPAZ system

The DIADEM/TOPAZ monitoring and forecasting system is currently being developed with support from projects funded by the European Commission and national research councils. The overall long term objective is:

> **to develop an operational ocean and ecosystem montoring and prediction system for the North Atlantic, Nordic Seas and Arctic Ocean, using state of the art numerical model tools and data assimilation methodologies.**

The focus has so far been on the development of methodologies for assimilation of remotely sensed data into ocean and ecosystem models. In contrast to other ongoing developments of real time assimilation systems we have chosen to work with rather sophisticated assimilation methodologies which also predict multivariate error covariance statistics.

The implementation of the assimilation system has been completed and applied in hindcast experiments for validation purposes and further calibrated for use with different data types. Currently, the assimilation system is operated in real time and provides forecasts of physical ocean parameters.

In this section the DIADEM/TOPAZ monitoring and prediction system will be described in some detail.

3.2.1 Background

The DIADEM/TOPAZ system is being developed to meet the needs from future users of marine parameters. It involves both the implementation and validation of state of the art ocean circulation models and marine ecosystem models, and the development of novel data assimilation methodologies. The system development has been centered around two ongoing European Commission funded projects, i.e. the DIADEM and TOPAZ projects, which are briefly explained next. The model domain used for the DIADEM/TOPAZ prediction system is shown in Figure 3.1. The grid is created using a conformal mapping of the poles to two new locations using the algorithm outlined in *Bentsen et al.* (1999). Within the TOPAZ project there is also a downscaling to coastal seas by nesting of high resolution regional models (illustrated by the frames in the figure).

DIADEM project

The EC MAST-III project DIADEM has focussed on the implementation of a data assimilation system for the North Atlantic and the Nordic Seas. The project involves partners from six European countries working with ocean and ecosystem modeling, data assimilation and processing of remotely sensed observations.

The major objective of the project was to implement and demonstrate novel sophisticated data assimilation methods such as the Ensemble Kalman Filter (EnKF) (*Evensen*, 1994), the Ensemble Kalman Smoother (EnKS) recently proposed by (*Evensen and van Leeuwen*, 2000), and the Singular Evolutive Extended Kalman Filter (SEEK) (*Pham et al.*, 1998), with the Miami Isopycnic Coordinate Ocean Model (MICOM) developed by *Bleck et al.* (1992) and a 3-dimensional implementation of the ecosystem model by *Fasham et al.* (1990) which has been coupled with MICOM by *Drange* (1994b).

The assimilated data were remotely sensed sea level anomalies and sea surface temperatures, used in the MICOM model, and ocean color data in the ecosystem model.

The assimilation systems have been applied in hindcast experiments for validation purposes and calibrated for use with different data types. The use of so called advanced methods have introduced the possibility of performing a multivariate and physically consistent analysis with statistical covariance functions which vary in space and time. This allows us to extract a maximum amount of information from observations of surface quantities. The real time

data flow from existing satellite observing systems and their capabilities in providing observations that can be used with the data assimilation system in an operational mode are currently being evaluated.

The DIADEM project led to a prototype of a validated pre-operational monitoring and prediction system for the North Atlantic and the Nordic Seas. Estimates of error statistics such as correlation scales and cross-correlation between different variables which are crucial information in all data assimilation systems and are produced routinely. The DIADEM project established, for the first time, an operational capability for coupled physical and ecosystem models in the North Atlantic and the Nordic Seas, where satellite information is assimilated using advanced data assimilation methods.

ToPAZ project

A new project ToPAZ has recently been funded by the European Commission under the fifth Framework Program. ToPAZ extends the developments of DIADEM to a more realistic operational system. The project has a strong focus on end-user requirements and has established a particular link with off-shore oil industry operating in deep waters along the Atlantic Margin north of Scotland. A major objective of ToPAZ is to establish a generic operational prediction system for ocean currents which will be applied and demonstrated for end users operating along the Atlantic Margin.

To meet the end users needs it is necessary to introduce nested regional models to allow for downscaling to very high resolution in the target areas where mesoscale processes must be properly resolved. In addition, the MICOM model used in DIADEM has been replaced by the recently developed Hybrid Coordinate Ocean Model (HYCOM) HYCOM wep page: http://panoramix.rsmas.miami.edu/hycom/.

This model integrates the properties of the isopycnal MICOM model for the deep ocean with a level model for the surface boundary layer. Thus it is designed to work equally well for the coastal shelf areas as for the deep ocean.

The data assimilation system developed in DIADEM will be further extended to use new data types. A capability will be developed for assimilation of *in situ* data from the ARGO program as well as additional remote sensing products such as ice concentration (SSMI), ice thickness (Cryosat), sea surface salinity (SMOS) and the improved sea level anomaly data which can be derived with the new geoid from the GOCE mission.

With the inclusion of a nesting capability and the assimilation of both *in situ* data and data from a variety of satellite sensors, the ToPAZ project will develop a state of the art and flexible operational ocean prediction system. The model system has been designed to be easily extendible to other geographical areas including the global domain and it allows for nesting of an arbitrary number of regional high resolution models.

3.2.2 Participants

The DIADEM project involves seven European partners with different responsibilities in the project:

1. The Nansen Environmental and Remote Sensing Center (NERSC) is coordinating the project and is supplying the model systems used by all partners. NERSC is responsible for model validation and an implementation of the EnKF with MICOM and the ecosystem model.
2. The Institute for Marine and Atmospheric Research, University of Utrecht (IMAU), is implementing the EnKS with the physical model.
3. The Universite Joseph Fourier, Laboratoire des Ecoulements Geo-physiques et Industriels, (LEGI) is developing a SEEK Filter with MICOM and the ecosystem model.
4. Calibration of the model parameters in the ecosystem model is done by the Alfred Wegener Institute (AWI).

5. Gridded fields of sea level anomalies and sea surface temperature data are delivered by Collecte Localisation Satellites (CLS).

6. Gridded ocean color data from SeaWifs are processed and delivered by the Joint Research Centre (JRC).

7. The HALO Laboratory for Oceanic and Atmospheric Sciences is maintaining the project web-page, where the forecasts from the assimilation systems are displayed during the real time operation of the system, and has also developed a marine information system which is used for the data management in the project.

The partnership in TOPAZ is a subset of the DIADEM consortium with major activities distributed among partners as follows:

1. NERSC is coordinating the project and maintaining the HYCOM model system. The operation of the nested high resolution models are done at NERSC and the development of an assimilation system for ice is the responsibility of NERSC.

2. LEGI is mainly responsible for the development of an assimilation system for *in situ* observations.

3. CLS is developing and operating a processing capability for ARGO *in situ* data and the remotely sensed products used in the project.

4. AWI will work on ecosystem model development and validation.

3.2.3 Hind-cast experiment

A hind-cast experiment has been carried out over a three months period using a low resolution version of the physical model. In the experiment gridded fields of sea surface temperature and sea surface anomaly (produced by CLS) were assimilated. The gridded data were available every 10 days. Three assimilation methods were used: the EnKF, the EnKS, and the SEEK filter. All three methods provided a realistic prediction for the model error statistics which were consistent with the innovation sequence. Further, the assimilation methods could take the multivariate statistics into account and provided a realistic analysis where the whole model state was updated from the surface measurements. As an example, the impact of an SST measurement is to correct the mixed layer temperature in the model but in addition it will also introduce an update of the mixed layer thickness and the location of the thermocline. Similarly the SLA data contain more information about the mesoscale structures, e.g., the rings in the Gulf Stream extension, which could now be updated consistently with the thermocline depth etc.

The model with its limited horizontal resolution has a tendency to a too northward location of the Gulf Stream separation. However, an interesting result was that the assimilation system was capable of correcting the location of the Gulf Stream axis and separation point. This is an important result since a successful assimilation system must be capable of correcting obvious model deficiencies. The results from hindcast experiments with the MICOM model has been published in *Brusdal et al.* (2003).

A twin experiment was carried out using the assimilation system with the marine ecosystem model. It proved that surface observations of ocean color or Clorophyll could be used to control the evolution of the marine ecosystem to some extent. When testing the assimilation system with real data from SeaWifs a similar conclusion could be drawn. The remotely sensed ocean color data do have an impact on the accuracy of the ecosystem predictions, however, the accuracy of the data are currently rather poor and a greater impact is expected using data from more accurate future sensors. The results from hindcast assimilation experiments with the ecosystem model coupled to MICOM are published in *Natvik and Evensen* (2003a,b).

Further, we have developed an implementation for assimilation of remotely sensed ice data which is described in *Lisæter et al.* (2003).

3.2.4 Real time operation

A real time prediction experiment was initiated November 2000. It has been ongoing until August 2001 with funding from the DIADEM project, and is now continued with funding from TOPAZ. The system was initially spun up into real time, first in simulation mode, but in a step wise process the assimilation system for the physical model was introduced with assimilation of sea level anomalies from TOPEX and ERS and Reynolds SST data. During spring 2001, the marine ecosystem model was coupled to the system and following a spinup simulation we started an experiment with real time assimilation of SeaWifs ocean colour data which lasted until end of June 2001.

The system has now been operated with a weekly assimilation cycle since November 2000.

A major conclusion is that the system can be operated in real time and provides resonable forecast results using available remotely sensed observations.

3.2.5 Summary

The DIADEM/TOPAZ system provides a prototype of a pre-operational monitoring and prediction system for the North Atlantic and the Nordic Seas. The system is based on sophisticated modeling and data assimilation tools and is set up for real time or near real time operation, and now assesses the real time data flow as well as the impact of the remote sensing products on the predictions.

The real time operation of the system has proved to be feasible and relies on the availability of remote sensing products in near real time, and atmospheric forcing fields from the meteorological forecasting centers.

Finally it should be stated that the DIADEM and TOPAZ projects comply with and contribute to the plans of international programs such as GODAE and EuroGOOS. The system developed has similarities with the other major initiatives in GODAE and will in many respects be complementary to these. Further, the system is one of the major initiatives contributing to the EuroGOOS task teams, in particular *the Atlantic Task Team* by developing an assimilation system for predicting the ocean circulation in the Atlantic, *the North West Shelf Task Team* by introducing high resolution regional models for physics and ecosystem covering the European shelfs, and finally the *Arctic Task Team* by the focus on ice modeling and assimilation of ice variables in the Arctic. The project has also realized the importance of merging these different areas into one model system to allow for a proper representation of the interactions between the different seas and oceans.

3.3 The Ensemble Kalman Filter

To assimilate observational data into a model system, it is important to have a knowledge of the errors present in the model and observations. In Kalman Filtering the needed statistics are the error covariance matrices of the observations and the model state. Describing the error covariances correctly is therefore a crucial component to the performance of sequential data assimilation schemes.

The Ensemble Kalman Filter (EnKF; *Evensen*, 1994) uses an ensemble of model states to estimate the model error statistics. It was originally proposed as an alternative to the Extended Kalman Filter (EKF) for solving strongly nonlinear problems. The EKF uses linearized equations to separately solve for the error covariance matrix, while the EnKF members are advanced using the original model equations. An advantage of the EnKF approach is that it retains the effect of nonlinear model behavior on the error covariance matrix. In an EnKF approach, few assumptions are made, and complex correlation patterns will evolve according to the evolution of the ensemble members. The simplified closure of the EKF has, on the other hand, been shown to be problematic when used in some non–linear models (*Evensen*, 1992). A brief explanation of the EnKF is given below, for a more thorough description of the EnKF see *Evensen* (1994); *Burgers et al.* (1998).

Let $\boldsymbol{\psi}_i^f \in \mathbb{R}^{n\times 1}$ be the n-dimensional model forecast of ensemble member number $i \in \{1, 2, ..., N\}$. This forecast evolves in time from a best estimate at time t_k, using the model

$$\boldsymbol{\psi}_i^f(t_{k+1}) = \boldsymbol{g}\left[\boldsymbol{\psi}_i^a(t_k)\right] + \boldsymbol{\beta}_i(t_k)\,. \tag{3.1}$$

Here $\boldsymbol{\beta}_i \in \mathbb{R}^{n\times 1}$ is a random component drawn from a $N(\mathbf{0}, \boldsymbol{\sigma}_{\boldsymbol{\beta}})$ distribution, and represents the effect of model errors on the evolution of the ensemble members. The covariance $\boldsymbol{\sigma}_\beta \in \mathbb{R}^{n\times n}$ will have to be specified based on knowledge about the model errors. The operator $\boldsymbol{g} : \mathbb{R}^{n\times 1} \rightarrow \mathbb{R}^{n\times 1}$ can, as already mentioned, be a nonlinear function of the model state. The individual ensemble members in equation (3.1) form a Monte Carlo approach for solving the Fokker–Planck equation (*Jazwinski*, 1970), which describe the evolution of the probability density function for a state $\boldsymbol{\psi}$. In order to infer the error evolution of the model state, knowledge is needed of the "truth". As the true model state is unknown, some assumptions must be made regarding its properties. In the EnKF, the best estimate of the truth is represented by the ensemble mean state. It follows that the model state error covariance used in the EnKF is that given by the ensemble covariance. Then, at any time, an estimate of the model state error covariance matrix can be computed from the ensemble of model states as

$$\boldsymbol{P}^f \approx \boldsymbol{P}_e^f = \overline{\left(\boldsymbol{\psi}_i^f - \overline{\boldsymbol{\psi}^f}\right)\left(\boldsymbol{\psi}_i^f - \overline{\boldsymbol{\psi}^f}\right)^T} \tag{3.2}$$

where $\overline{\boldsymbol{\psi}}$ is the ensemble estimated mean state, and the overbar denotes the expected value. In other words, the ensemble covariance matrix is taken to be representative of the error covariance matrix.

At the time observations are available an analysis is computed. The observations $\boldsymbol{d} \in \mathbb{R}^{m\times 1}$ have an associated uncertainty $\boldsymbol{\epsilon}$, and an observation error covariance matrix $\boldsymbol{R} = \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T$, where the observation error covariance matrix must be based on prior knowledge of the observation errors. Let $\boldsymbol{H} \in \mathbb{R}^{m\times n}$ be a linear operator that transforms the model state to the observation space. Then the analysis update is given by the following variant of the traditional Kalman filter equation (*Jazwinski*, 1970);

$$\boldsymbol{\psi}_i^a = \boldsymbol{\psi}_i^f + \boldsymbol{P}_e^f\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{P}_e^f\boldsymbol{H}^T + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{d}_i - \boldsymbol{H}\boldsymbol{\psi}_i^f\right) = \boldsymbol{\psi}_i^f + \boldsymbol{K}_e\left(\boldsymbol{d}_i - \boldsymbol{H}\boldsymbol{\psi}_i^f\right)\,. \tag{3.3}$$

$\boldsymbol{K}_e$ is called the Kalman Gain, and is given as

$$\boldsymbol{K_e} = \boldsymbol{P}_e^f\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{P}_e^f\boldsymbol{H}^T + \boldsymbol{R}\right)^{-1}\,. \tag{3.4}$$

Equation (3.3) tells how each ensemble member is updated using the model error covariance matrix estimate (3.2), and the error covariance matrix of the observations. Special notice should be taken of the observation vector $\boldsymbol{d}_i$

used in equation (3.3); as indicated by its subscript it is different for each ensemble member. This is because the observations need to be perturbed to get an analysis error covariance matrix consistent with the original Kalman Filter. As shown by *Burgers et al.* (1998) the analysis with $\boldsymbol{d}_i$ taken from a $N(\boldsymbol{d}, \boldsymbol{R})$ distribution gives the following analysis covariance matrix;

$$\boldsymbol{P}_e^a = (\boldsymbol{I} - \boldsymbol{K_e H})\, \boldsymbol{P}_e^f \,, \tag{3.5}$$

which is consistent with the covariance of the analysis in the traditional Kalman Filter. Without perturbation of the original observations $\boldsymbol{d}$, the analysed covariance would be systematically underestimated, an effect which could lead to so–called filter divergence. Filter divergence means that the error estimate of the model in the EnKF is too small, and the analysis will have little impact on the model.

Recently there has been a debate on the drawbacks of using perturbed observations in the Ensemble Kalman Filter, connected to noise introduced by the perturbations. This has led to the development of variants of the EnKF which do not need perturbed observations, e.g. *Anderson and Robinson* (2001); *Whitaker and Hamill* (2002). For our application we note that the effects of random perturbations should only be significant for very small ensemble sizes, and should probably be of lesser importance for our 100 member ensemble.

3.4 WP 3100: Assimilation of SLA and SST data

Currently, we are using the Ensemble Kalman filter and the HYCOM model to investigate the impact of assimilating SLA and SST data on the predictability of the ocean physics. The TOPAZ near real-time forecasting experiment covers the full Atlantic ocean. It started in January 2003. Within this experiment, the importance of the SLA and SST data has been studied.

The weekly assimilation cycle for the TOPAZ system is a two steps sequence. First, an ensemble of one hundred realistic system states is propagated by the HYCOM model forced by the ECMWF atmospheric data (downloaded from the ECMWF ftp site). We assume that the uncertainties in this atmospheric data are the main cause of model errors and therefore we perturb them by random noise having adequate statistical properties both in space and time. The second step updates the resulting forecast ensemble with remotely sensed SLA and SST data (also downloaded from public ftp sites). The EnKF analysis scheme provides an ensemble of analysed states accounting for errors of both the model forecast and the data by a least squares method as outlined above. The analysed ensemble returns to the first step to be further integrated one week forward. All results of both steps for all the physical ocean variables at all depths will be freely downloadable on the TOPAZ website: http://topaz.nersc.no.

To study to which extent the model captures the information in the data and keeps the model on a ”track” close to the satellite data, we have plotted remote sensed observations of SST and SLA in Figure 3.6, the difference between the observed and the modeled SST in Figure 3.7 (left panel), the impact of their assimilation on the model upper layer temperature in Figure 3.7 (right panel) and the evolution of both observed and modeled zonal SSH averages during seven assimilation cycles in Figure 3.8. In the left panel of Figure 3.7 we find the largest discrepancies in three regions: the Gulf Stream region where the model shifts the warm current Northwards compared to the satellite data, the Labrador Sea where the model produces too cold water and the North-West African coastal region where the model produces too warm water compared to the satellite data. The right panel of Figure 3.7 shows the impact of the assimilation of SST observations on the model top layer temperature. It generally fits well the features of the left panel and fills the gaps caused by the cloud cover in agreement with the ocean physics. The discrepancies are dramatically reduced in most parts of the domain, and especially in the North Atlantic which indicates that the assimilation of SST brings the model results closer to the observations. We note that both panels do not fit perfectly since the estimation accounts for measurement errors. These errors have two main sources: 1) the relative instrumental precision and 2) representativity error since a surface image does not represent exactly an average of the upper layer (of about ten meters depth). The SST data update the ocean temperature in the upper layers where the variability is higher, but they also have an impact on the other physical variables according to the multivariate correlations in the ensemble. It is important to notice that the EnKF assimilation is not affected by the inevitable presence of clouds. The corresponding gaps are filled according to the ensemble statistics that are produced by the physical model and therefore the interpolation is automatically in agreement with the system equations.

If we compare the CLS SST to the CLS SLA data in Figure 3.6 it appears that more mesoscale structures can be retrieved from the SLA data provided by CLS. Indeed, the SLA remote sensing images are not affected by the cloud cover and clearly exhibit the mesoscale eddies and meanders.

The assimilation of remote sensing SLA data is more complicated than that of SST in the sense that they can not be linked directly to ground data. The observed sea level is relative to the satellite altitude which depends on the gravity field. Therefore, the altimetry data are provided as anomalies (SLA) with respect to a constant reference level. In order to associate these anomalies to the SSH modeled by HYCOM, we need to know the reference sea level (or Mean Sea Surface MSS) which also depends on the local gravity field in the water column. Therefore, the use of the SLA data is strongly linked to the knowledge of the earth geoid, which is presently relatively inaccurate. An error in the choice of the MSS leads to a measurement bias, which in practice cannot be distinguished from a model bias.

In addition the SSH is computed as a diagnostic variable in MICOM and HYCOM, and it is not entirely clear how the model MSS should be computed to be consistent with the MSS used when deriving the SLA data.

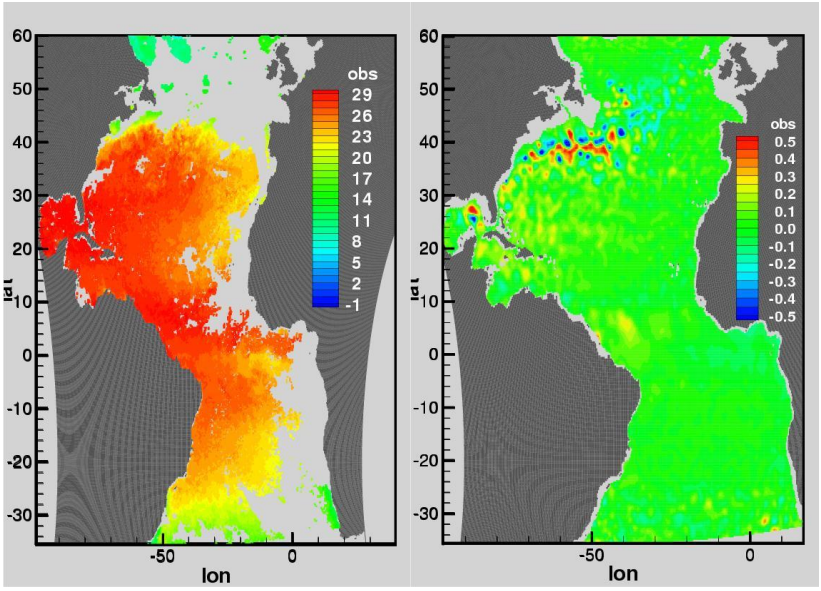


Figure 3.6: Left: Observed sea surface temperature on July 29th 2003, CLS SST data, the holes correspond to unobserved cloudy regions. Right: Observed sea level anomalies for the same day, CLS data.

As a first alternative one could rely entirely on the model and define the reference sea level by a long term model run. In this first alternative the observed anomalies are relative to the model MSS and it is possible to get all the mesoscale structures contained in the SLA data well assimilated into the model. However, it would not be possible to correct for biases contained in the model MSS, for example a wrong path of the Gulf Stream.

The other alternative, using a best possible MSS constructed from observations, will allow for correction of biases, but one could expect some technical problems if this MSS is inconsistent with the way the MSS is computed in the model. Anyway, this is the approach we have used in SIREOC, and it has turned out to work well.

On Figure 3.8 we see the successive comparisons every second week from the 21st of January to the 4th of March 2003 of the zonal SSH averages (seven assimilation cycles have occured in that period). The red line represents the observed SSH (a mean sea surface plus the anomalies observed from the satellite) and the purple line represents the corresponding forecast SSH. Cycle after cycle, the forecast SSH progressively converges towards the observed values, which means that the data assimilation system can bring the model nearer to the data in spite of the initial bias.

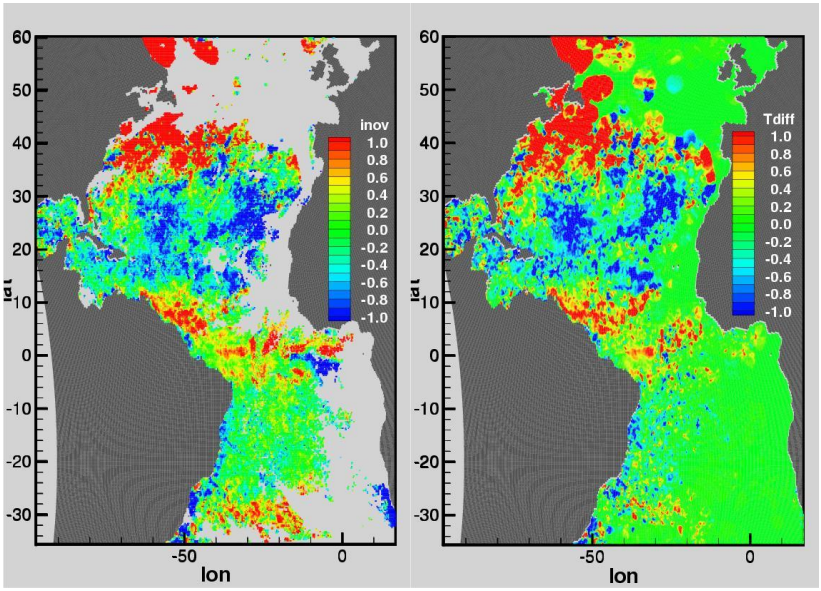


Figure 3.7: Left: July 29th 2003 innovations (Observed minus forecast SST before assimilation); Right: the subsequent surface temperature increments (analyzed minus forecast model SST).

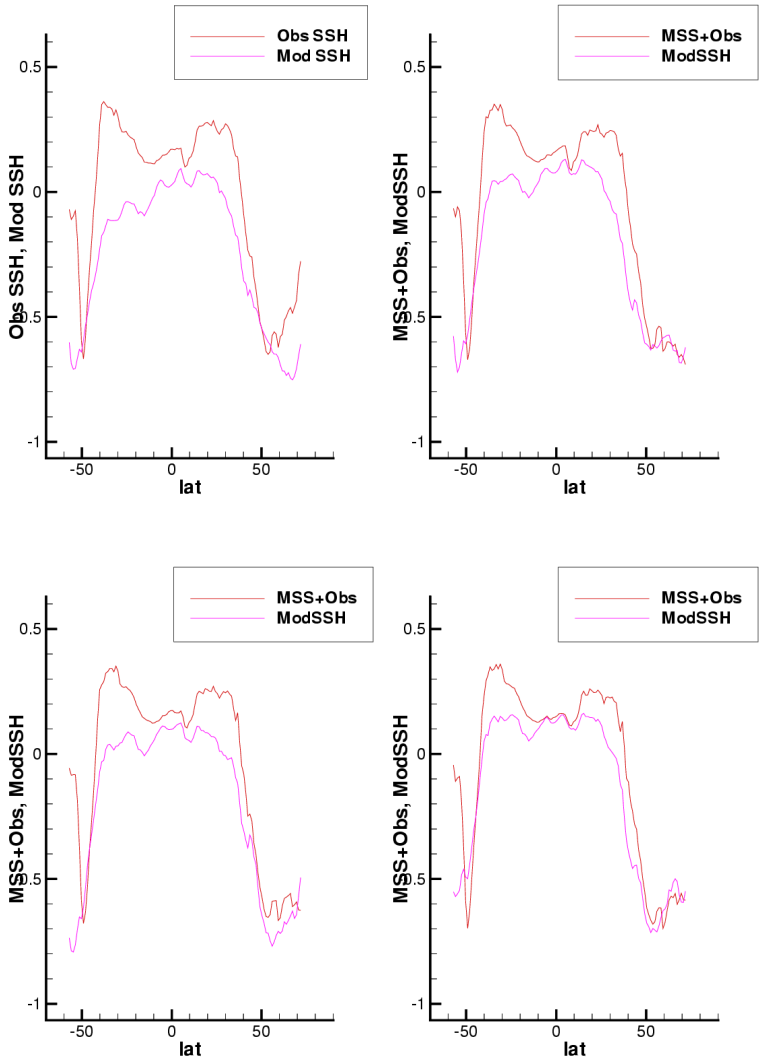


Figure 3.8: Successive zonal averages of sea surface heights from the 21st of January (top left panel) to the 4th of March 2003 (bottom panel), every second week. Red line: observed SSH (observed SLA plus mean sea surface height); Purple line: modeled SSH.

3.5 WP 3100: Assimilation of brightness temperature and salinity

The main objective of the ESA: Soil Moisture and Ocean Salinity (SMOS) mission is to globally observe soil moisture and ocean salinity, two crucial variables for modelling our weather and climate *Berger et al.* (2002). Salinity is fundamental in determining ocean density and hence thermohaline circulation. Furthermore, ocean salinity plays a part in establishing the chemical equilibrium, which in turn regulates the CO_2 uptake and release. Unlike sea surface temperature (SST) and sea level anomalies (SLA), it has not yet been possible to measure salinity from space.

The SMOS instrument will be launched early 2007, and is designed to provide brightness temperature (T_B) data for 3–5 years. The instrument is microwave radiometer using the frequencies in the L-band, corresponding to 1.4 GHz. The satelite orbit, instrument design and dataprocessing prosedures is designed to provide data every third day with a 35–50 km resolution at all the various data-processing levels defined by the Committee on Earth Observing Satellites (CEOS). The accuracy requirment of the ocean salinity observations has been set to 0.1 practical salinity units (1 psu = 1g salt in 1kg of seawater), every 10 days at 200 km spatial resolution, based on results from GODAE (Global Ocean Data Assimilation Experiment).

As the SMOS instrumentation will not provide SSS data directly we need the inversion from brightness temperature (T_B) data to ocean salinity. However, T_B is a non-trivial function of sea surface temperature (SST) and salinity (SSS) and parameters determining the interaction between the atmosphere and the ocean. This makes the inversion to salinity complex and difficult. Another potential use of the T_B data is the assimilation directly into dynamical ocean circulation models used for ocean monitoring and prediction. In this way the inversion to salinity is avoided but the dependency on the emissitivity model, and the corresponding model error, is still present through the establishment of the modelled T_B observations (forward modelling). The accuracy of the observations will now be the noise in the T_B measurements. The NERSC data assimilation system is based on an implementation of the Ensemble Kalman Filter (EnKF) with HYCOM, which is currently used for assimilation of SST, SLA and ocean color data. In this work package we have refined this system to adapt the future T_B measurments.

All the ocean-atmosphere fluxes are determined in the ocean model using the atmospheric forecasts from numerical weather prediction centers together with the current model state. Thus, we can compute T_B from the model. The integrated use of all these parameters allows us to optimally introduce the information from T_B into the assimilation system and to properly update the model state in a way which is consistent with all prior information and error statistics. A nonlinear measurement functional for T_B is used by augmenting the model variables with the T_B, and then computing covariances between the T_B and the rest of the model state. It is then possible to examine and estimate how model variables are influenced by the measured T_B as well as the strength of this influence. The objective is to demonstrate assimilation of T_B data using simulated observations, and the work consists of the following sub-task:

1. Establish simulated T_B data fields.
2. Update the assimilation scheme to adapt T_B data.
3. Run assimilation experiments to study the impact on temperature and salinity when assimilating T_B data. This task will provide an assessment of the potential of a methodology for assimilating T_B into a preoperational ocean monitoring and prediction system.
4. Run numerical experiments with direct assimilation of Sea Surface Salinity data.

In previous ESA projects *Barnier et al.* (2002) it is found that an accuracy of 0.1 psu over a grid cell of 100 km times 200 km for a period of 7 days is sufficient for description and quantification of many central ocean processes both on monthly and seasonal scales.

In this work package we will test what accuracy in T_B observations are needed for assimilation into the ocean modelling and data assimilation system used at NERSC.

3.5.1 Establish simulated T_B data.

The T_B data function

Passive microwave radiometers measure the emission of electromagnetic waves from the earth in the wavelength domain from 1.5 cm to 30 cm. The measured wave field has contributions from emissions from the sea surface, the atmosphere, and from reflection of incoming wavefield at the sea surface from the atmospheric and solar emission and deep-space background radiation. The emission can be modelled by more or less complex radiation models. The accuracy of these models depends on how accurate the sea surface roughness is descibed and modelled and on how good the electromagnetic scattering is modelled. T_B can be retrieved by forward modelling using the descibed emission model in *Barnier et al.* (2002).

In the work performed in SIREOC we have used the 6-dimensional T_B data grid computed from an emissivity model from IFREMER using a small slope approximation model combined with a spectral model for the ocean surface. Details of these computations and description of the model can be found in *Barnier et al.* (2002). The T_B values were computed for the following conditions:

1. 10 m wind speed from 0 to 20 m/s with a step of 1 m/s.
2. SST from 0 to $30°$C with a step of $1°$C.
3. SSS from 32 to 38 psu with a step of 1 psu.
4. incidence angle θ_{inc} from 0 to $60°$ with a step of $5°$.
5. azimuth angle θ_{azi} from $0°$ to $360°$ with a step of $20°$.

The tabulated T_B values were organized in a large number of ascii files, and arrived at NERSC along with a multi-linear interpolation code used to calculate the brightness temperature. Thus, providing six parameters as input (see below), the program would open the correct files and extract T_B values from the table. These T_B values are then interpolated to provide an estimate of the T_B value corresponding to the exact values of the input parameters. The T_B data were merged into one large six-dimensional matrix which is stored in a single binary file. This file is read once and kept in memory every time we initiated an assimilation step. The required T_B values could then be quickly extracted from memory and passed as input to the multi-linear interpolation function.

The brightness temperature function

The T_B as function of SST and SSS is plotted in Figure 3.9 for different angles of incidence (5, 20, 40 and 50 degrees) under no wind conditions. By intercomparing the plots it is seen that the dynamic range of the TB values is around 5 K, but that the range is swithced towards higher values with increasing angle of incidence. The gradients with respect to SSS and SST is more or less conserved at temperatures above $20°$C for different angles of incidence. However, at temperatures less than $20°$C the gradients are changing significantly when considering the different incidence angles. In all the plots it is seen that the T_B value decreases with increasing SSS for all temperatures. Furthermore, it is seen that the dependency on SSS is smallest for temperatures less than $10°$C. The dependency on SST is slightly more difficult as the T_B increases with temperatures less than $20°$C and decreases at temperatures greater than $20°$C. The brightness temperature changes more rapidly with SST at relatively low salinities.

The impact of variable wind speed and azimuth angle is presented for two selected incidence angles , $5°$, in Figure 3.10 and, $40°$, in Figure 3.11. At the low wind speed (10 m/s) the change from wind along track(left) to cross wind(right) do not change the T_B values dramatically for the incidence angles considered here. At wind speed of 20 m/s the effect of changing the azimuth is more pronounced at the smalest angle of incidence than at the largest angle. For the two

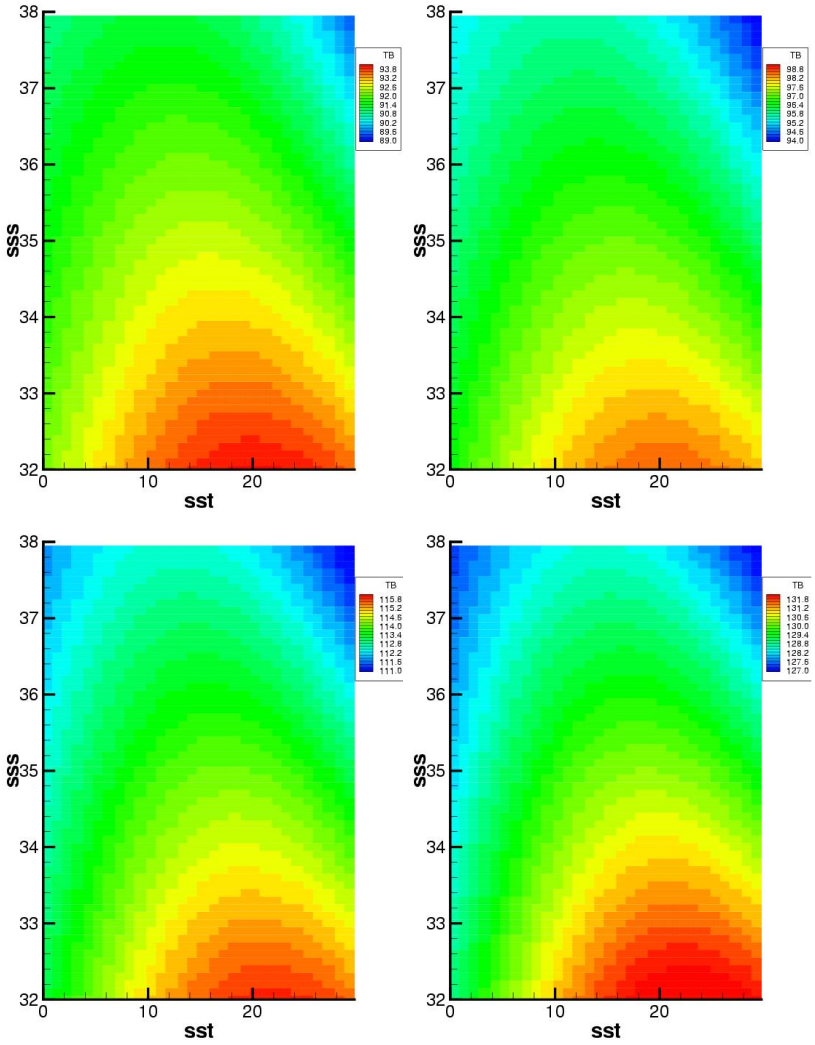


Figure 3.9: Brightness temperature as function of SST and SSS in vertical polarisation. All the plots include zero wind conditions (wind speed is 0 m/s and azimuth is 0 degrees). In the upper row the incidence angle is 5 degrees in the left plot and 20 degrees in the right. In the lower row the incidence angle is 40 degrees in the left plot and 50 degrees in the right.

incidence angles and the two different azimuth angles considered in Figure 3.10 and Figure 3.11 a significant increase of T_B values is observed for all SST and SSS when the wind speed is increased from 10 to 20 m/s.

In the emission model the sea surface roughness is connected directly to the wind speed and wind direction. This is a very simple way of modelling the sea surface roughness which is also influence by foam, rain, natural surface films, water salinity and more distant waves. Furthermore, calculation of scattering from rough surfaces introduces relatively large model errors, which are unknown when we model the T_B. Therefore in this study we will not study the ”exact” impact of the wind in the assimilation process. However, from the plots above we can conclude that if we underestimate the wind when performing the assimilation, the analysed field will have too high positive update in salinity and temperature. This has also been verified through assimilation experiments.

Some initial conclusions can be made:

- The dynamic range for T_B is around 5° K for each set of incidence angle, polarisation and wind condition and the definition area of the T_B function with respect to SSS and SST.

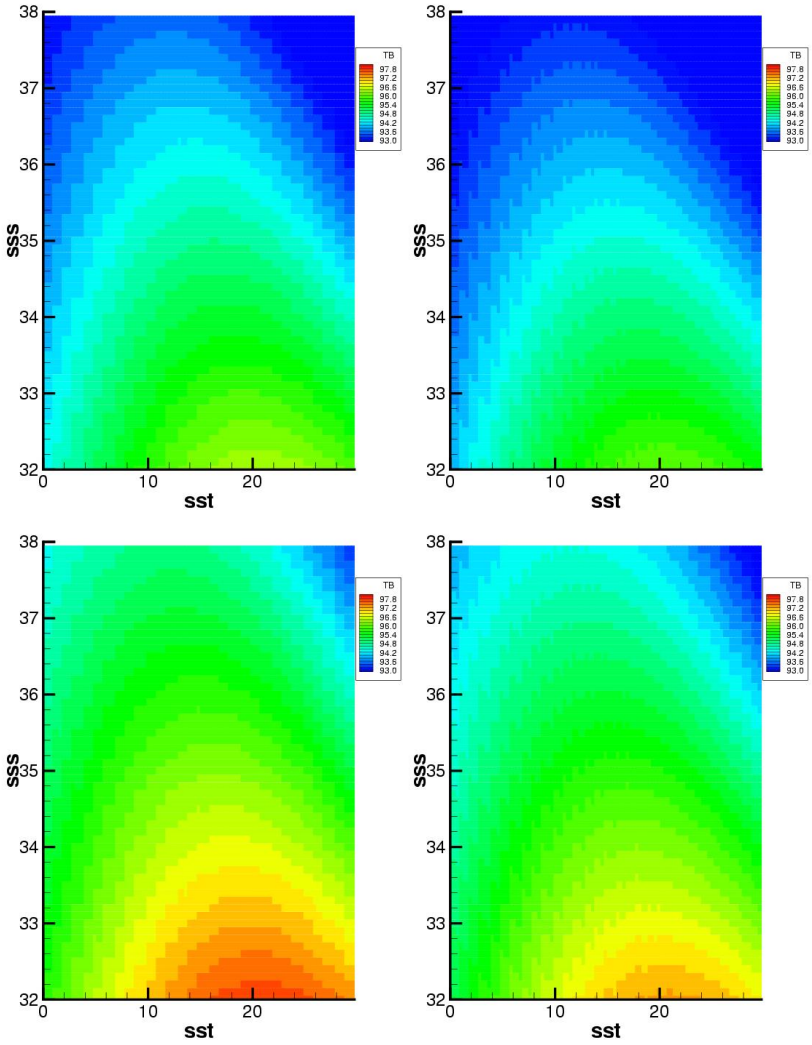


Figure 3.10: Brightness temperature as function of SST and SSS in vertical polarisation. All plots in this figure have an incidence angle of 5 degrees. In the upper row of plots the wind is 10 m/s while it is 20 m/s in the lower row. The left plots has azimuth equal to 0 degrees while the right plots have a 90 degree azimuth (cross-wind)

- The T_B value decreases monotonically with salinity for all temperatures, incidence angles and winds considered here.
- In the cold water (less than $10°$C) the T_B function is more sensitive to salinity at small incidence angels than at larger angles.
- The monotonic T_B dependency on salinity is stronger at temperatures greater than $10°$C.
- The T_B dependency on temperature is non-monotonic such that a specific salinity and T_B value give two "suggestions" for the temperature.
- The effect of changing azimuth increases with wind speed.
- The effect of wind speed is stronger than the effect of changing azimuth angle.

Keep in mind that only vertical polarisation is considered in this report.

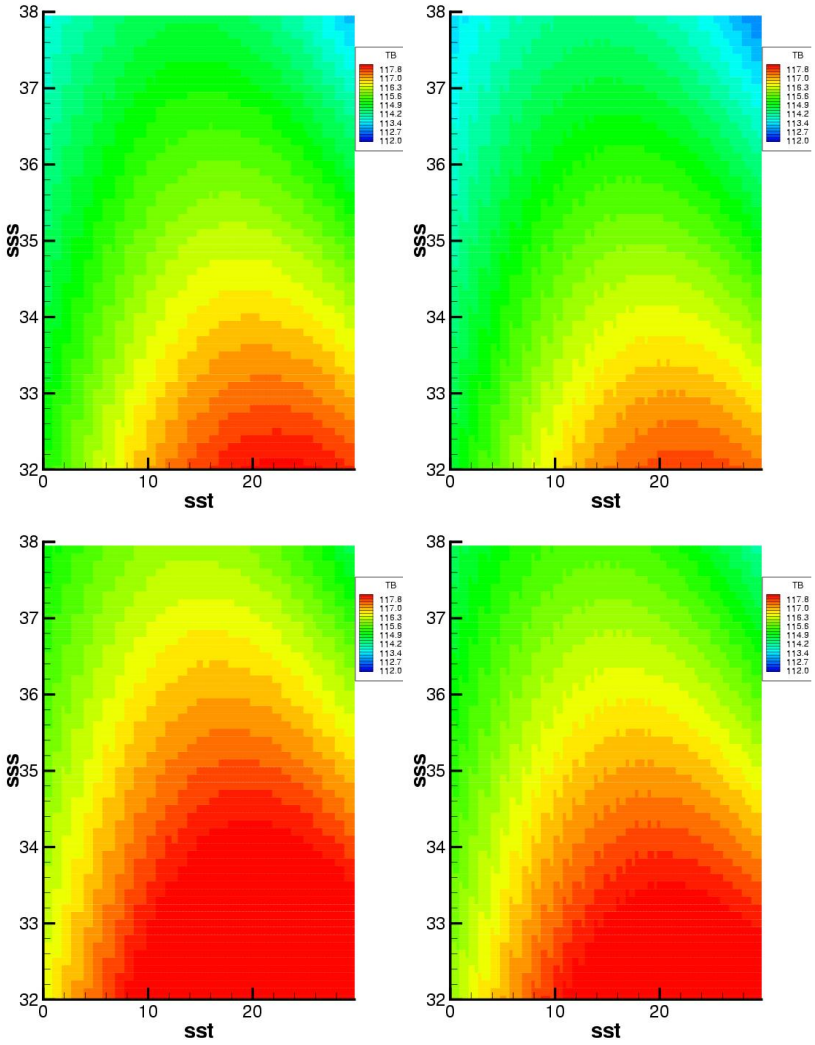


Figure 3.11: Brightness temperature as function of SST and SSS in vertical polarisation. All plots in this figure have an incidence angle of 40 degrees. In the upper row of plots the wind is 10 m/s while it is 20 m/s in the lower row. The left plots has azimuth equal to 0 degrees (along track wind) while the right plots have a 90 dergree azimuth (cross-wind).

Simulated brightness temperature data

For the assimilation experiments we needed to establish procedures to create synthetic observations with a known percentage of noise, and the following was used:

1. First the T_B values were computed on an observational grid defined with 1 by 1 degree resolution using the T_B table as discussed above. In the T_B function the SST and SSS fields were obtained from a selected forcast model restart file. The angle of incidence was set to $40°$, the wind speed was set to 0 m/s, the azimuth angle was set to $0°$ and the polarisation was set to vertical, and these parameters were kept constant throughout this study.

2. A two dimensional smooth random noise field, $N(i, j)$, with mean zero and variance one, was generated using standard routines. The noise field was then multiplied by the measurement error standard deviation σ which was given as a percentage of the dynamic range in T_B values. In our cases the range were found to be $5°$ K.

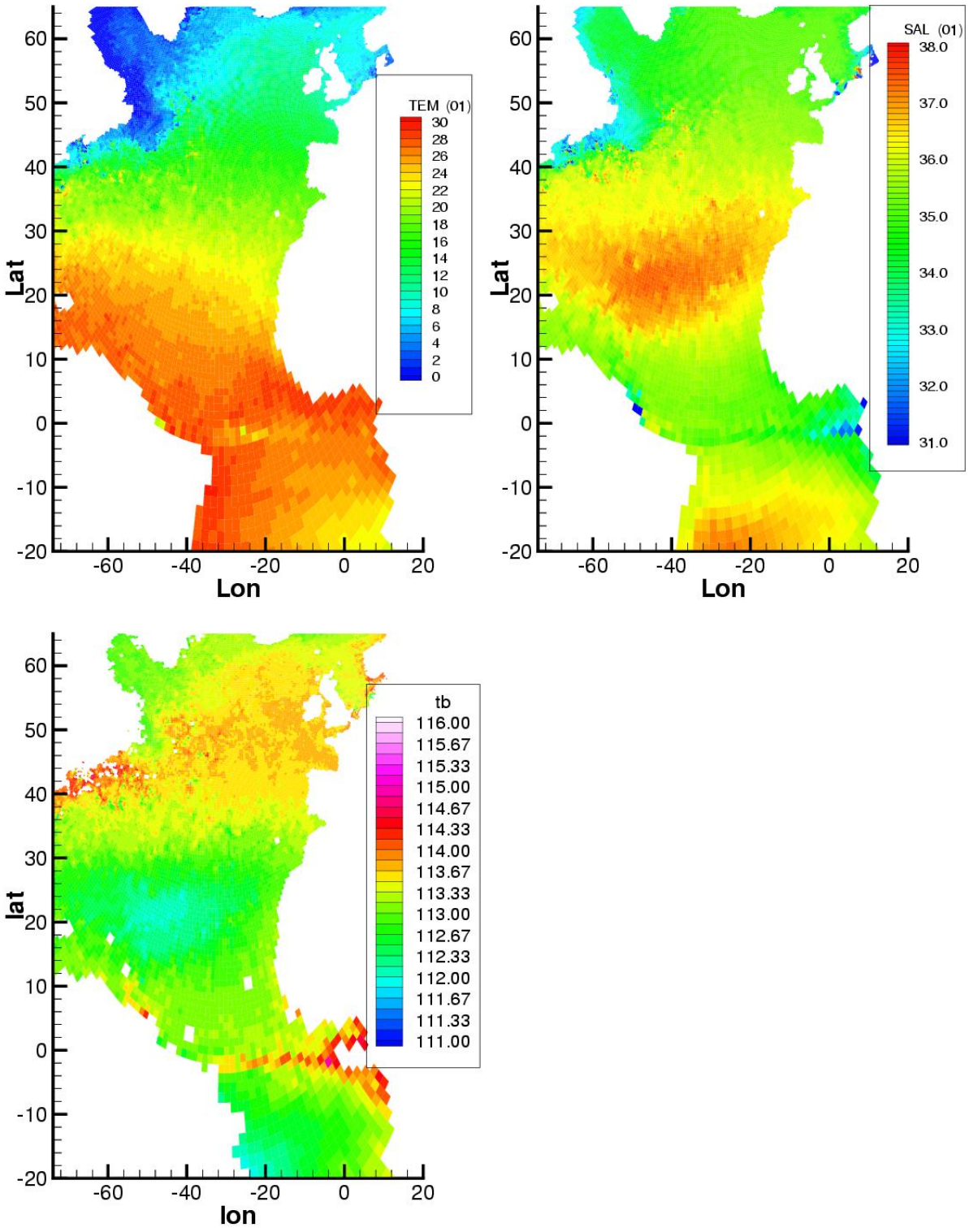


Figure 3.12: The ensemble mean temperature and salinity for January 1 2002 in the DIADEM/TOPAZ experiment plotted on the model grid. The mean Sea Surface Temperature (SST) is shown in the left column and mean Sea Surface Salinity (SSS) to the right. The corresponding T_B values using incidence angle 40 degrees are shown in the lower plot.

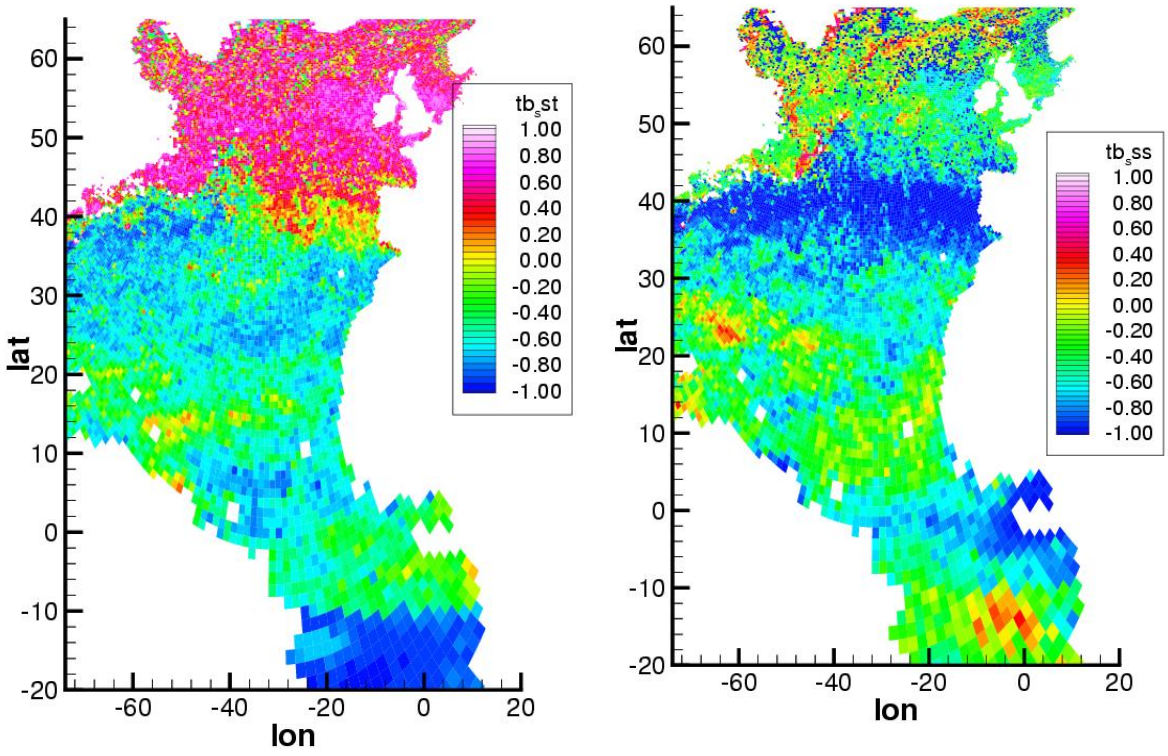


Figure 3.13: The ensemble correlation between temperature and the corresponding T_B at each model grid point is plotted to the left. To the right the covariance between salinity and T_B is plotted.

Finally, the grid of perturbed T_B observations to be used in the EnKF analysis was calculated as

$$d(i,j) = T_B(i,j) + \sigma * N(i,j). \tag{3.6}$$

The data, geographical location, a given error variance, and location are then stored in a binary file (”observation.uf”) which is later read by the EnKF analysis program.

Accessing the brightness temperature data within the EnKF scheme

Once the new T_B function was established it was relativly easy to update the EnKF scheme to adopt the new type of data. The synthetic T_B data is read from the ”observation.uf” file and stored in a vector. The ensemble of model states is stored in a matrix A. From the model states we are then computing the corresponding modelled T_B values using the same T_B function as when establishing the synthetic observations, these are stored in the matrix HA. This allows for the computation of the innovation (or difference) between the T_B observations and the model prediction.

3.5.2 The modelled SST/SSS data and synthetic brightness temperature data.

Figure 3.12 shows the ensemble mean SST and SSS model data for the 1st of January 2002 as taken from the DIADEM/TOPAZ real time system. The temperature at high latitudes north of 30°N is typically less than 20°C while it is between $20-30$°C at low latitudes. The salinity has two strong anomalies at mid-latitudes with values around 37

psu in the centre of the anomalies. These plots will be interpreted as the model forecast, which will later be updated in the analysis scheme.

When real data from the SMOS satellite are available, several incidence angles will be included in the processing and this will improve the accuracy of the T_B data. Another complicating factor is the wind condition which will change along the track and with time. This will require advanced processing algorithms for the T_B data. As the major focus in this project is not the data processing algorithms we have selected only one particular incidence angle of 40 degrees, vertical polarisation and zero wind conditions when establishing the synthtic T_B data to be used in assimilation experiments.

In this case, the T_B data takes values between 111 and 116°K. The functional form of T_B with respect to SSS iand SST is displayed in the lower left plot in Figure 3.9. The plot shows a monotonically decreasing T_B with increasing SSS. Whereas the dependence on SST is more complex since T_B is increasing as a function of SST for low temperatures up to $10 - 20$°C (dependent on the SSS), where it reaches a maximum and then decreases again. Thus, there may be two different water masses corresponding to the same value of T_B, and this could potentially be a problem in the assimilation scheme.

The brightness temperature data corresponding to the mean temperature and mean salinity of the predicted model state is shown in the lower left plot in Figure 3.12. Notice first that T_B is not defined in many points in the Gulf of Guinea due to very fresh water, and it is also undefined in areas along the coast of South America due to high temperatures. In grid points with undefined T_B values there will be no assimilation of data. The particular features to notice in the T_B data are:

- low T_B values are found at mid-latitudes due to high SSS values,
- low T_B values are also found at very high latitudes caused by very low SST values,
- high T_B values are found in regions with temperatures between $10 - 20$°C and fairly normal salinities, s.a. in the Gulf Stream extension and the North Atlantic Current,
- The maximum T_B values are located in the eastern Equatorial regions which are dominated by very warm and fresh water,

Furthermore, it is seen that the T_B features in the north look like the features in the SST plot, while the salinity features are “preserved”, south of 40 degrees in the plot of SSS. This “geographical” change in T_B sensitivity is also reflected in the correlation plots in Figure 3.13, where it is observed that the correlation to temperature is relatively high in the north while it is lower when correlating it with the salinity. Between 30-40° N there is a strong negative correlation between salinity and T_B, while the correlation with SST is less but still significant. The picture is more complicated in the south because the T_B can either increase with temperature (0–20°C) or it may decrease with temperature (20–30°C), in addition to a stronger sensitivity to SSS. On the other hand, the T_B decreases with salinity for all temperatures, and therefore the salinity features are preserved in the south.

In addition to the mean over the 100 member we have also calculated the variance over member in each model grid point (Figure 3.14) for SST, SSS and the corresponding T_B. The model variance is essential for the impact of the model result in the assimilation process.

In the upper left plot we see that north of 40° N the variance in SST is gradually reduced from 0.08°C^2 to 0.004°C^2, moving northward. Considering the variance in SSS in the righthand plot we observe that it is very low between 0.001–0.002 psu^2 north of 30° N, except in the Gulf Stream outside the coast of North America and in the Skagerak where the variability is between 0.04–0.2 psu^2. The calculated variance in T_B is very low between 30–40° N because of the strong correlation to SSS at these latitudes (Figure 3.13) and the low variance in SSS. The variance in TB increases slighly towards the north because of reduced dependence on SSS and increased dependence on SST which has a larger variance than the salinity in this region.

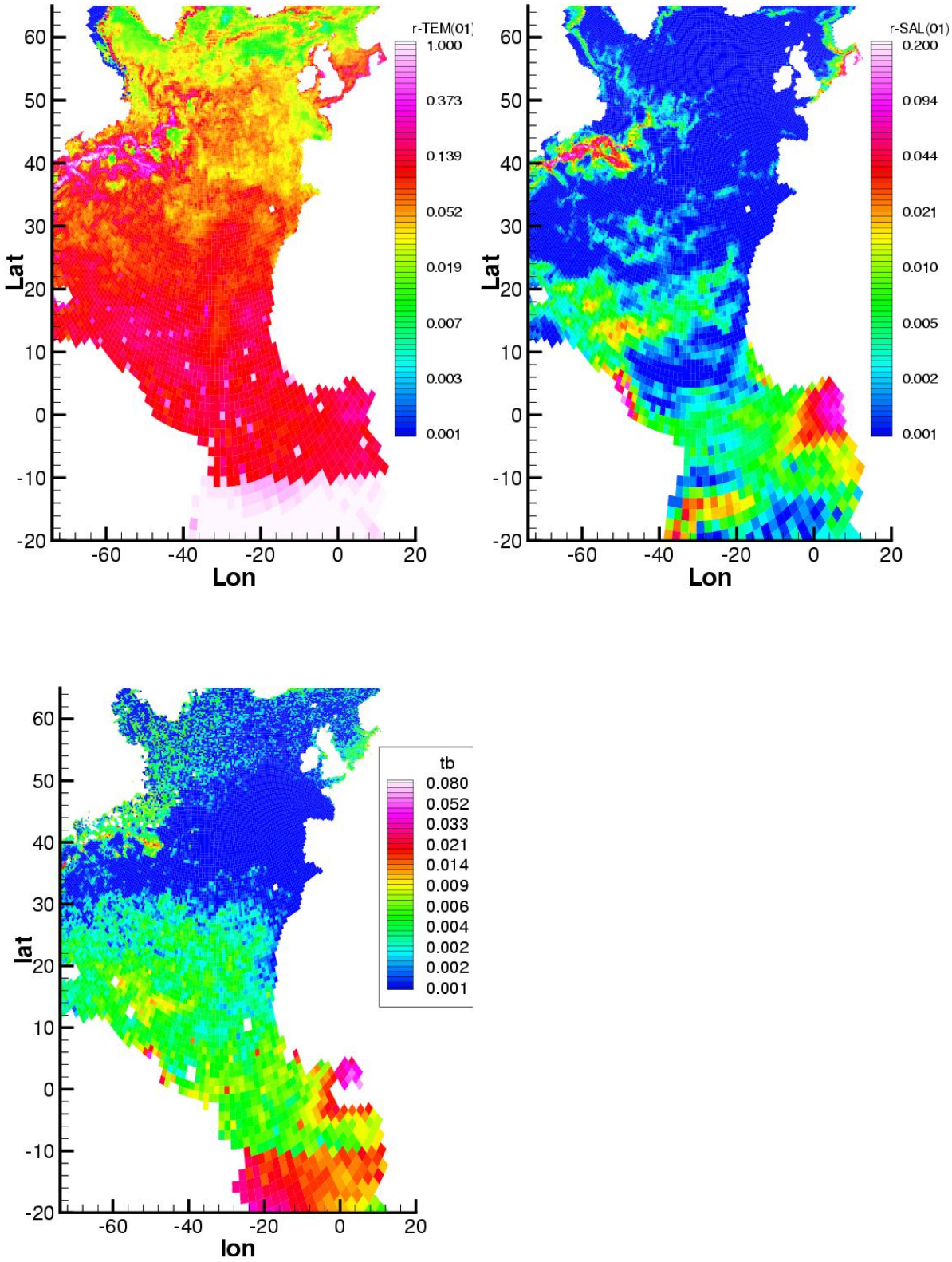


Figure 3.14: The variance in the model data is plotted for temperature (left) and salinity (right) . The lower plot is the variance in T_B data, it contains masking of points where the TB value is undefined for one or more of the 100 members.

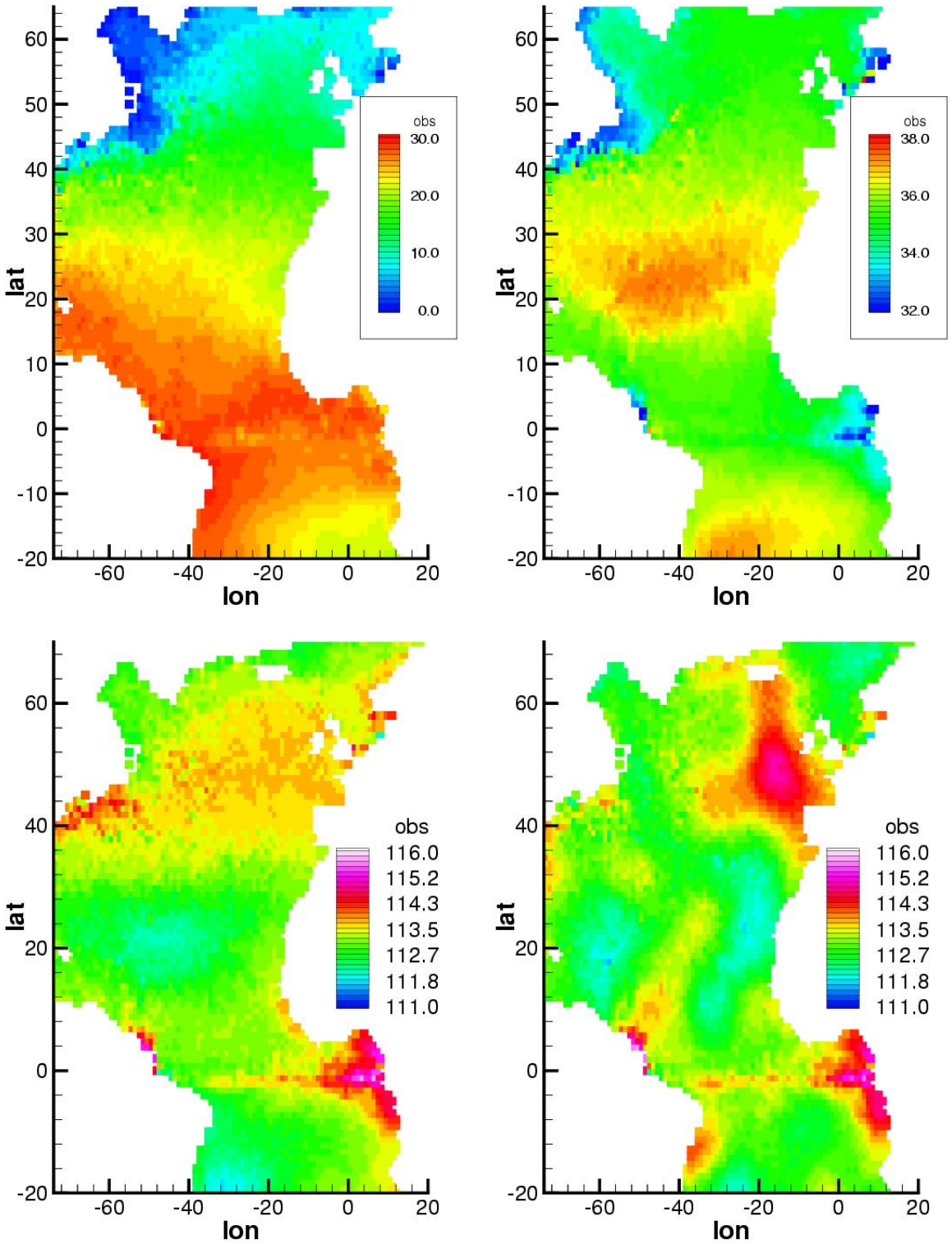


Figure 3.15: The upper left plot shows original SST (left) and SSS (right) data from member 40 in the ensemble. The lower left plot shows the corresponding T_B values computed from the T_B function. T_B The lower right plot is generated by adding 10 % noise to the T_B plot to the left.

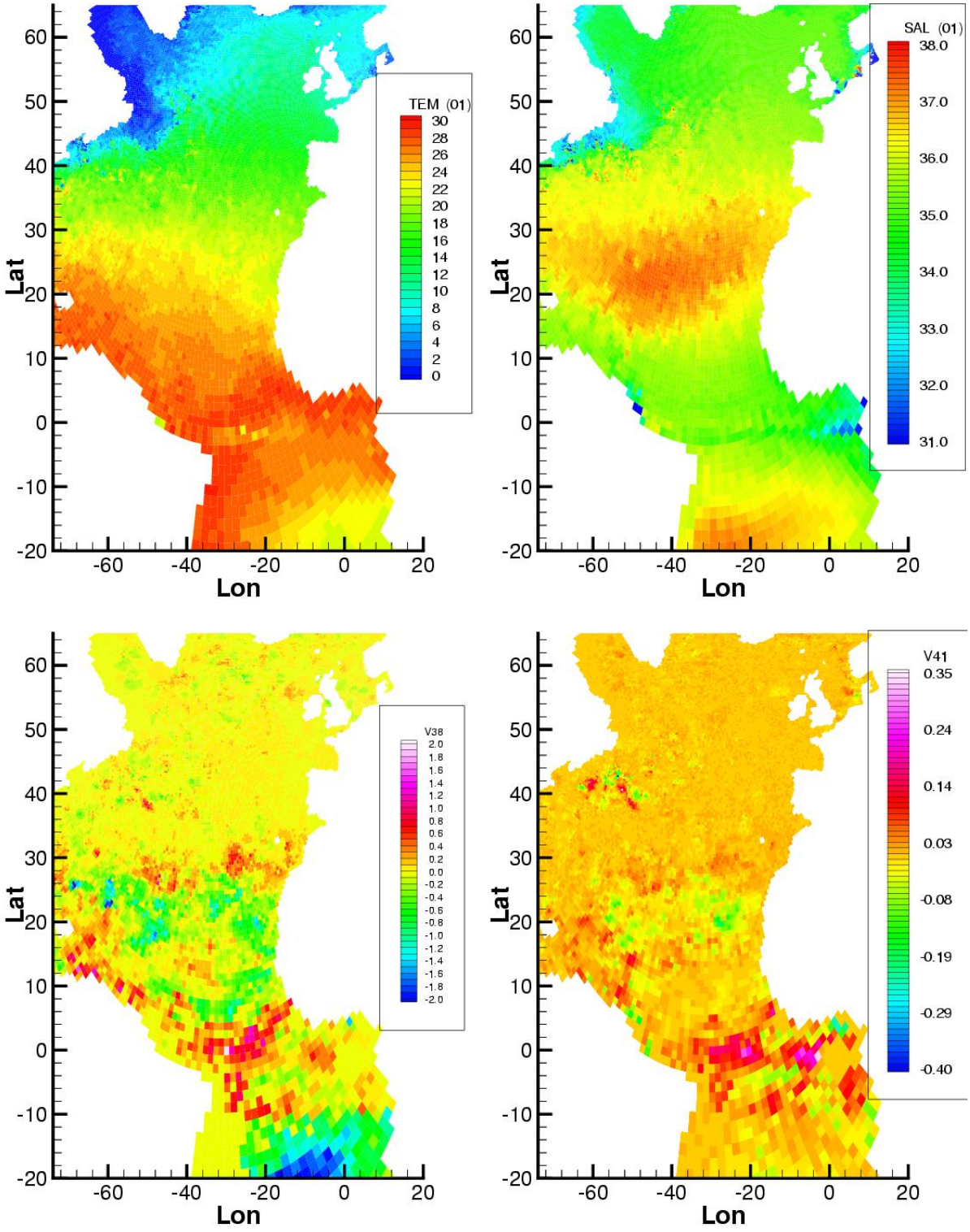


Figure 3.16: (Case 1) The analysed SST (left) and SSS (right) are shown in the upper plots. The lower plots show the difference (analysis minus forecast) for SST (left) and SSS (right).

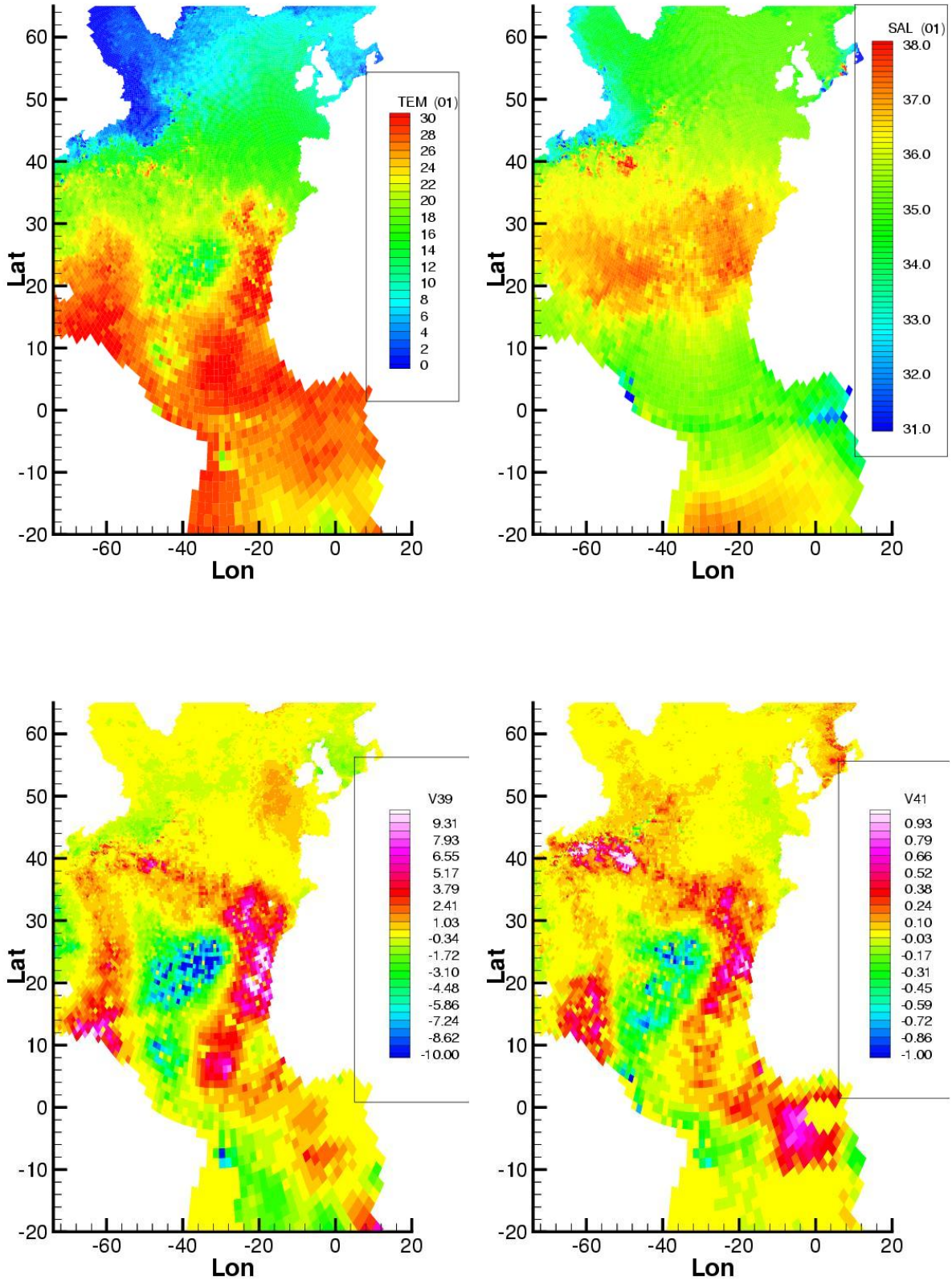


Figure 3.17: (Case 4) The analysed SST (left) and SSS (right) are shown in the upper plots. The lower plots show the difference (analysis minus forecast) for SST (left) and SSS (right).

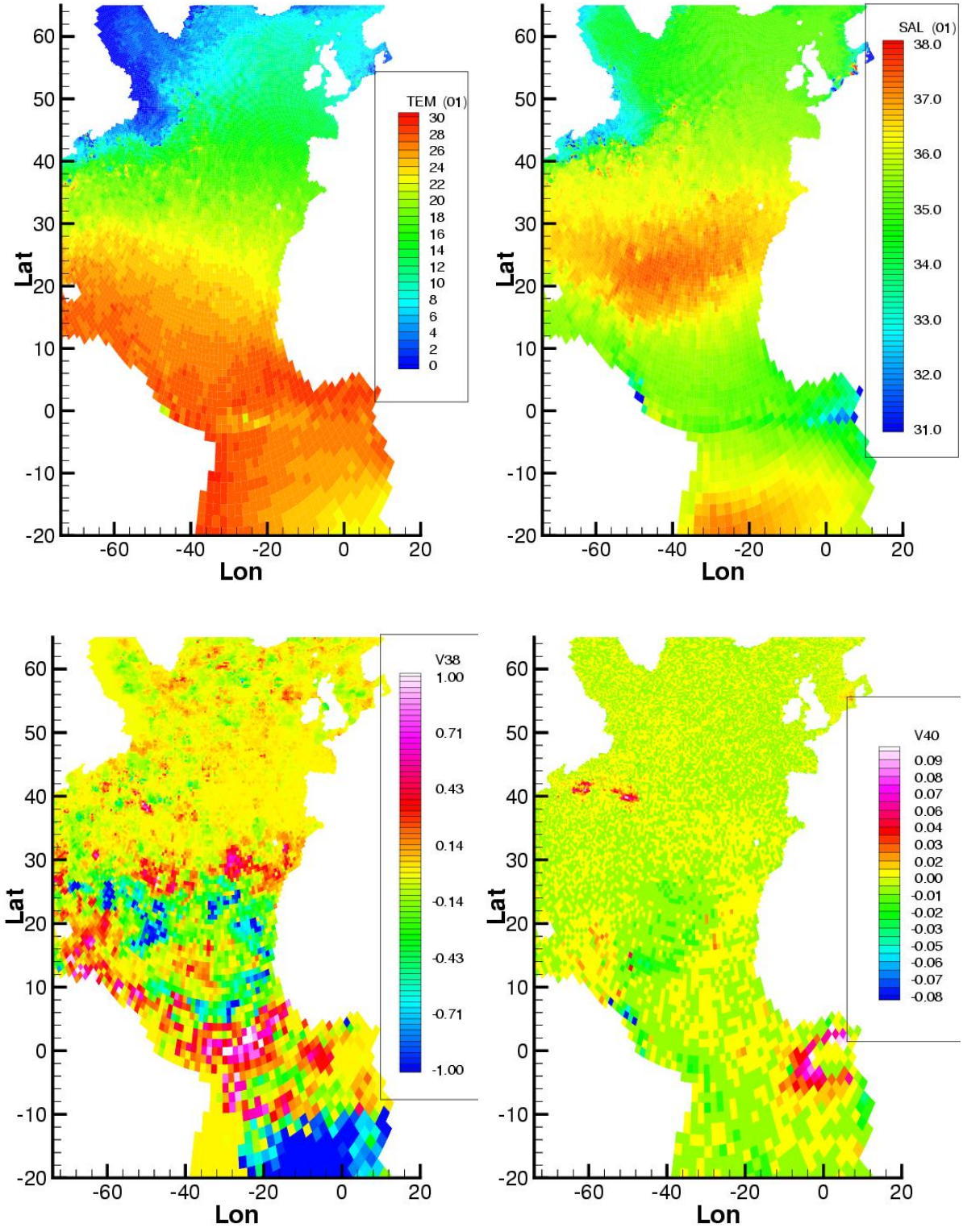


Figure 3.18: (Case 5) The analysed SST (left) and SSS (right) are shown in the upper plots. The lower plots show the difference (analysis minus forecast) for SST (left) and SSS (right).

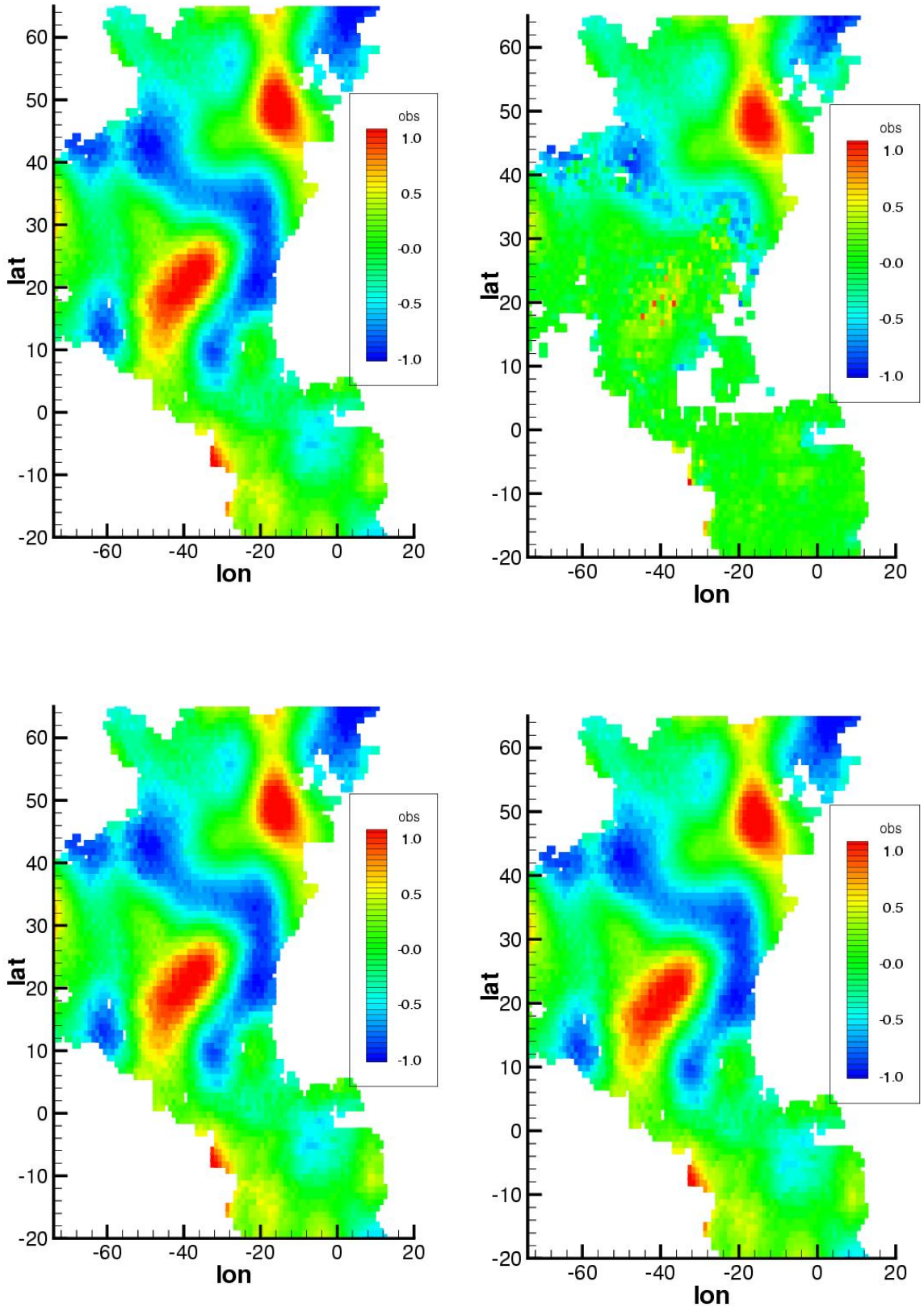


Figure 3.19: Forecasted innovation(left) and analysed innovation (right) for Case 4 (upper) and Case 5(lower). The lower plots show the difference (analysis minus forecast) for SST (left) and SSS (right).

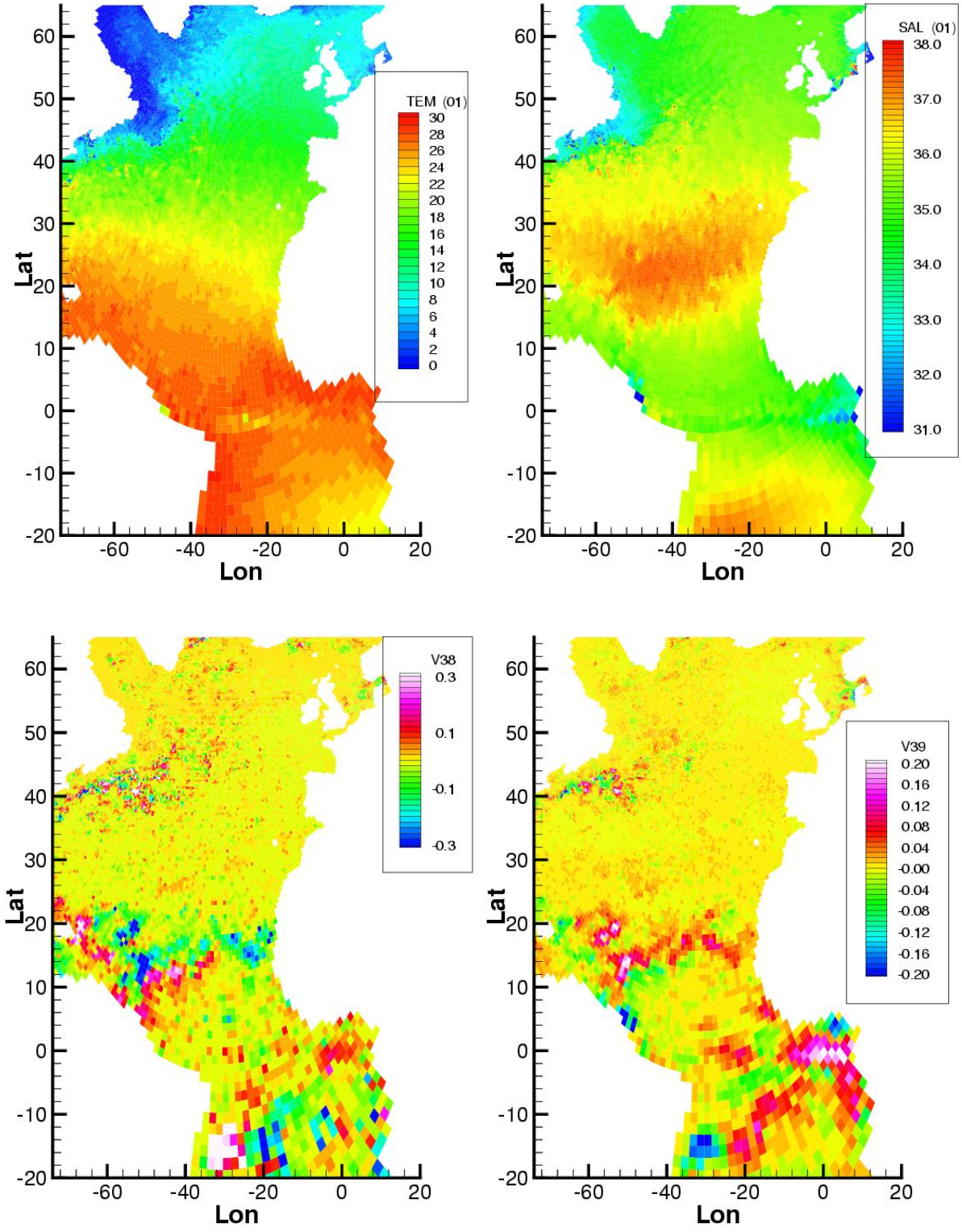


Figure 3.20: (Case 6) The analysed SST (left) and SSS (right) are shown in the upper plots. The lower plots show the difference (analysis minus forecast) for SST (left) and SSS (right).

In the south the variance in SST is generally higher than in the north. The variance is between $0.052 - -1.1°C^2$ (0.2–1.0°C in standard deviation). The strongest variance in SST is found south of 10° S, this variance has been set high because no assimilation of SST data is done south of this latiude in the DIADEM/TOPAZ experiment.

In the south the variance in SSS vary much more than in the north, the variance takes values between 0.002 psu^2 and 0.2 psu^2. The largest variance is found in and ouside the Gulf of Guinea.

When considering the variance in T_B we see by comparing the other plots in Figure 3.14 that high variances are generally found where the variance in SSS is relatively high at the same time as the variance in SST is stable and high. One clear exception from this is south of the 10 S latitude where variance in SST is much higher than elsewhere and the correlation is close to -1.

From the above discussion we conclude that the variance in T_B is controlled by

- the correlation to SSS/SST,
- the magnitude of the variances in SSS and SST

3.5.3 Assimilation experiments

The most important part of the study is to investigate the impact of assimilating T_B and SSS data on the model SST and SSS fields. As seen from the above, the T_B function of SST and SSS is complex and the mapping from SST and SSS in the model grid requires careful interpretation.

To generate the synthetic T_B data to be assimilated, we have used the SST and SSS from the 40$^{\text{th}}$ ensemble member, as shown in Figure 3.15. The data were interpolated to the observation grid, and then used as input to the T_B function. The lower left plot shows the corresponding T_B data. The model fields are similar to those given in Figure 3.12, but differs within the one or two standard deviations in the ensemble which define the accuracy of the model. When intercomparing the SST and SSS in Figures 3.12 and 3.15 only small differences are visual. The SST in Figure 3.15 is slightly warmer in the south, and the SSS anomalies in Figure 3.15 are wider than in the forecast. The lower plot in Figure 3.15 shows the corresponding T_B data.

The brightness temperature T_B can be written:

$$T_B = T_B\left(\theta_{inc}, \theta_{azi}, SST, SSS, u, p\right), \tag{3.7}$$

where θ_{inc} is the incidence angle, θ_{azi} the azimut angle, u the wind speed and p the polarization, respectively.

Nine assimilation experiments have been carried out where the following T_B data and parameters have been used:

Case 1 Assimilates the original T_B data as given in Figure 3.15, the data variance is set to $0.0025°K^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 0$m/s, p is vertical and SST and SSS are chosen for ensemble member 40, respectively.

Case 2 Assimilates the T_B data based on constant SST=16°C, the data variance is set to $0.0025°K^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 0$m/s, p is vertical and SSS is chosen for ensemble member 40, respectively.

Case 3 Assimilates the T_B data based on constant SSS=34°C, the data variance is set to $0.0025°K^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 0$m/s, p is vertical and SST is chosen for ensemble member 40, respectively.

Case 4 Assimilates the T_B data which were corrupted by noise, the data variance is set to $0.0025°K^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 0$m/s, p is vertical and SST and SSS are chosen for ensemble member 40, respectively.

Case 5 Assimilates the T_B data which were corrupted by noise, the data variance is set to $0.25°K^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 0$m/s, p is vertical and SST and SSS are chosen for ensemble member 40, respectively.

Case 6 Assimilates the SSS data as given in Figure 3.15, the data variance is set to 0.01psu^2, SSS are chosen for ensemble member 40.

Case 7 Assimilates the SSS data as given in Figure 3.15, the data variance is set to 0.32psu^2, SSS are chosen for ensemble member 40.

Case 8 Assimilates the original T_B data as given in Figure 3.15, the data variance is set to $0.0025°\text{K}^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 20$m/s, p is vertical and SST and SSS are chosen for ensemble member 40, respectively. In the assimilation scheme the wind information is set to be $u = 20$m/s $\theta_{azi} = 0°$ in all model gridd points.

Case 9 Assimilates the original T_B data as given in Figure 3.15, the data variance is set to $0.0025°\text{K}^2$, $\theta_{inc} = 40°$, $\theta_{azi} = 0°$, $u = 20$m/s, p is vertical and SST and SSS are chosen for ensemble member 40, respectively. In the assimilation scheme the wind information is set to be $u = 10$m/s and $\theta_{azi} = 0°$ in all model gridd points.

Four of the assimilation experiments are examined in detail in this report: Case 1, Case 4, Case 5 and Case 6. The others are additional experiments used to support the interpretation of the others.

We first describe the results from the Case 1, Case 4 and Case 5 experiments which all assimilate T_B data. The T_B observations are in all cases evaluated from the model SST and SSS (displayed in Figure 3.15) using the 40^{th} ensemble member. In the first experiment the T_B is used without any perturbation, while in the next two experiments we add 10 % smooth and random noise to get more significant changes in updates. Finally, in Case 6 we observe the effect of assimilating directly the SSS data.

The observation vector included 7344 data points in "wet" model grid points. Some points were excluded from the assimilation because the T_B function is only defined for temperatures between 0 and $30°$C. This reduces the number of observations included in the EnKF analysis to 6793.

Results: Assimilation of T_B.

The variance in the data is a measure of their accuracy, and it determines the influence of the data in the analysis. In Case 1–4 the variance is 0.0025 $°\text{K}^2$ for all data points. To test the effect of a poorer accuarcy in the T_B data we use a variance of 0.25 $°\text{K}^2$ in Case 5.

The analysed SST and SSS resulting from the Case 1 experiment are given in the upper plots in Figure 3.16. In the lower plots we give the differences between the analyses and forecasts for SST (left) and SSS (right). It is seen that the SST update (lower left) is generally less than $\pm 0.6°$C, except for a circular region in the south where the update is negative and close to $2°$C. In the north the update in temperature is less than $\pm 0.2°$C, while relatively large areas in the south of $40°$ N have updates between $\pm 1.0°$C. The difference between analysed and forecasted SSS is generally between -0.08 and 0.03 psu, except for an area along the Equator and in the Gulf Stream where the update is positive and close to 0.2 psu.

So the general observation in Figure 3.16 is that the SST and the SSS fields are more strongly updated south of around $40°$N than north of this latitude.

Because the difference between the mean of the ensemble members and the 40^{th} ensemble member is relatively small (see Figures 3.12 and 3.15), 10 % random noise was added to the data. The resulting temperature brightness data to be assimilated is shown in the lower right plot in Figure 3.15. The noise clearly dominates the T_B observations. A very pronounced T_B anomaly of ($+2$ °K) is seen at ($40°$ N, $20°$ E). In the south both negative anomalies (-2 °K) and positive anomalies ($+2$ °K) in the Gulf of Guinea are observed. These perturbed observations were assimilated with a variance of 0.0025 $° \ K^2$ for all observation points in Case-4.

The results after EnKF analysis are visualised by plotting the SST field to the left and the SSS field to the right in the upper part of Figure 3.17. The lower plots in Figure 3.17 show the residual between the analysed mean field and the

predicted mean field. Significant updates are observed both in SST and SSS south of 40° N. The SST is corrected with up to ± 10.0 °C, while the salinity is updated with up to ± 1.0 psu. Also in this case it is observed that the updates are more significant south of 40° N than at higher latitudes. In the north the SST is updated by upto 1.0°C in SST and the SSS by up to ± 0.10 psu. This corresponds to a 10 times larger update in the south than in the north. It is clearly observed that the the features seen in the perturbed T_B data (Figure 3.15) are reflected in the residuals between the analyzed model state and the forcasted model state (Figure 3.17).

In the Case 5 experiment we used the same T_B observations as in Case 4 (shown in the lower left plot in Figure 3.15). The variance is increased to $0.25^\circ \mathrm{K}^2$ for all data points, thus, the data will have less impact in this case. This corresponds to a standard deviation in T_B of 0.5° K^2 which corresponds to the previously recommended accuracy of the T_B for the final SMOS data. Results after EnKF analysis are shown in Figure 3.18. The upper plots show SST (left) and SSS (right) after analysis. The lower plots show the difference between the analysed and predicted SST (left) and SSS (right). The SST is updated with less than 0.3°C in absolute value. The strongest update in SST is found in the southern part of the plot. The difference between forecasted and analyzed SSS is less than 0.08 psu. This is a much smaller update than in the previous Case 4.

The T_B innovation plots before (left) and after (right) the analysis for Case 4 and Case 5 are plotted in Figure 3.19. The innovation are in all cases less than 1.0°K in asolute value. The two left plots shows the innovation before the analysis and they are identical for both cases. The two plots on the right hand side show the innovation after the analysis. In the upper right plot the innovation is reduced by around 50 % south of 40° N. The strong anomalies in the innovation (both negative and positive) is significantly weaker at low latitudes after the assimilation, while the anomalies in the innovation persists in the north. This again illustrates that the impact of the analysis is stronger at mid latitudes than at high latitudes. This is due to the regional differences in the ensemble predicted variance of the modelled SST and SSS. Since the EnKF predicts an accurate model forecast at high latitude this will ensure lower impact of the TB data, than in regions with large forecast variance. In the lower plots of Figure 3.19 we rarely observe any reduction of the innovation after the analysis in Case 5. Only very small reductions are observed in the Gulf of Guinea.

The experiments above demonstrate the importance of high accuracy in the T_B observations, and the very strong effect of the modelled variance in the model.

Direct assimilation of SSS

In the final experiment (Case 6) unperturbed SSS data from the 40^{th} emsemble member is assimilated to demonstrate the effect of direct assimilation of the SSS data. The result is shown in Figure 3.20. The variance of the SSS data was set to 0.01 psu^2. This accuarcy corresponds to the recommended accuracy and includes both the errors in the underlying measurements and the model error introduced in the inversion to SSS data. The corresponding update in salinty is between ± 0.2 psu in the south, while the update is close to 0 psu in the north. In temperature the update is close to 0.0°C in the north (except in the Gulf stream), and south of 20° N the update vary between ± 0.2 °C. The direct assimilation of salinity have the strongest impact on the SSS while it has a smaller impact on the SST.

Summary of the results from assimilation of T_B data

In this study we have presented results from 4 different experiments where simulated T_B and SSS data were assimilated into an OGCM using the EnKF. The experiments were designed assuming the following instrumental settings:

1. incidence angle $\theta_{inc} = 40.0^\circ$;
2. azimuth angle $\theta_{azi} = 0^\circ$;
3. wind speed $u = 0$ m/s;

List of results								
		Standard dev			SSS		SST	
Case		(K)	(psu)	(° C)	Update (psu)	Location	Update (C)	Location
Free-run	S	$0.004 - 0.2$	0.045	$0.2 - 1.0$				
Free-run	N	< 0.004	$0.03 - 0.04$	$0.2 - 0.3$				
Case 1	S	0.05			-0.8–0.2	Along equator	± 1.0	Along 30 N Near equator Below 10 S
Case 1	N				-0.08–0.03	Gulf stream	± 0.2	
Case 4	S	0.05			± 1.0	TB anomalies reflected in updates	± 10.0	TB anomalies reflected in updates
Case 4	N				± 0.1		± 1.0	
Case 5	S	0.5			± 0.01	Gulf of Guinea	-1.0–0.8	Along 30 N Near equator Below 10 S
Case 5	N				± 0.01		± 0.15	
Case 6	S		0.1		± 0.2	Gulf of Guinea Around 15 N	± 0.3	Along 15 N
Case 6	N				± 0.0	Gulf Stream	± 0.0	Gulf Stream

Table 3.1: List of results of assimilation of T_B and SSS data. Updates are given as overall range of updates within two areas defined as S: for areas south of 40° and N: for areas north of 40°. Anomalies are areas with large negative or positive updates.

4. polarisation p vertical;
5. variance in observations $\sigma^2 = 0.0025°\mathrm{K}^2$ (Case 1 and 4), $\sigma^2 = 0.25°\mathrm{K}^2$ (Case 5) and $\sigma^2 = 0.01\mathrm{psu}$(Case 6).

Two different variants of the T_B data sets were created using the modelled SST and SSS. The first data set was generated by using the 40^{th} ensemble member (Case 1) and the second by adding random perturbations to the first data set (used in Case 4 and Case 5). In Case 6 unperturbed SSS data from the 40^{th} ensemble member was assimilated directly. The sequence of experiments described above allowed us to give an assessment of the impact on the model state of assimilating T_B data sets representing different innovations (difference between observations and model data) and different accuracy of the data. We focussed the attention on the effect on the SST and SSS variables.

The results of the assimilation experiments are summarized in Table 3.1, where the accuracy of the model and observations are provided in standard deviation. Furthermore the table provides the general updates for north and south of 40° N latitude in each experiment and anomalies in the updates are listed.

In all cases considered in this study the strongest update in both temperature and salinity are found south of 40° N. In the cases with strong impact of the data the update in the south is 5–10 times larger than in the north. By increasing the difference between data and forecast by adding 10% (with respect to dynamic range) random noise to the T_B data the updates are generally increased by a factor 10. In the case of introducing the T_B data with a poor accuracy, 10%

of the dynamic range, in Case 5 the updates in SSS caused by the perturbed T_B is reduced by a factor of 0.01 and in SST by a factor 0.1 compared to the updates in Case 4 (with accuracy of 1% of the dynamic range).

In Case 6 we study the impact of direct assimilation of unperturbed SSS using the recommended data accuracy of 0.1 psu from the GODAE experiment. Almost no update is observed in the north, while clear updates are seen in the south. Compared to Case 5, which had a poor accuarcy in T_B but stronger innovations, the overall updates are significantly smaller in both SSS and SST. On the other hand the size of the updates in Case 6 are similar to Case 1 for SSS while the update in SST is 3 times smaller. So, given that the SSS data can be provided with an accuracy of 0.1 psu it will impact the SSS in the EnKF analysis in the same order as when assimilating T_B data with accuracy of $0.0025^{\circ}\mathrm{K}^2$. However, before taking any consequences of this result one should test the connection between accuracy in inverted SSS and the accuracy in the measured T_B.

There is no simple and overall connection between the accuracy in the SSS data and the accuracy in the T_B. In the case of using SSS data the accuracy of the observations is a combination of the measurement errors in the T_B data and the errors introduced when inverting T_B data, caused by limitations of the emission model and errors in the SST data and atmospheric data which is nessecary input to the inversion. When considering the T_B observations, the accuracy corresponds to measurement errors which can be improved by an averaging process. So from the observational point of view it would be better to use the direct observations of T_B.

In the assimilation process the variances/errors in the model data are also important. When performing the direct assimilation of SSS data the matrix, $\mathbf{HA}$, is established by interpolating the modelled SSS to the observation grid and the general error introduced is the modelled variance in the SSS data represented by the spread of the 100 ensemble members. In the case of representing the modelled T_B data the error budget becomes more complicated. The total error consits of

- the modelled T_B variance, which is established through the nonlinear T_B functional using the modelled variance in SSS and SST and errors in the atmospheric fields as input,
- the shortcomings in the underlying emission model used to establishing the T_B functional will introduce another error called representation error. This error is unknown, to us.

In our study the error estimate in the modelled T_B have only included the modelled variance and not the representation error. This means that we have overestimated the accuracy of the modelled T_B.

To proceed with the assimilation of the T_B or SSS data it is nessecary:

- to get improved emission models (including sea surface roughness)
- establish better estimates of the emmision model error statisitics, to get
- better estimates of the measurement errors.

To summarize our results:

- The assimilation of the T_B data proved efficient for controlling the model SSS and to some extent also the SST.
- The assimilation of T_B and SSS data had less impact at high latitudes than at mid latitudes. This is a combined effect of a typically lower variability in the model predicted SST and SSS at high latitudes compared with mid and low latitudes. In a real time monitoring system the predicted variance in the model will vary according to how long time it is since the last update. It is also resonable that the variance will vary according to the seasonal variations. In cases where the variance in the north increases the impact of assimilating T_B will increase in this region.
- The results of the assimilation depends strongly on the the error statistics in observations and in the model.

3.5.4 Perspectives and recommendations for future work

This work must be characterized as a first preliminary study where we demonstrate the capability of the system to assimilate T_B data into an ocean model. Based on this study we conclude that future satellite observed T_B data are likely to serve as an important source of information in operational ocean data assimilation systems, provided that the data can be delivered with sufficient accuracy and resolution and a proper emission model exists. However, our present study have several simplifications.

The synthetic T_B data have been generated by considering one angle of incidence, vertical polarisation and zero wind conditions. The real T_B data will be a composition of signals for different incidence angles and will contain significant influence of the wind speed and azimuth. The inclusion of several incidence angles will require a weigthing function when establishing the modelled T_B parameter. This, together with averaging in space and time, will increase the accuracy of the modelling T_B.

Our study have used a flat zero wind field, and this is a very strong simplification. In an operational monitoring system we will have access to wind field data which will be important input to the T_B function when establishing the modelled T_B field. As for other data types it will be important to get proper error statisitics for the wind data. Also, we have assumed that the emission model provides a T_B functional which is accurate when representing the modelled T_B in the **HA**. This strong simplification causes the modelled T_B to have strong impact compared to the synthetic T_B in the assimilation.

Based on the results and the described limitations in our study we propose, in forthcoming studies, to

- perform elaborate sensitivity studies where T_B data are generated simulating realistic processing techniques,
- examine the impact of assimilating T_B data together with data from other sensors, e.g., SLA and SST data in a hindcast experiment,
- establish improved emsission models to reduce errors in the inversion to SSS data and in representing modelled T_B datai,
- perform studies to improve the spesification of errors in the integrated data and modelling system,
- implement studies to improve knowledge of how the errors in observations and models propagate through the system.

3.6 WP 3100: GOCE MISSION and impact on operational oceanography

The shape of the geoid surface defines the local horizontal. On land it provides the reference surface for "topography", and over the ocean it would correspond to sea-level as if no currents were flowing. Present knowledge of the Earth's gravity field and its geoid, as derived from various observing techniques and sources, is incomplete. However, within a reasonable time, substantial new understanding will be derived by exploiting satellite based gravity observations from the GRACE mission launched in second half of 2002 and the GOCE mission approved for launch in 2006. GOCE is specifically designed for the determination of the stationary gravity field and its geoid to high accuracy (1 mGal and 1 cm) and spatial resolution (100 km) *SP-1233(1)* (1999); *Johannessen et al.* (2003).

The typical elevation scales of the dynamic ocean surface topography range from 0.1 m to 1 m. In comparison, the accuracy of present geoid models is also multi-decimetric on the scale of many ocean circulation features, in particular at wavelengths shorter than 400 to 500 km. Hence the mean ocean circulation cannot be properly estimated from satellite altimetry at these wavelengths. The use of imprecise geoid models for the determination of the absolute dynamic topography at shorter spatial-scales consequently result in computation of false multi-decimetric topographic signals which, in turn, will lead to false transport calculations of several 10's of Sv (1 Sv = 10^6 m^3/s and is typically corresponding to a heat transport of 5×10^{13} W in the North Atlantic basin). In contrast, the accurate and high-resolution marine geoid, as derived from GOCE, will in combination with satellite altimetry enable new precise estimates to be made of the absolute ocean topography as schematically illustrated in Figure 3.21.

As suggested in Figure 3.21 the altimetric signal will be directly related to the marine geoid provided by GOCE, thus allowing the complete height measurements to be utilised. Complemented with in-situ data and numerical ocean models, the new knowledge of the absolute ocean topography will, in turn, act as an important constraint in assimilation and subsequent calculations of ocean circulation and its volume transport and heat fluxes.

In a recent simulation study by *Provost et al.* (1999) the ocean surface topography obtained using the 1/12° resolution MICOM ocean model of the North Atlantic were partitioned into: a) 1000 km wavelength corresponding to the present situation in which the spatial separation of the dynamic topography from the geoid height is performed adequately;

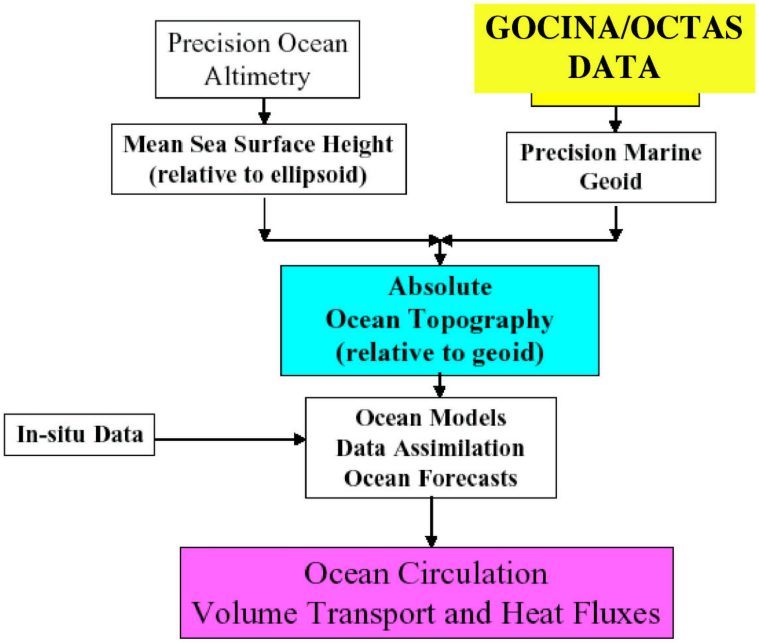


Figure 3.21: Absolute ocean circulation studies from combination of GOCE geoid with precision altimetry.

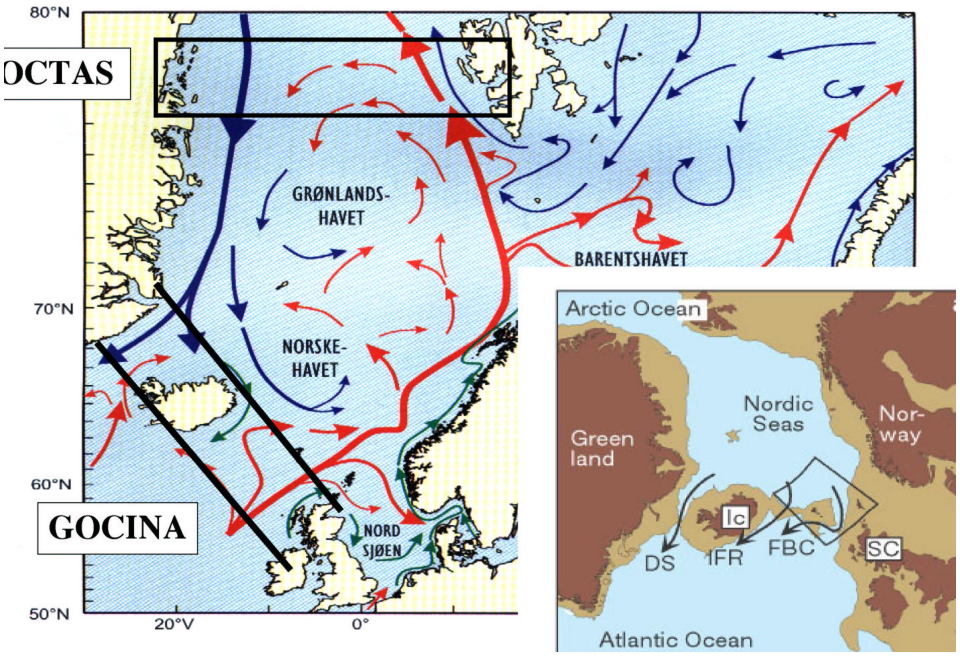


Figure 3.22: Schematic illustration of the upper ocean circulation pattern into and out of the Nordic Seas. In the insert the deep water outflow pattern from the Nordic Seas to the Atlantic Ocean is indicated.

b) 250 km wavelength corresponding to that part of the topography which can be well resolved with GOCE; and c) 100 km being that part of the dynamic ocean topography at the shortest scales which will remain unresolved to satisfactory precision after GOCE. The key findings suggest that many interesting features, such as mesoscale fronts and eddies associated with the Azores Current and western boundary currents can be resolved with GOCE. The only exception is the fairly intense decimetric signals associated with the intense and short-scale Florida Current and Gulf Stream front.

The expected advanced knowledge of the eddy statistics of the real ocean from altimetry, together with knowledge of the precise positions of the ocean jets from altimetry plus gravity (rather than from the assumption of frontal positions by means of sea-surface temperature or hydrographic information at present), will enable much better determination of the role played by the eddies in maintaining the jet components of the circulation. This was demonstrated by *Hughes and Ash* (2001) using satellite derived sea surface temperature fields of the sharp frontal areas within the ACC, most of which will have decimetric signals in the absolute ocean topography, and which it is intended to localise precisely and measure by means of GOCE in combination with altimetry.

Based on the end-to-end simulation results obtained by *Sünkel et al.* (2000) (and briefly reported in *SP-1233(1)* (1999) it is suggested that the accuracy of the marine geoid height provided by GOCE will be about 2.5 mm at 100 km, rather than 1 cm. This means that a 1 cm accuracy can be expected at finer spatial resolution around 70–80 km. An important question in this context is the potential application of the GOCE geoid height data in combination with precision altimetry for ocean circulation studies along isteep and narrow continental shelf breaks and subsequent shelf seas as well as in semi-enclosed seas such as the Mediterranean. The currents topographically steered along the shelf break acts as an open boundary between the coastal regions and the deep ocean and any exchanges of mass and energy between the two regions may therefore manifest itself along the shelf break.

How much of a constraint will GOCE place on the strength of the shelf slope current, and how small a scale can GOCE resolve in this region, given the covariance between gravity, bathymetry, and dynamic topography? The advanced DIADEM/TOPAZ data assimilation systems are based on dynamically consistent estimates of error statistics *Evensen* (1997); *Evensen and van Leeuwen* (2000). These methods ensure multivariate and physically consistent analyses with statistical covariance functions varying in space and time. This allows the extraction of maximum

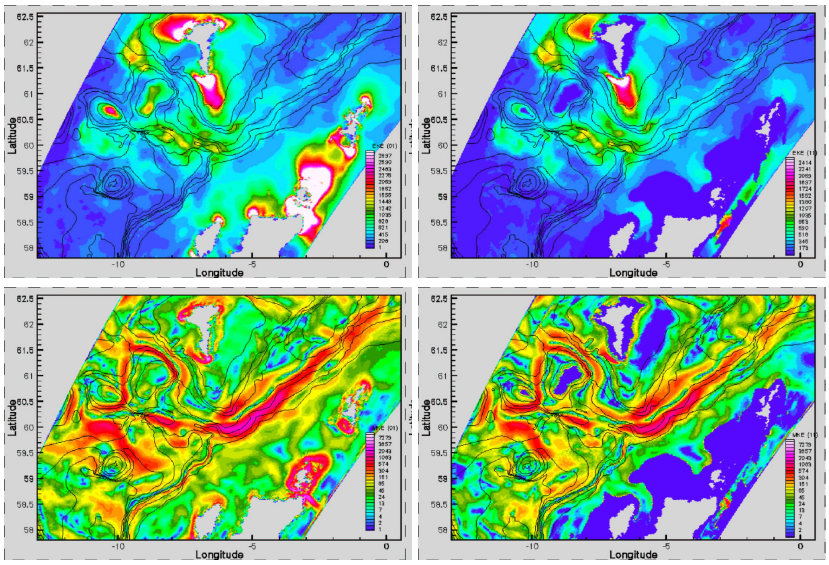


Figure 3.23: Intercomprasion of modelled derived eddy kinetic (upper) and mean kinetic (lower) energy for the surface layer (10 m) left and sub-surface layer (100m) right.

amount of information from satellite surface observations, since vertical projection of information is controlled by the dynamically evolving error covariances in the system. A brief investigation has been initiated with model run to produce fields of mean and eddy-kinetic energy (EKE) for the ocean and shelf region southwest of the Faereo-Shetland Islands (Figure 3.22 and Figure 3.23). These fields reveal how much of the time-varying altimetric signal (sea surface height anomalies and eddy kinetic energy) is found along the continental shelf break versus the deeper ocean basins. The remaining time-invariant mean signal derived from the ocean model is in turn reflecting the strength and importance of the mean flow which is largely controlled and steered by the relative steep shelf break topography with a typical cross-shelf width of 70-100 km.

The partition of the inflow of warm Atlantic Water to the Nordic Seas (Figure 3.21) into the amount of water transported with the mean versus fluctuating part of the inflow (Figure 3.22) clearly reveals the expected impact of topographic steering. The mean kinetic energy in the surface (10 m) and in the sub-surface (below 100 m) layers have practically the same pattern and strength due to the barotropic signal, and equals the eddy kinetic energy in magnitude. The current use of altimeter for studies of the inflow of Atlantic Water to the Nordic Seas is therefore hampered by this strong topographic steering effect.

3.6.1 Summary

The results briefly presented above demonstrate promising impact of the GOCE mission for oceanography. Still we need to build on these and advance the impact simulation studies further in order to adequately promote the mission in the oceanographic community as suggested in Table 3.2.

As revealed in the table, impact on the understanding of the role of the positions, strengths and dynamics of the short-spatial-scale fronts and jets in controlling the ocean circulation has been achieved for some ocean basins, notably in the North Atlantic and Southern Ocean. With the provision of GOCE data the opportunity to carry out real analyses of this kind for the entire global ocean, in combination with auxiliary data, will ensure greater confidence to the construction of the next generation of ocean and climate models. It is furthermore clear that studies should focus on regional and semi-enclosed seas as well as continental shelf break regions to explore the impact at finer spatial scales in the limit for the GOCE observation capabilities. In so doing, it will also be very relevant to consider additional use of local gravity data such as currently being jointly undertaken in the EU funded GOCINA and Norwegian Research

	Open ocean	**Regional and semi enclosed seas and Shelf breaks and coastal regions**
Mean flow	Positive impact documented for North Atlantic	Relevant scales + strength and interaction with bathymetry needs to be further examined (Partly aim of GOCINA and OCTAS
Jets and eddies	Positive impact documented at scales from 250 to 100 km Positive impact on interaction of mean flow with jets and eddies in the Azores current region	Interaction of mean flow with jets and eddies. Can shorter scales 50-100 km features be explored and how do they interact with bathymetry? How important is addtion of local gravity data. (Partly aim of GOCINA and OCTAS)
Heat and volume fluxes	Positive impact documented for the upper ocean ($< 1000m$) and across sharp fronts	The impact of GOCE for heat and volume fluxes needs to be further studied. Usually signals show strong variability on seasonal and interannual scales. Also strongly connected to fisheries in the Nordic Seas
Data assimilation	Positive impact on predictive skills Need to conduct multivariate assimilation using GOCE, altimetry, Argo, SMOS, IR, etc	Assimilation experiments are nedded to study the impact of the GOCE derived geoid The importance of local gravity data must also be examined. (Partly aim of GOCINA and OCTAS

Table 3.2: GOCE Impact Matrix for Oceanography

Council OCTAS projects (refs). Finally, in combination with in-situ data and ocean models, the new understanding gained from present and the future suggested studies, will in turn, act as an important constraint for calculations of oceanic mass and heat transport as well as its impact on global sea level.

The new gravity field knowledge derived from GRACE and GOCE will be very timely in the context of the development of the integrated observing system for the ocean in which continuity of precision altimetry (Jason–1, Envisat RA–2, and possibly Jason–2) together with the Argo profiling float program play a fundamental role.

3.7 WP3200: Assimilation of ice parameters

3.7.1 Introduction

In the polar regions ice is an important component of the climate system, and it strongly modifies exchanges of heat, momentum and moisture between the ocean and the atmosphere. This effect is perhaps best seen in wintertime heat fluxes in the interior of the Arctic ice pack where the atmospheric heat fluxes over open water can be two orders of magnitude larger than over ice.

The thermohaline circulation is affected by the presence of sea ice. This is due to the insulating effect of the ice cover, and because of salt fluxes between ice and ocean when sea ice freezes or melts. The ocean stratification is also an important factor in determining the impact of sea ice on the thermohaline circulation. In the Central Arctic Ocean, the stratification is strong, due to the presence of a cold halocline layer. This layer does, to some extent, reduce the thermohaline circulation in this region. The majority of the thermohaline circulation is therefore believed to be connected to processes in the subpolar seas.

Among the subpolar seas where deep water formation takes place are the Siberian and Laptev shelf seas. Here, it is believed, sea ice formation plays an important part in creating dense waters. The brine released by freezing ice creates dense shelf water, which is believed to sink into the Arctic Ocean, entraining water masses on their way to the deep ocean (*Rudels et al.*, 1999). In the Labrador and Greenland Seas, sea ice often has the opposite effect as seen on the shelves. These seas are “production sites” for deep and intermediate water masses of the North Atlantic. Here, ice is transported from other regions, and excessive transport has the capability of producing buoyancy anomalies in these regions, such as the “Great Salinity Anomaly” (*Dickson et al.*, 1988; *Belkin et al.*, 1998). The buoyancy anomalies can reduce the deep water production, and the variability of the ice transport is therefore an important factor in the variability in deep water production. The effect of such variability on the thermohaline circulation has previously been demonstrated by *Holland et al.* (2001).

In addition to the importance of sea ice for the global climate there is also a need for shorter-term predictions of sea ice concentrations. Offshore operations and weather forecasters would be likely to profit from ice concentration estimates in polar regions. At present we have a reasonably good understanding of polar sea ice concentration through data from passive microwave sensors. The data from the Special Sensor Microwave/Imager (SSM/I) carried on board the satellites of the Defense Meteorological Satellite Program makes it possible to compute nearly daily snapshots of the sea ice concentration.

In contrast to the sea ice concentration data, relatively little information is available on sea ice thickness. The launch of the ESA Cryosat satellite will provide an important new dataset for climate researchers. The availability of the Cryosat sensors will provide users with regularly sampled ice thickness estimates, and therefore much improved estimates of long term changes of the ice cover. This is in contrast to the current situation, where the best dataset is the ice thickness estimates from Upward Looking Sonars, carried onboard British and American submarines. Based on such a dataset, *Rothrock et al.* (1999) inferred a reduction of the average ice thickness of the Arctic from 3.1m in the 1950–1970s period, to 1.8m in the 1990s. This result is in some contrast to results from the Soviet Union ”North Pole” drifting stations. The *Rothrock et al.* (1999) data indicate a trend of -5cm/y, while the ”North Pole” estimates indicate a trend of -2cm/y (*Nagurnyi et al.*, 1994). It should be noted that the conclusions of data from the ”North Pole” stations are unclear as well, since these stations are based on measurements of elastic waves in the ice, which can be hard to directly connect to ice thickness. Furthermore, the ice thickness trends are also investigated by modeling results. *Holloway and Sou* (2002) found that at least some of the trends from *Rothrock et al.* (1999) could be due to changes in the large scale distribution of thick ice, due to changes in wind patterns. They argue that the results of *Rothrock et al.* (1999) are more a result of a statistical ”fluke”, than of a large ice melt in the Arctic. In short, there is some controversy regarding the ice thickness trends in the Arctic. In this situaton the ice thickness estimates from the Cryosat mission will be of great use.

The ability to extend available knowledge, and forecast ice variables, is possible through the use of numerical models

describing the thermal and dynamic properties of the ice. However, in addition to producing forecasts it is essential also to produce reasonable estimates of the uncertainty in the model forecast. Furthermore, to reduce uncertainties in the model forecast it is essential to have an initial model state which is as close as possible to the truth. This is where the different data assimilation techniques come into play. Furthermore, the use of models and data assimilation is not only related to forecasting, but can also be used to produce a reanalysis.

The focus of this work is to demonstrate the assimilation of sea ice concentration and sea ice thickness in a coupled sea ice ocean model using the Ensemble Kalman Filter (EnKF; *Evensen*, 1994). The EnKF is a sequential and ensemble–based data assimilation technique; an ensemble of model states is run in parallel and the analysis is computed at discrete times, using only information available at that time. Model error statistics, necessary for the analysis, are calculated from the ensemble of model states. This makes the EnKF attractive due to its relatively easy implementation compared to other assimilation schemes. There is, for instance, no need to compute adjoint equations or a tangent linear operator. This point is even more important for coupled models, where different time scales can be important for the different model components. Because of its simple implementation and its capability for describing error statistics in nonlinear models, variants of the Ensemble Kalman filter have been widely used in geophysical applications (*Evensen*, 1994; *Houtekamer and Mitchell*, 1998; *Reichle et al.*, 2002; *Haugen and Evensen*, 2002; *Natvik and Evensen*, 2003a). A short presentation of the EnKF was given in Section 3.3

The model system used in this study is the HYbrid Coordinate Ocean Model (HYCOM; *Bleck*, 2002), together with a dynamic-thermodynamic sea–ice model. The sea ice model consists of the Elastic-Viscous-Plastic dynamic model of *Hunke and Dukowicz* (1997), and a thermodynamic module from *Drange and Bleck* (1996). The system is forced by atmospheric variables from the European Centre for Medium–range Weather Forecasting (ECMWF).

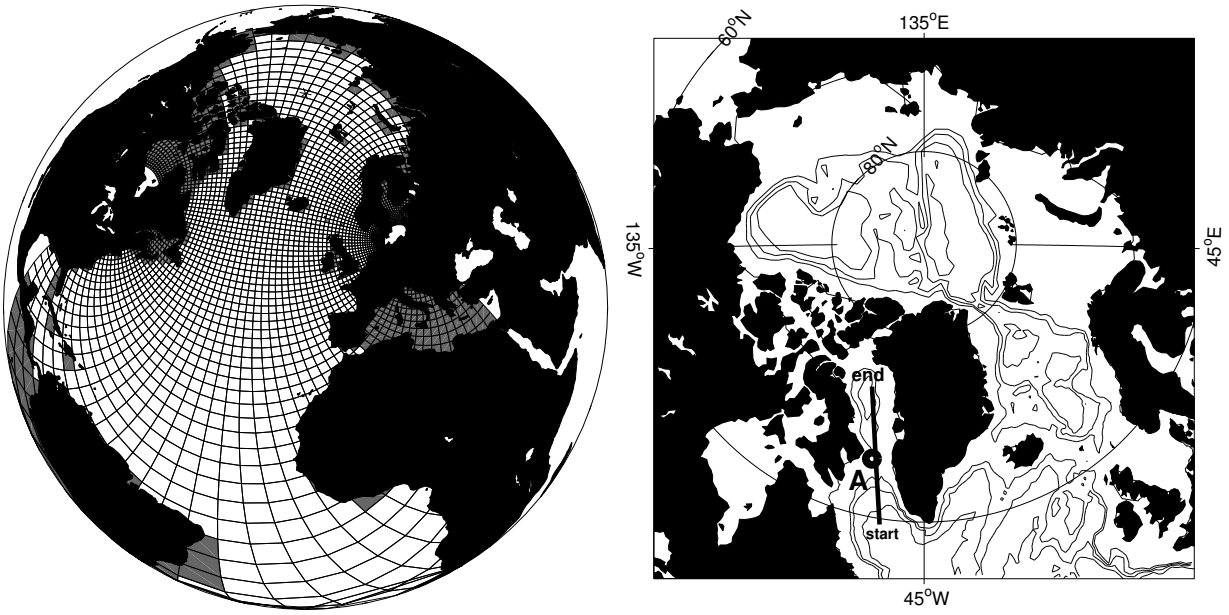
The assimilation is performed with the EnKF, where the assimilated variables are sea ice concentration and sea ice thickness. We do not assimilate ice concentration and ice thickness at the same time, but treat assimilate each variable in seperate experiments. The assimilation scheme is multivariate, so both ocean and ice model variables are updated in the analysis. The use of variables from both ocean and ice model components in the assimilation scheme is essential, as the properties of the oceanic mixed layer are strongly dependent upon the presence of ice. An analysis update in ice concentration/thickness with no accompanying modification of surface layer properties could result in a state where newly introduced ice melts immediately.

To our knowledge, there have been no previous attempts to perform multivariate ice variable assimilation into this type of model system with the EnKF scheme. The focus of the study is therefore largely upon the methodology and the realism of the multivariate scheme. The outline of this study is as follows. In Section 3.7.2 we will briefly describe the model system, whereas the EnKF was described previously in Section 3.3. The study continues with the experiences from the ice concentration experiment first, and then the experiences from the ice thickness assimilation. A short summary is given in the end.

3.7.2 Model setup

The ocean model is the HYbrid Coordinate Ocean Model (HYCOM; *Bleck*, 2002), which is a further development of the Miami Coordinate Ocean Model (MICOM; *Bleck and Smith*, 1990). The MICOM model uses density as the vertical coordinate. The main advantage of isopycnic coordinates lies in their ability to maintain the properties of water masses which does not communicate directly with the surface mixed layer. In the interior of the ocean, mixing is believed to mainly occur along neutral surfaces (*Montgomery*, 1938), which for most situations are relatively close to isopycnic coordinate surfaces. One of the drawbacks of the MICOM model is its relatively simple parametrization of the oceanic mixed layer, described by a bulk mixed layer model (*Gaspar et al.*, 1990).

The major improvement in HYCOM relative to MICOM is the introduction of so-called hybrid layers, which does not necessarily need to keep a given target density. This approach does not allow density layers to become massless, and allows for high vertical resolution close to the surface of the ocean. This has in turn facilitated the introduction of more sophisticated vertical mixing schemes, such as the K-Profile Parametrization (KPP; *Large et al.*, 1994), which



(a) (b)

Figure 3.24: Figure 3.24(a) Illustration of the model grid. Areas masked in the model are marked as gray, while areas outside the grid are not included in the model. Note that only every second grid line is shown. Figure 3.24(b) Part of the model domain with the Arctic Ocean. Also shown is a section used in the text (section A), and a point on this section (circle), which we will refer to as point 'O'.

is used in this study.

The model grid used in this study has a resolution focus in the Nordic Seas with closed boundaries in the Bering Strait and the South Atlantic. A part of the model grid is shown in Figure 3.24(a). It was created with the conformal mapping tools of *Bentsen et al.* (1999), and has grid sizes ranging from 100 to 150 km in the Arctic. The vertical discretization uses 22 isopycnal layers, which in σ_0–coordinates range from 21.8 to 28.11. Note that the lightest layers in this discretization are primarily used to describe the surface mixed layer, as they are usually too light to describe interior water masses of the ocean. The lightest layers become the surface layers in the hybrid coordinate formulation.

The ice thermodynamics model used has many features in common with the "0–layer" ice thermodynamic formulation of *Semtner* (1976), which ignores the specific heat of the ice. In the limit of zero heat capacity of ice the heat conduction equation gives the vertical temperature profile in the ice as a linear function. The conductive heat flux has, as a result of this, the same absolute value at the surface and bottom of the ice slab. The thermodynamic model also includes a snow layer, and a linear temperature profile is prescribed through the snow as well. A detail which will be commented later on is the thermodynamical formulation of heat exchange between water and ice. The present model uses a simplified formulation of heat exchange using so-called "infinite diffusivity" (*Holland and Jenkins*, 1999). In this formulation, any heat available for sea ice melt in the upper ocean layer is immediately used to melt ice. The available heat is determined by the upper ocean layer temperature deviation from the freezing point of the ocean. One consequence of this is that no ice will be present in a model grid cell when the sea surface temperature is above the freezing point of the ocean. For full details of the ice thermodynamic model we refer to *Drange and Bleck* (1996).

The ice dynamics model is the Elastic–Viscous–Plastic (EVP) ice rheology of *Hunke and Dukowicz* (1997). The EVP model presents an alternative to solving the traditional Viscous–Plastic model (VP; *Hibler*, 1979) by introducing an elastic component to the rheology equations. The elastic waves dampen out when solving the dynamical equations,

and the resulting solution approach the one obtained by the VP model. The benefits of the EVP rheology is that it allows for an explicit parallel solution of the VP equations. The numerical implementation of the EVP model also shows better response to rapid changes in forcing of the sea ice component, relative to numerical implementations of the VP model (*Hunke and Dukowicz*, 1997). The ice dynamic and thermodynamic models have been solved for the same model grid as the ocean model, shown in Figure 3.24(a).

The synoptic forcing used were temperature, winds and humidity determined from dew point temperatures, fields which all were acquired from the European Center for Medium–range Weather Forecasting (ECMWF). Clouds are based on climatologies from the Comprehensive Ocean and Atmosphere Data set (COADS; *Slutz et al.*, 1985), while precipitation is based on the climatology of *Legates and Willmott* (1990). River input is modeled as a negative salinity flux, and the river sources in the Arctic include the Lena, Ob, Kotuy, Dvina, Yenisei and the Mackenzie rivers (*Dümenil et al.*, 1993; *Aagaard and Carmack*, 1989). At the surface the ocean model uses temperature and salinity relaxation towards Levitus Climatologies (*Levitus et al.*, 1994; *Levitus and Boyer*, 1994), with a common relaxation time scale of 50 days.

The focus of this study is upon the effects of the Ensemble Kalman Filter Scheme in the coupled model context. The model fields were initialized from a realistic model state used in the EC MAST–III project DIADEM (MAS3-CT98-0167) and integrated for one year prior to the experiment. The generation of the full ensemble of model states used in the EnKF will be described in the next section.

3.8 WP3200 – Ice Concentration assimilation

3.8.1 Ice concentration assimilation – Experiment Setup

The state vector $\boldsymbol{\psi}$ was chosen to include both ocean variables and ice variables as these are closely related, particularly for the ocean surface. For each of the 22 layers of the ocean model we included salinity, temperature, layer thickness and velocity. In addition the barotropic velocity and barotropic pressure components were included along with ice concentration and ice thickness from the ice model.

This gives a total of 115 variables for each grid cell. The analysis (3.3) can be calculated for the full model state vector $\boldsymbol{\psi}$. For this application the dimension for the state vector $\boldsymbol{\psi}$ is

$$n = \dim \boldsymbol{\psi} = 140 \times 130 \times (22 \times 5 + 5) \approx 2 \times 10^6 \,, \tag{3.8}$$

which can result in problems if we only have 100 ensemble members forming a basis for this vector space (the analysis update is in reality a linear combination of the ensemble members $\boldsymbol{\psi}_i$). A common practice in data assimilation for high–dimensional systems is therefore to look at the problem locally, meaning that each grid cell value is updated using observation values in a radius of influence r_0 around the grid cell. In this way the 100 ensemble members will better describe the vector space (now with dimension 115). Thus the local analysis should make the problem better behaved and was chosen for this study. Note that the inclusion of variables other than ice concentration will also force an update of these variables because they can be negatively or positively correlated with ice concentration through the ensemble covariance matrix.

During the integration of the ensemble members there is a need to incorporate the effect model errors have on the evolution of the ensemble members. How to best do this is a study in its own right. One approach could be to add random fields to the ocean layer interfaces and ice thickness in the model at the analysis time. The approach used here is to add pseudo–random fields, with a prescribed length and time scale, to the ECMWF forcing fields. This procedure is a crude way of incorporating the effect of model errors $d\boldsymbol{\beta}_i$ in equation (3.1) into the ensemble, and is the same as assuming that the dominant errors are in the forcing data. Furthermore, this approach does not disturb model dynamics in an unphysical manner. See the appendix in *Evensen* (1994) for a description of the pseudo–random fields added to the forcing, and Table 3.3 for a description of the random forcing component parameters.

Table 3.3: The table shows some of the parameters used in the assimilation experiment.

Parameter	Explanation	value
N	Number of ensemble members	100
r_0	Observation radius of influence	100 km
σ_c	Obs. error standard deviation	0.05
r_c	Obs. error decorrelation length scale	10 km
r_β	Decorr. length for random forcing	~ 1000 km
τ_β	Atm. stress standard deviation	0.015 N/m^2
T_β	Atm. temp. standard deviation	3.0 K

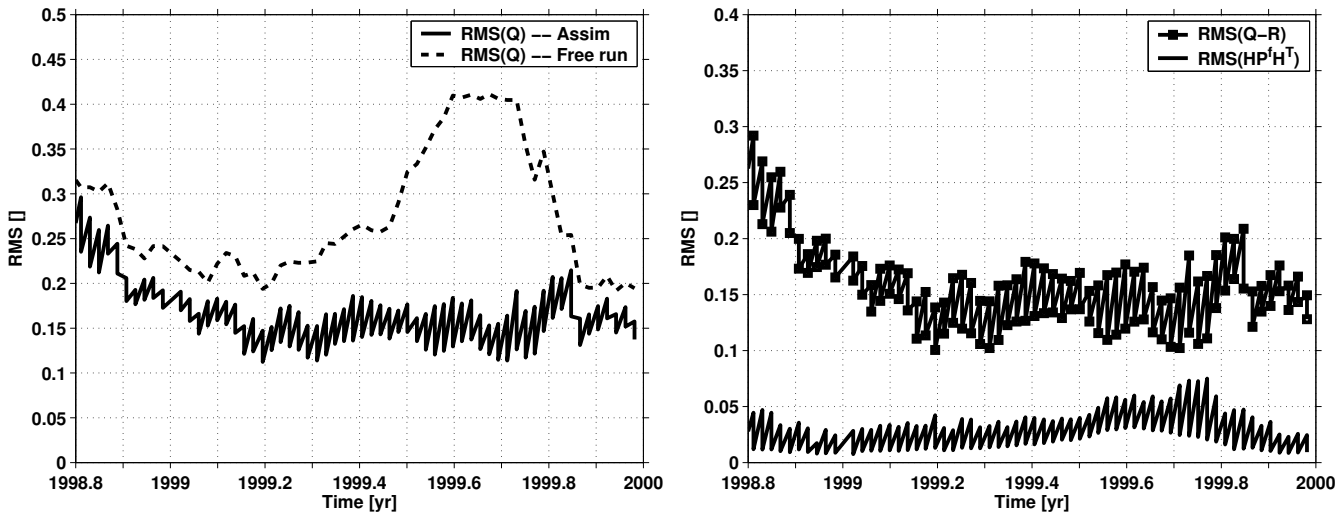
To calculate sea ice concentration, brightness temperatures from the Special Sensor Microwave/Imager (SSM/I) are used. The brightness temperatures are supplied by the National Snow and Ice Data Center (*http://nsidc.org*) on a 304×448 grid using a polar stereographic projection. A variant of the NORSEX algorithm (*Svendsen et al.*, 1983) is used to calculate the ice concentration from brightness temperatures, where channels at 19.35, 22.24, 37.00 and 85.80 GHz are used. The ice concentration was calculated on the same grid as the original brightness temperature grid from the NSIDC, and then averaged over 5×5 grid cells to give an observation vector $\boldsymbol{d}$ of somewhat smaller size and resolution characteristics closer to that of model grid cells in the Arctic. The observation error variance for the elements of $\boldsymbol{d}$ was set to 0.0025 giving an error standard deviation of 0.05 for the ice concentration. This error variance was fixed throughout the year, although the real error is likely to increase in summer due to the presence of melt ponds on the sea ice. *Emery et al.* (1994) give a value of $\sim 5\%$ for the winter time ice concentration error and bias when using the TEAM and bootstrap algorithms. The error for the NORSEX algorithm is probably of the same magnitude and larger in summer. The observation error decorrelation length scale r_c was set to 10 km for the experiments. Compared to the spacing between the observations this is relatively low, and the result will be an error covariance matrix which is approximately diagonal.

The initial ensemble was generated from a realistic model state, to which we added random fields. For each ensemble member, new ice concentration and ocean model layer thickness fields were generated by adding smooth pseudo–random fields (*Evensen*, 1994) to the original model state. All other variables of the individual ensemble members were identical to the original model state. The random fields added to the ocean model layers were vertically correlated and had a horizontal decorrelation length of ~ 400 km. The standard deviation of the random fields were set to 10% of the original ocean layer thickness, and the ocean layers were finally adjusted so that no barotropic waves were generated by this procedure. In a similar manner, the random fields added to the ice concentration had the same decorrelation length, and the standard deviation were set to 10% of the original ice concentration. The ensemble of model states generated this way has a distribution which through equation (3.2) describes our confidence in the initial ensemble average.

The ensemble of model states was integrated for a month before starting the experiment. The experiment covers the period from Julian day 289 in 1998 up to January 2000. Parallel to the ensemble run a free-run model was also integrated for comparison. This model run used the standard ECMWF forcing fields with no pseudo–random fields added, and has an initial state equal to the initial ensemble mean. Some of the relevant parameters for the assimilation experiment are given in Table 3.3.

3.8.2 Impact of assimilation

The assimilation of ice concentration has an effect upon all the variables of the coupled model. In this section we examine the analysis updates and the cumulative effect of assimilation. The focus is put on the ice model variables, and the surface layer of the ocean model.



(a) (b)

Figure 3.25: 3.25(a) The solid line shows the evolution of the innovation RMS for the assimilation experiment, while the dashed line shows the innovation RMS for the free-run model. 3.25(b) The solid line with square marks shows the RMS of the innovation covariance matrix minus the observation error covariance matrix. The solid line shows the RMS of the error covariance matrix of the model. Note that all RMS estimates are taken only over observation points where either the model or the observations have an ice concentration larger than 0.05

3.8.2.1 Innovation vector

Given that the object of the assimilation is to introduce observations into the model state, we should see the impact of the data assimilation in the distance between observations and model state. The success of the assimilation can to some extent be examined from this distance. The vector difference between observations and model state in the observation space is frequently referred to as the innovation vector. The innovation vector of member k is given as

$$\boldsymbol{\lambda}_k = \boldsymbol{d}_k - \boldsymbol{H}\boldsymbol{\psi}_k \,. \tag{3.9}$$

The second–order moment of this becomes

$$\boldsymbol{Q} = \overline{(\boldsymbol{d} - \boldsymbol{H}\boldsymbol{\psi})\,(\boldsymbol{d} - \boldsymbol{H}\boldsymbol{\psi})^T} \,, \tag{3.10}$$

and a measure of the distance between observations and model is here given as

$$\mathrm{RMS}(\boldsymbol{Q}) = \sqrt{\frac{1}{m}\mathrm{trace}(\boldsymbol{Q})} \,, \tag{3.11}$$

which corresponds to the square root of the mean innovation variance, where the variance is taken in the measurement points. The evolution of the innovation RMS is shown as the solid line in Figure 3.25(a). To make it easier to interpret the results, the average innovation is taken only over observation points where at least one of the model members, or the original observation has an ice concentration larger than 0.05.

Noticeable in Figure 3.25(a) are the discontinuities which occur at the times when the analysis is performed, indicative of how the innovation RMS is reduced when the model ice concentration is adjusted by the analysis. Also to be noted is the strong RMS growth which occurs after the analysis is performed. In many cases the innovation growth between two analyses exceeds the RMS reduction from the first analysis. Further into the time series however, there is a clear cumulative effect of the assimilation. At the beginning of the experiment the innovation RMS exceeded 0.30, but a

while into the experiment it is reduced to approximately 0.15, and stays close to this value for the remainder of the experiment.

A free–run experiment was run for the same time period as the assimilation experiment. The evolution of the innovation RMS for the free–run experiment is shown as the dashed line in Figure 3.25(a). Comparing the assimilation experiment with the free–run experiment, we see that the RMS–values start out the same. They quickly diverge, however, and the innovation RMS of the free-run experiment increases relatively to the assimilation experiment. Note that both the free-run and the assimilation experiment show reduced innovation RMS in the winter of 1999, relative to their initial values.

In winter, the large atmospheric temperature gradient across the ice edge has a strong impact upon the modeled ice concentration fields, with and without data assimilation. This ice edge is prescribed in the ECMWF forcing fields, partly by use of SSM/I data, and as a consequence, the atmospheric forcing alone will serve to pull the model towards the observations. This explains some of the RMS reduction in winter, for both experiments.

Around time point 1999.2, the RMS of the free–run experiment starts to move further away from the observed values, and as summer approaches, we start to see large differences between the experiments. The atmospheric temperature gradient across the ice edge is smaller in summer, meaning that the ECMWF ice edge is not so effective in determining the modeled ice edge. This in turn means that the assimilation will have a larger impact in summer, relative to the free–run experiment. At time point 1999.6 the difference between the model and free-run reaches a maximum, as the free run has an innovation RMS of ~ 0.4 while the assimilation experiment has a innovation RMS of ~ 0.15. Towards the end of the time series the RMS–values of the free–run innovation are rapidly reduced, as the autumn freeze–up starts in the Arctic. We note that both the free–run and the assimilation experiment have lower innovation RMS–values at the start of 2000, than at the start of 1999, with the assimilation experiment being slightly closer to the observations.

The innovations also allow us to make rough estimates of the quality of the estimated errors. Consider the model estimate to be given as a true value plus an error $\boldsymbol{q}'$:

$$\boldsymbol{\psi}^f = \boldsymbol{\psi}^t + \boldsymbol{q}' \,. \tag{3.12}$$

Likewise, let the observation be given as the sum of the true estimate and an error $\boldsymbol{\epsilon}'$

$$\boldsymbol{d} = \boldsymbol{\psi}^t + \boldsymbol{\epsilon}' \,. \tag{3.13}$$

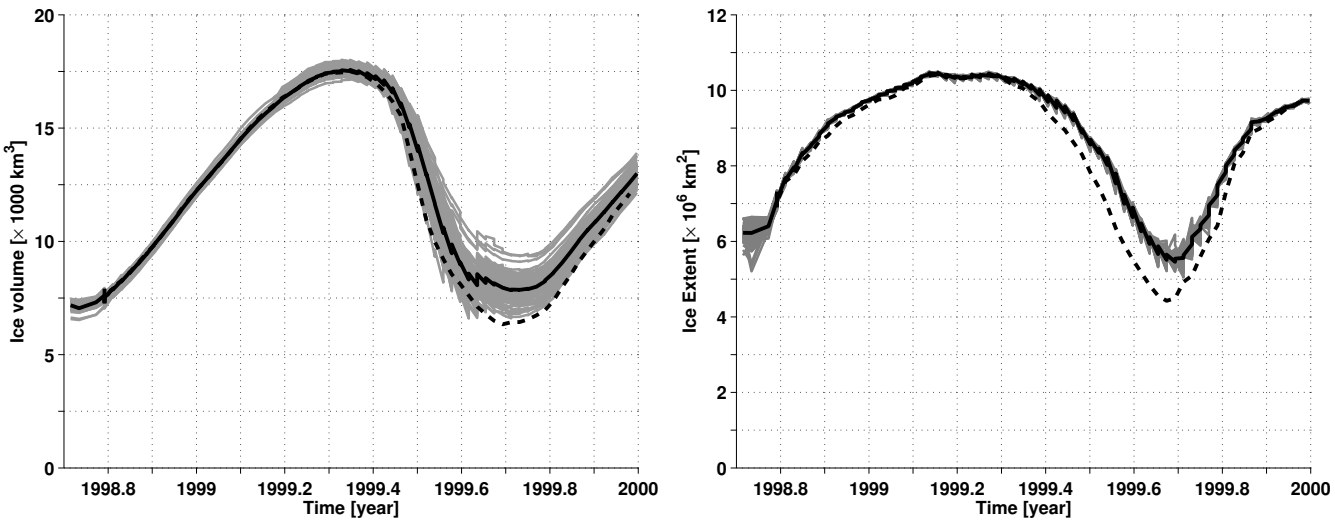
If we assume that the observation and model errors are uncorrelated, we have for the second order moment of the innovation vector

$$\boldsymbol{Q}' = \overline{(\boldsymbol{\epsilon}' - \boldsymbol{H}\boldsymbol{q}')\,(\boldsymbol{\epsilon}' - \boldsymbol{H}\boldsymbol{q}')^T} = \boldsymbol{R}' + \boldsymbol{H}\left(\boldsymbol{P}^f\right)'\boldsymbol{H}^T \,, \tag{3.14}$$

where the primes are used to distinct the theoretical value for the innovation covariance matrix from the real innovation covariance matrix, equation (3.10). By using equation (3.14), we can now get a simple test on the quality of the model error covariance matrix. This is done by comparing $\mathrm{RMS}\,(\boldsymbol{Q} - \boldsymbol{R})$ against $\mathrm{RMS}\left(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T\right)$. If the theoretical estimate given by equation (3.14) is correct for $\boldsymbol{Q}$ as well, then we should have

$$\mathrm{RMS}\,(\boldsymbol{Q} - \boldsymbol{R}) = \mathrm{RMS}\left(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T\right) \tag{3.15}$$

The RMS values of the innovation minus the observation covariance matrix, $\mathrm{RMS}(\boldsymbol{Q} - \boldsymbol{R})$, and the RMS–values of the model error covariance matrix, $\mathrm{RMS}(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T)$, are shown in Figure 3.25(b). Apparent in this figure is that $\mathrm{RMS}(\boldsymbol{Q}-\boldsymbol{R})$ is always larger than $\mathrm{RMS}(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T)$. This difference is most noticeable at the start of the experiment, when $\mathrm{RMS}(\boldsymbol{Q}-\boldsymbol{R})$ is about five times the size of $\mathrm{RMS}(\boldsymbol{H}\boldsymbol{P}^f\boldsymbol{H}^T)$. The values of $\mathrm{RMS}(\boldsymbol{Q}-\boldsymbol{R})$ are quickly reduced from the initial values of ~ 0.25, and seem to stabilize around a value of ~ 0.15. The RMS values of the model error covariance matrix, on the other hand, remain low throughout winter, with a slight increase in the summer of 1999. At this time, the RMS values of the innovation matrix minus the observation error matrix, are approximately three times



(a) Ice volume

(b) Ice extent

Figure 3.26: The evolution of ice volume and ice extent for experiments. The gray lines are individual ensemble members, the solid black line is the ensemble mean, while the dashed line shows the evolution of free–run experiment.

the size of the RMS value of the model error covariance matrix. As the freeze–up starts in the fall of 1999, we notice that the RMS of the model error covariance matrix is again reduced to values similar to the winter of 1999.

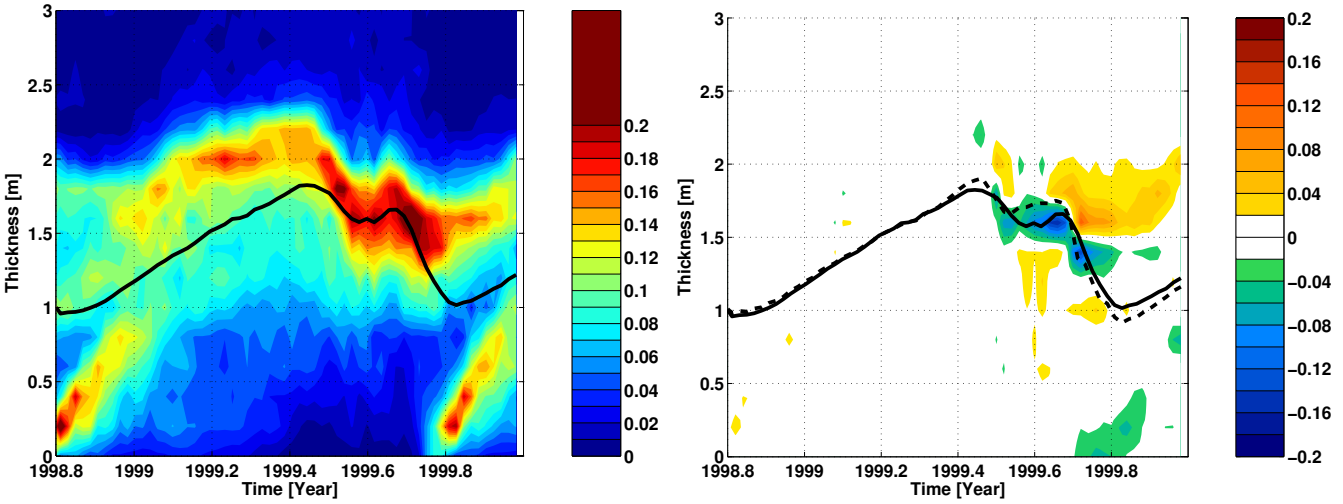
All in all, the model error covariance matrix appears to be underestimated compared to the theoretical estimate (3.15). Note however, that this conclusion can only be drawn if we are confident that the observation errors are of the correct size. We believe the observation errors are of the correct size in winter, while they may be too low in summer. The reason for the underestimate of the model errors can in part be traced back to too low ice concentration variance in the model, especially in winter. The low ice concentration variance results in an underestimate of the model error in the EnKF.

3.8.2.2 Large scale Impact on the ice cover

Two convenient measures of the modeled ice cover properties are ice volume and ice extent. Ice volume is the total volume of ice while the ice extent here denotes the area which has an ice concentration of 15% or more. Only the ice cover of the northern hemisphere is considered. The evolution of ice extent and ice volume is shown in Figure 3.26 and reveals that both the model ice volume and ice extent are changed as a result of the assimilation procedure. The effect of individual analysis updates can be seen through the jumps that occur in ice volume and ice extent for the different ensemble members and for the ensemble mean. Generally, the analysis updates for both ice volume and ice extent increase their values.

The jumps in ice volume are not merely an effect of analysis changes in ice concentration. When the analysis is performed, there is an ice thickness adjustment as well as an ice concentration adjustment brought on by the multivariate assimilation scheme. In other words, the increased ice volume given by the analysis is not just an effect of increasing ice concentration and keeping the ice thickness fixed.

Although the individual updates of ice concentration have a relatively small impact on the ice extent and ice volume, there is a cumulative effect of the assimilation, which is seen by comparison with the free–run experiment. The ice volume and extent is generally higher for the ensemble mean in the assimilation experiment than it is for the



(a) Forecast ice thickness distribution

(b) Difference Forecast and free-run distribution

Figure 3.27: Contour plots of the ice thickness distribution for the entire Arctic. At each time t the Arctic ice thickness pdf can be retrieved as the values along a vertical line. Figure 3.27(a) shows the ice thickness distribution for the forecast, while the solid line shows the mean ice thickness. Figure 3.27(b) shows the difference between the forecast and the free–run ice thickness distributions. The solid line shows the forecast mean thickness, while the dashed line shows the free–run mean thickness.

free–run experiment, although this difference is small prior to the summer of 1999. The highest ice volume for both experiments is seen in March/April, and as the effect of spring and summer melt sets in, the difference between the free-run and the assimilation experiment starts to increase. The difference is seen clearly in summer, when the ice extent and ice volume are at their lowest values. At this time the ensemble mean in the assimilation experiment shows an ice volume which is $\sim 2.000\text{km}^3$ larger than in the free–run experiment. The ensemble mean ice extent at this time is $\sim 1 \times 10^6$ km^2 greater in the assimilation experiment.

As time progresses and the autumn freeze–up starts in the Arctic, the ice volume differences and ice extent differences between the two experiments start to decrease. The impact of the assimilation has therefore the greatest impact upon the summer time ice cover for the time period considered here. The present model system generally underestimates the ice volume and ice extent as compared to observations, so the assimilation of ice concentration improves these estimates.

Both the ice thickness and the ice concentration are changed by the assimilation steps, and this modifies the Arctic ice thickness distribution. In Figure 3.27(a) the evolution of the Arctic ice thickness distribution for the ensemble mean forecast is shown. The contour plots show the distribution of ice with different thickness for the entire Arctic, and how it evolves in time. At a particular time t the probability density function for the ice thickness is given by the values on the vertical line extending from the time axis.

The Arctic ice thickness distribution is influenced by the existence of thick, multi–year ice and of thinner seasonal ice. The distribution of thick ice is increased/reduced by thermodynamic processes and increased by ridging of thinner ice. In autumn and early winter the thickness distribution is bimodal with the modes determined by the multi–year ice ~ 2 m and newly frozen seasonal ice < 0.5 m. The freezing and ridging processes change the thin ice into thicker multi–year ice. In spring/summer, as the melt period sets in, the thinnest and seasonal ice generally melts first, so the ice thickness distribution changes to a more uni–modal distribution.

The general features of the Arctic ice thickness distribution is similar in both the free–run and assimilation experiments. The small changes that occur reveal themselves if we look at the difference between the evolution of the

distributions. In Figure 3.27(b) we see the difference between the evolution of the Arctic ice thickness distributions where we have subtracted the free–run experiment from the assimilation experiment.

Figure 3.27(b) reveals that there are small differences in the thickness distributions before the spring/summer melt sets in. As the melt period sets in, however, the effect of the assimilation start to show. In Figure 3.27(b), the mean ice thickness in the assimilation experiment (solid line) is initially lower than the mean ice thickness in the free–run experiment (dashed line). In the free–run experiment, the thinner ice generally melts first, while in the assimilation experiment, a lot of this ice is retained through the assimilations. This causes the ice thickness distribution in the assimilation experiment to have more ice in the thickness range 0.5–1.4 m, and less ice of thickness 1.5–2.0 m relative to the free–run experiment, see Figure 3.27(b) around time mark 1999.6.

Later on, at the end of the summer, the effect of freeze–up in the Arctic begins to show, with the assimilation experiment having more thick ice (~ 2 m) and less thin ice (< 0.5 m) relative to the free-run experiment. This is connected to lower summer ice concentration in the free-run model. When the freezing sets in for the free-run model, there will be more open water and consequently more thin ice will be created relative to the assimilation experiment. The assimilation experiment has retained a greater fraction of the thick multi–year ice in the central Arctic during the summer, and this shows up as increased occurrences of ice thickness around 2 m.

3.8.2.3 Selected ensemble updates

When the analysis is computed, the variables in the ice and ocean models will be adjusted according to the difference between observations and model state, and according to the error covariance statistics. In the multivariate scheme the properties of the ocean model are updated along with the analysis update in ice concentration. Due to changes in model behavior over the season, there are large differences in the analysis update at different days of the year. A controlling factor in this respect is the north–south migration of the ice edge, which has a strong impact upon the spatial distribution of the ice concentration variance. This impacts the magnitude and location of the analysis updates over the season. To illustrate this we show in Figure 3.28 the analysis update of ice concentration along with updates of upper layer salinity and temperature for the ocean model. The analysis updates are shown for two different times of the year, a typical situation for summer and winter. All plots illustrate the update for the ensemble mean.

For the winter time ice concentration update, Figure 3.28(a), we see that the biggest impact occurs close to land boundaries or along ice edges, which are marked by the forecast 0.7 ice concentration contour. At this time of year the analysis gives a reduction of ice concentration in the Greenland and Labrador Seas. We also see an increase in ice concentration along the coast of Newfoundland, in the Barents Sea and for a small region in the Denmark Strait. The updates in ice concentration are of the order 0.5 and smaller, and cover a relatively small spatial area compared to the total ice area. For the regions far from the ice edge and land boundaries, the ice concentration update is below the plotting threshold shown in the color scale of Figure 3.28(a).

The situation for a winter time update of temperature in the upper layer of the ocean model is shown in Figure 3.28(b), and the updates in surface layer temperature occur over a smaller region than the updates in ice concentration. To some extent the area difference in the figures for ice concentration and temperature updates are due to the plotting threshold, but as will be shown later, it is also due to the statistics of the ensemble, and model limitations. By comparing with the ice concentration analysis update in winter, it should be noted that the sign of the temperature update is mostly opposite to the sign of the ice concentration update. That is, when ice concentration is increased the sea surface temperature is decreased and vice versa. The temperature updates are of order 1° Kelvin or less.

The plot of winter time updates of salinity in the upper layer of the ocean model is shown in Figure 3.28(c). Again we note that the updates occur mainly in the ice edge regions with very small updates far away from the ice edge and land boundaries. If we compare the salinity update with the ice concentration update however, we find that the region with significant salinity updates is somewhat broader along the ice edge. We also see that the connection between the sign of the ice concentration update and the salinity update is less clear than for the connection between temperature and ice concentration update. There is a weak tendency for the sign of salinity updates to be the opposite to that of

the ice concentration updates at this time of year. The salinity corrections are of the order 0.1 psu.

The analysis update in ice concentration for a typical summer situation is shown in Figure 3.28(d). Compared to the winter time situation given in Figure 3.28(a) we see that the update occurs over a larger area. Although the update is still strongest close to the ice edge, the region with significant updates is much broader as compared to the situation in winter. For the specific time given in Figure 3.28(d) we see mostly an increase in ice concentration for the region from the Fram Strait to Sevarnaya Zemlya, in the Baffin Bay and close to the Canadian Arctic Archipelago. Reductions in ice concentration are seen in regions of the Beaufort and East Siberian Seas.

Comparing the summer time update in sea surface temperature, Figure 3.28(e) with the situation in winter, Figure 3.28(b), we see that the temperature update covers a larger region. However, the winter to summer increase in areas with significant updates are much larger for ice concentration than for sea surface temperature. In fact, the sea surface temperature updates are located mainly along the ice edge, as in winter. The sign of the temperature update is mostly opposite to the sign of the ice concentration update, which is the same behavior as in winter.

Finally we see the summer update in sea surface salinity, Figure 3.28(f). We note that the regions with a significant update in salinity are larger than the regions with significant update in ice concentration. It is also important to note that the magnitude of the updates in summer are larger than the magnitude of updates in winter. Differences can also be seen in the sign of the updates as compared to the situation in winter. Whereas the winter time salinity updates had no clear connection with the sign of the ice concentration updates, we see that the summer time salinity updates have the same sign as the ice concentration updates.

The ensemble update plots give hints of the complex statistical behavior described by the model ensemble. Behavior which show considerable changes both in time and space. This behavior will be considered more closely in following sections.

3.8.3 Second order statistics of the model ensemble

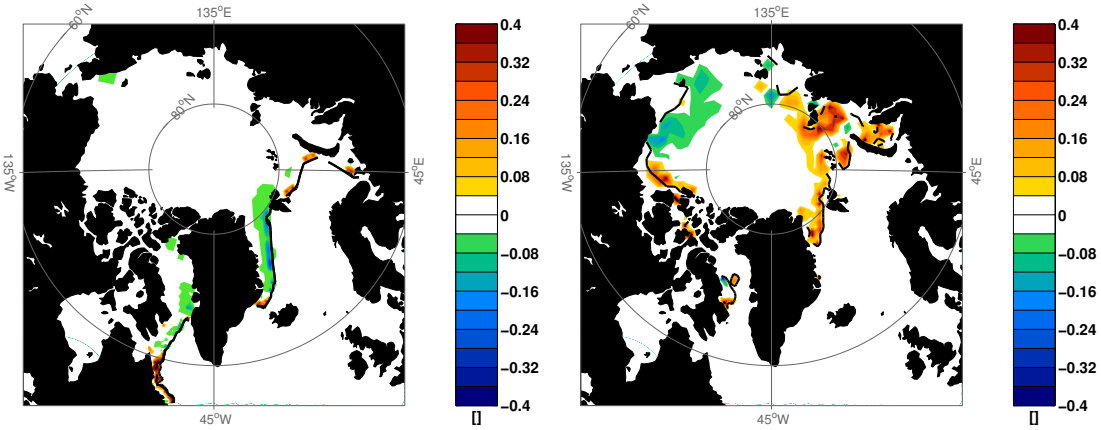
The representation of error statistics is an important property of the EnKF. For a multivariate assimilation scheme, as presented here, it is of particular importance that the covariance between different variables, for instance temperature and ice concentration, is realistic. In this section we examine how the variance and covariance fields from the ensemble change in time and space, trying to illustrate some aspects of the ensemble behavior. An understanding of the processes which determine the statistics is important, and can make it easier to evaluate which model improvements should be implemented for the system at a future stage.

Concerning the evolution of the second order ensemble statistics, it is important to realize that the evolution is dependent upon the prior ensemble statistics and the random forcing applied. The random forcing induces most of the variation in the ensemble. A lot of the following discussion will therefore focus upon the effect the random forcing has upon the ensemble.

In this section, we will use the following notation to denote the covariance, variance and correlation:

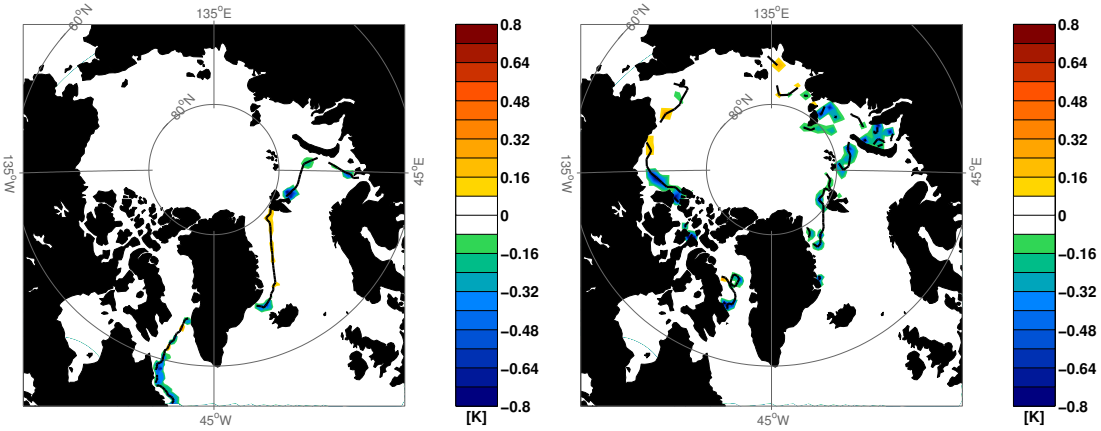
$$\begin{aligned} \mathrm{cov}\left\{F(\boldsymbol{x}_1), G(\boldsymbol{x}_2)\right\} &= \overline{\left(F(\boldsymbol{x}_1) - \overline{F(\boldsymbol{x}_1)}\right)\left(G(\boldsymbol{x}_2) - \overline{G(\boldsymbol{x}_2)}\right)} \\ \mathrm{var}\left\{F(\boldsymbol{x})\right\} &= \mathrm{cov}\left\{F(\boldsymbol{x}), F(\boldsymbol{x})\right\} \\ \gamma\left\{F(\boldsymbol{x}_1), G(\boldsymbol{x}_2)\right\} &= \frac{\mathrm{cov}\left\{F(\boldsymbol{x}_1), G(\boldsymbol{x}_2)\right\}}{\mathrm{var}\left\{F(\boldsymbol{x}_1)\right\}\mathrm{var}\left\{G(\boldsymbol{x}_2)\right\}} \ . \end{aligned} \tag{3.16}$$

One should note that the positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ can denote two different locations. The covariance and correlation can therefore be between two different fields at two different points.



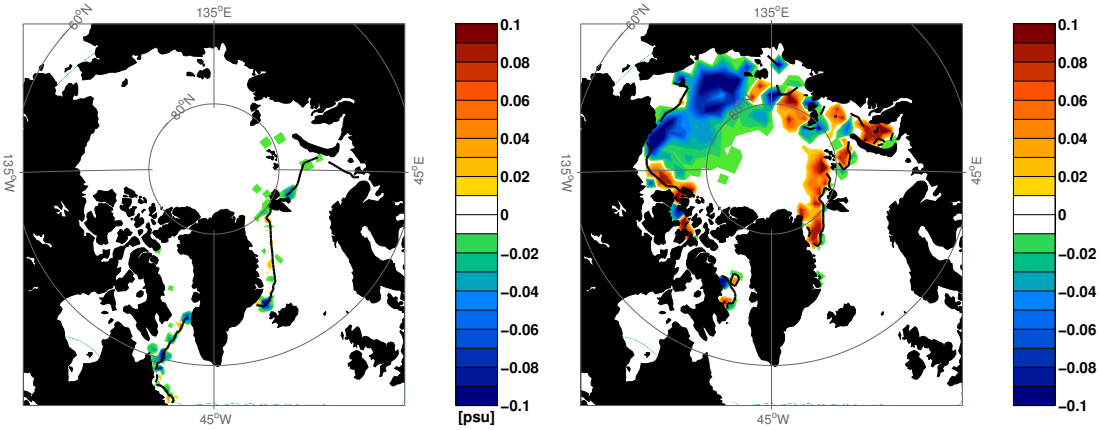
(a) Ice concentration update, day 15

(d) Ice concentration update, day 218



(b) Temperature update, day 15

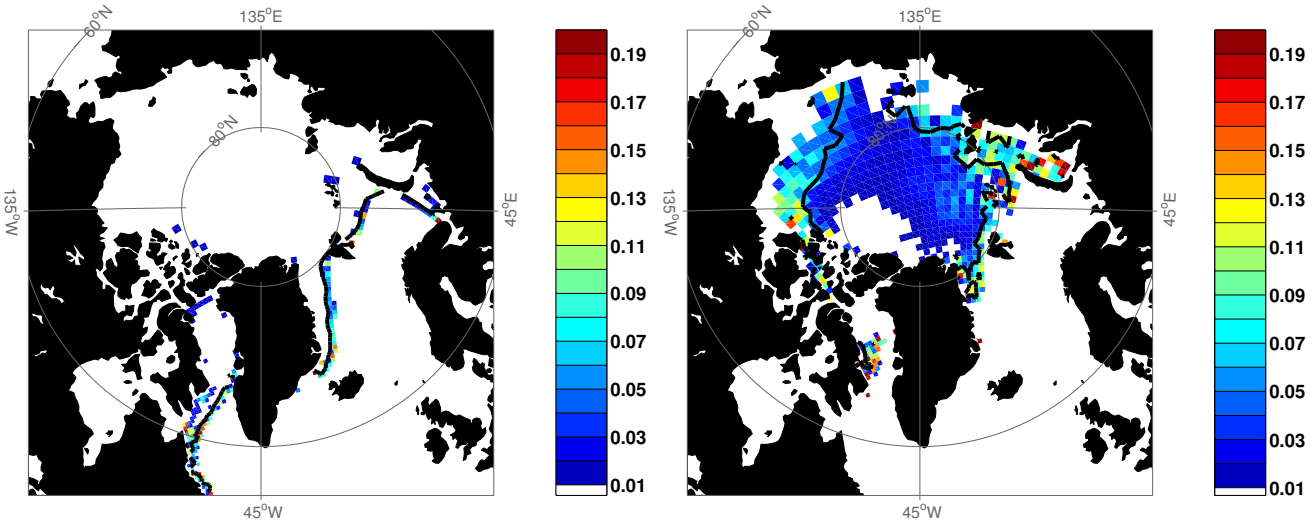
(e) Temperature update day, 218



(c) Salinity update, day 15

(f) Salinity update, day 218

Figure 3.28: Selected analysis updates for the ensemble average. Typical situations for summer and winter in 1999 are shown. Figure 3.28(a), 3.28(b), and 3.28(c) show typical winter time updates in ice concentration sea surface temperature and salinity, respectively. Figure 3.28(d), 3.28(e), and 3.28(f) show typical summer time updates in ice concentration sea surface temperature and salinity, respectively. The contour lines indicate the 0.7 ice concentration contour of the ensemble mean.



(a) 1999 day 15

(b) 1999 day 218

Figure 3.29: Figure 3.29(a) show the ice concentration standard deviation for the individual grid cells as calculated from the ensemble. Also shown is the position of the 0.7 ice concentration contour (solid line). The statistics are based on the prior ensemble, and at the time Julian day 15 in 1999 (winter conditions). Figure 3.29(b) is similar to Figure 3.29(a) but shows the summer situation (Julian day 218 in 1999).

3.8.3.1 Ice concentration variance

The variance of ice concentration in the model is important for the analysis update since it determines the Kalman gain matrix $\boldsymbol{K}_e$. For a local analysis, which operates on individual grid cells, there will be no analysis update of other variables in a grid cell (temperature, salinity etc) unless we have a nonzero ice concentration variance there. The ensemble ice concentration variance in each grid cell therefore deserves some attention.

An example of the ice concentration standard deviation for winter is shown in Figure 3.29(a). Along with the variance plot we have also indicated the 0.7 ice concentration contour of the ensemble average. In Figure 3.29(a) we have masked out the lowest values of ice concentration, leaving only a limited area with significant ensemble variance. Not surprisingly we see that the regions with highest values are the regions which are close to the ice edge, e.g. in the Barents Sea, the Greenland Sea, and the Labrador Sea. The analysis updates of ice concentration, temperature and salinity in winter, e.g. Figure 3.28(a), illustrated that the changes in these variables happened in proximity of regions with high ice concentration standard deviation in Figure 3.29(a).

Thorndike and Colony (1982) showed that a lot of the ice motion over small time scales can be explained by the geostrophic wind speeds, especially far from regions with thick ice. Similar observations apply for the ice dynamics model, and a lot of the ice concentration variance along the ice edge can be explained by the random forcing component applied to each member. As the ensemble members have different degrees of transport across the ice edge this will lead to increased ice concentration variance by the following simple mechanism: Ice advection across the ice edge into warmer waters will, through ice melting, cool the surface waters and make them fresher. If this effect is stronger than the warming effect of mixing (due to a possible increase in the applied surface stress) then the net effect will be a cooling and freshening of the surface layers. Consequently, the more ice is transported across the ice edge, the more likely it is that the ice–free waters will be cooled to such a degree that ice can persist there. This effect links the ice concentration variance to the random forcing because different degrees of transport across the ice edge lead to different ice concentrations. A similar argument applies for the ice-pack-side of the ice edge. If the net heat balance for the water interface is close to zero, then different degrees of ice divergence will lead to variations in the local ice

concentration. Additionally an increased water fraction could lead to enhanced mixing of warm waters from below the sea ice which could amplify local changes in ice concentration.

Another effect of the varying wind forcing for the ensemble members is the increased ice concentration variance close to land boundaries. This can be seen in the Labrador Sea region, where offshore ice transport reduce the model ice concentration in the western part of Davis Strait. Again, the different degrees of onshore/offshore winds in the random forcing component of the individual ensemble members will increase the ice concentration variance in the ensemble.

Important in the ice concentration standard deviation in Figure 3.29(a) are the small values far away from the ice edge, for example in the Central Arctic Ocean. In the present model system there is virtually no ice concentration variance during winter in the Central Arctic Ocean. This is connected to deficiencies in the ice model, where any open water fraction due to ice divergence is quickly closed to the maximum allowed ice concentration, c_{max}, by freezing of new ice. This means that all ensemble members have an ice concentration of c_{max} for large areas in the Central Arctic Ocean, and the ice concentration variance is therefore zero. In reality there are small variations of wintertime ice concentration in the central Arctic Ocean due to small-scale opening of leads (e.g. *Kwok*, 2002). These variations, although small, are important because of the large heat fluxes which occurs over leads in winter.

The low ice concentration variance give ensemble based error estimates which are essentially zero. This means that the EnKF ensemble statistics predict a perfect model and no analysis update is done in these regions, an effect which explain the small analysis updates in the central Arctic Ocean during winter, e.g. Figure 3.28(a). In the future we hope to improve the wintertime behavior of the ensemble in the Arctic Ocean by improving both the ice dynamics and ice thermodynamics models.

If we look at the situation for ice concentration variance in summer we see a somewhat different picture. Figure 3.29(b) shows the ice concentration standard deviation for summer, and it is apparent that there is a much larger region with ice concentration standard deviation above the 0.01 threshold. As with ice concentration variance during winter, the highest values are mainly found close to and seawards of the ensemble–average 0.7 ice concentration contour. In summer, the leads opening up in the Central Arctic Ocean do not freeze immediately as they do in winter. This way the ice concentration variance is allowed to increase in response to differences in ice divergence and melting among the members. The ice transport across the ice edge is not as important in summer as in winter. This is an effect of winds being weaker, and of ice concentration variance imposed by ice melt becoming a more important contributor to ice concentration variance.

Ice concentration variance and the ice edge As illustrated in the wintertime and summertime situations, Figure 3.29(a) and 3.29(b), the spatial pattern of ice concentration variance varies strongly through the season. Moreover, the highest values of ice concentration variance are largely controlled by the location of the ice edge and by land boundaries. To understand how the ice concentration variance pattern evolves in time it is useful to obtain a simple measure of the distance between the ice edge/land boundaries and regions with high ice concentration variance. To this end we will use the field

$$d_e(\boldsymbol{x}) = \text{distance from point } \boldsymbol{x} \text{ to ice edge or land}\,, \tag{3.17}$$

where the ice edge is defined by the 0.01 ensemble average ice concentration contour.

A measure of the positions with high ice concentration variance relative to the ice edge or land boundaries can now be obtained, utilizing the field $d_e(\boldsymbol{x})$. Taking the normalized ice concentration variance as weights, we calculate a mean value of $d_e(\boldsymbol{x})$

$$\overline{d_e(\boldsymbol{x})} = \int d_e(\boldsymbol{x}) \,\mathrm{var}\left\{c(\boldsymbol{x})\right\}\, d\boldsymbol{x} \Big/ \int \mathrm{var}\left\{c(\boldsymbol{x})\right\}\, d\boldsymbol{x}\,. \tag{3.18}$$

Here $\overline{d_e}$ is a weighted average distance from the ice edge/land boundary where regions with high ice concentration variance has been given the most weight. It will be indicative of how far into the ice pack areas of high ice concentration variance are located. Based on Figure 3.29(b) and 3.29(a) we anticipate high $d_e(\boldsymbol{x})$ in summer, and low $d_e(\boldsymbol{x})$ in

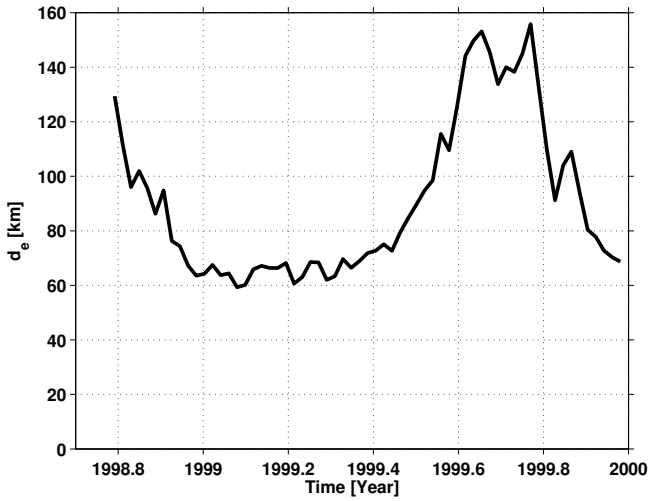


Figure 3.30: The weighted average distance from the ice edge/land boundaries to regions of high ice concentration variance, $\overline{d_e}$ from equation (3.18).

winter. Note that $d_e(\boldsymbol{x})$ will vary throughout the season due to the moving ice edge. Therefore we tried to scale the variable $d_e(\boldsymbol{x})$ with a seasonally dependent parameter describing this effect. However, the analysis did give similar results as without such a scaling, and we therefore show the case without scaling here.

The time series of of $\overline{d_e}$ is shown in Figure 3.30, and shows how $\overline{d_e}$ varies in the experiment. At the start of the experiment $\overline{d_e}$ is relatively high, but when the winter season sets in, $\overline{d_e}$ is reduced and is close to its minimum value at the end of 1998. As was apparent in Figure 3.29(a) the regions with high ice concentration variance are close to the ice edge/land boundary during the winter season, giving low values for $\overline{d_e}$. At the onset of the melt season, the ice concentration variance increases. The ice melt in turn affects most of the Arctic ice pack at the peak of the Arctic summer. The increase in d_e indicates that the ice concentration variance is no longer confined to the close proximity of the ice edge, although it is of largest magnitude there, as was seen from Figure 3.29(b). Towards the next winter the value of $\overline{d_e}$ is again reduced, and is approximately at the same level at the start of year 2000, as at the start of year 1999.

The temporal behavior of the ice concentration variance shown here is largely determined by the background heat fluxes of the ice model. As was mentioned on page 85, the ECMWF forcing has a strong gradient across the ice edge in winter, with very strong negative heat flux in the central ice pack. Because of this, the random forcing component used in the EnKF is not large enough to create any significant ice concentration variance, except close to the ice edge. As a consequence, our ensemble will have a reduced ice concentration variance in winter. In summer the ice cover can develop more freely, which translates into increased ice concentration variance for this time of year.

Ice concentration PDF In Kalman Filter methods, a variance minimizing estimate is used as the final analysis. It is important to remember that the maximum likelihood estimate is not necessarily equal to the variance minimizing estimate. If the forecast probability density function is Gaussian, however, then the analyzed estimate becomes equal to the maximum likelihood estimate.

Figure 3.31 shows a box–whisker plot of the ice concentration in point 'O' in Figure 3.24(b). One can clearly see the collapse of the ensemble ice concentration for this point in summer. In winter, as the ice edge is at some distance from the point, one can also see that the ice concentration variance is very low.

At the times when the ice edge is closer to point 'O', there is significant ice concentration variance close to the ice edge. Note, however that at many times the sample distribution is skewed, e.g December 1998 and late December 1999. The skewness of the sample is most frequently encountered when the median value is close to the zero or maximum ice concentration. At many times, one can also see that the range of values the samples take is very large

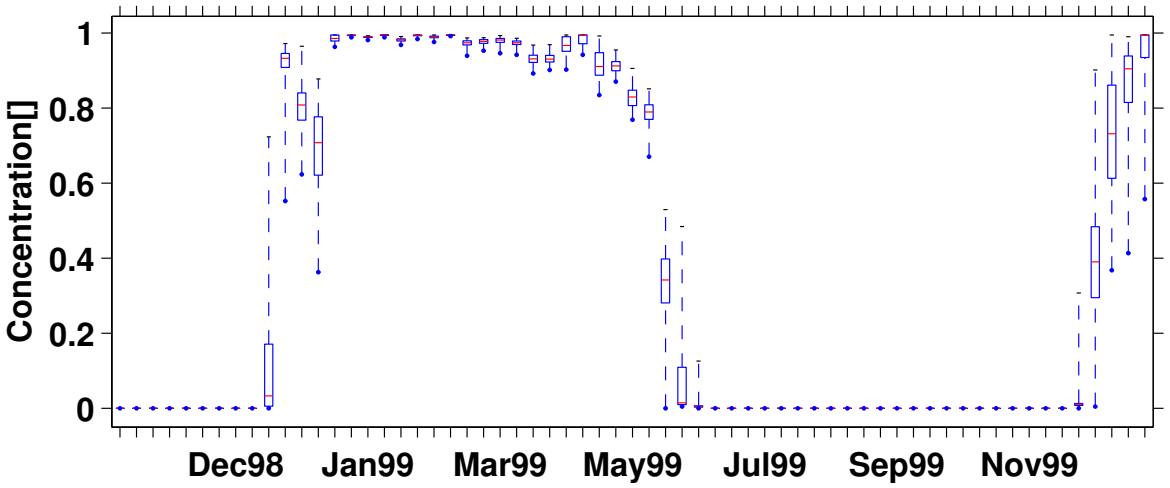


Figure 3.31: A box–whisker plot for the sample of the ice concentration probability density distribution. The samples are from point O for different times in the assimilation experiment. The plots show the lower, middle (median) and upper quartile, and the lines extending from the box gives the extreme values of the data.

compared to the distance between the lower and upper quartile. The skewness, and the large range of values both indicate that the sample is taken from a non–Gaussian distribution.

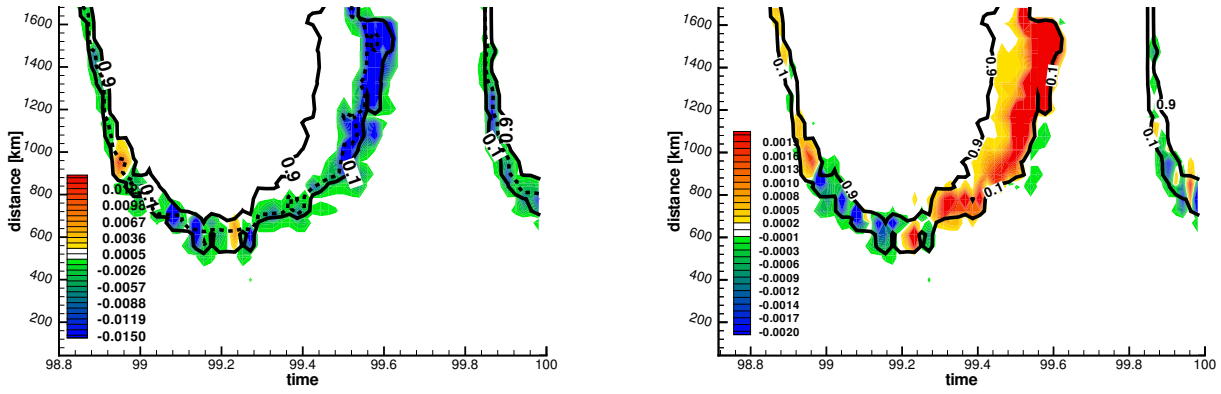
The distribution of ice concentration can of course never be Gaussian, because it is limited to values between 0 and 1. The behavior of the sample distribution of ice concentration should therefore be quantitatively similar to variables which are constrained by upper and lower limits, be it saturation values or other physical constraints. Examples of variables with this type of behavior are the concentration of biological material in biochemical models (*Natvik and Evensen*, 2003a) or soil–moisture (*Reichle et al.*, 2002). An improvement of the Gaussianity of the sample could be possible if suitable transformations are applied to the model state variables prior to the analysis (*Bertino et al.*, submitted).

3.8.3.2 Ensemble Covariance

An important aspect of the assimilation scheme is how well the surface layers of the ocean model are updated. The temperature of the surface needs to be reduced when new ice is introduced, and the update of surface salinity should take into account that sea ice has a lower salinity than that of the ocean surface. All these calculations for the sea surface should be taken care of by the analysis equation (3.3). In this section we will focus upon ensemble–based ice concentration covariances for variables taken from the same grid cell (local covariance), for instance ice concentration and sea surface temperature from the same model grid cell. In the notation from (3.16) the covariance between e.g. ice concentration and temperature of ocean layer 1 in position $\boldsymbol{x}$ can be written as $\mathrm{cov}\left\{c(\boldsymbol{x}), T(\boldsymbol{x}, 1)\right\}$. Due to the nature of the covariances, which are generally located close to the moving ice edge, and whose position depends upon the season, we will make some use of Hovmöller plots in the following.

A Hovmöller plot of the grid cell covariance between ice concentration and sea surface temperature is shown in Figure 3.32(a), while a corresponding plot of the covariance between ice concentration and sea surface salinity is shown in Figure 3.32(b). These plots are for section A in Figure 3.24(b), going from the Labrador Sea into Baffin Bay. In both of these figures we have indicated the location of the ensemble–mean 0.1 and 0.9 ice concentration contours.

This particular section in the Baffin Bay was chosen because it appeared best suited to demonstrate the mechanisms causing the covariance signals, which will be explained later.



(a) $\mathrm{cov}\,\{c(\boldsymbol{x}), T(\boldsymbol{x}, 1)\}$

(b) $\mathrm{cov}\,\{c(\boldsymbol{x}), S(\boldsymbol{x}, 1)\}$

Figure 3.32: Figure 3.32(a) shows a time/distance (Hovmöller) plot for the grid cell covariance between ice concentration and surface layer temperature along section A in Figure 3.24(b). Also shown is the ensemble–mean 0.1 and 0.9 ice concentration contour. Figure 3.32(b) is similar to Figure 3.32(a) but shows covariance between ice concentration and surface salinity. For temperature the dashed line denotes the contour for which all ensemble members have an ice concentration of 0.01 or larger.

Ice concentration – SST covariance The covariances between surface temperature and ice concentration in Figure 3.32(a) are mostly negative or zero throughout the season, and as for ice concentration we find the highest absolute values close to the ice edge. The ice thermodynamic model used here has some limitations which restrict the covariance between ice concentration and sea surface temperature. The most important restriction is the requirement that the sea surface temperature is at the freezing point in order for ice to exist. This effectively groups the connection between ice concentration and sea surface temperature into three different situations.

1. $c > 0$ and $T = T_f$ for all ensemble members
2. $c = 0$ and $T \geq T_f$ for all ensemble members
3. A mixture of members with properties from 1 and 2

Consider first the situation where all ensemble members in a grid cell have a positive ice concentration, i.e, situation 1 above. In this case, variations in sea surface temperature is determined by the freezing point dependence upon salinity, and the covariance with ice concentration should be small in magnitude. From Figure 3.32(a) we see that this occurs in summer; As the region with ensemble–averaged ice concentration between 0.1 and 0.9 broadens, there is also a large region with a very low covariance magnitude between sea surface temperature and ice concentration. This region is roughly bounded by the dashed line, and the ensemble 0.9 ice concentration contour, where the dashed line denotes the limit for which all ensemble members have $c > 0$. In other words, this region corresponds to situation 1 above.

Then consider a situation where some ensemble members have $c > 0$ and some members have $c = 0$, situation 3 above. In this situation we can have ensemble–variations in both ice concentration and sea surface temperature, because there are temperatures above the freezing point associated with $c = 0$. Above freezing point temperatures are associated with $c = 0$ and freezing point temperatures are associated with $c > 0$, resulting in a negative covariance. This negative covariance between ice concentration and sea surface temperature, is seen between the dashed line and the 0.1 ice concentration contour in Figure 3.32(a).

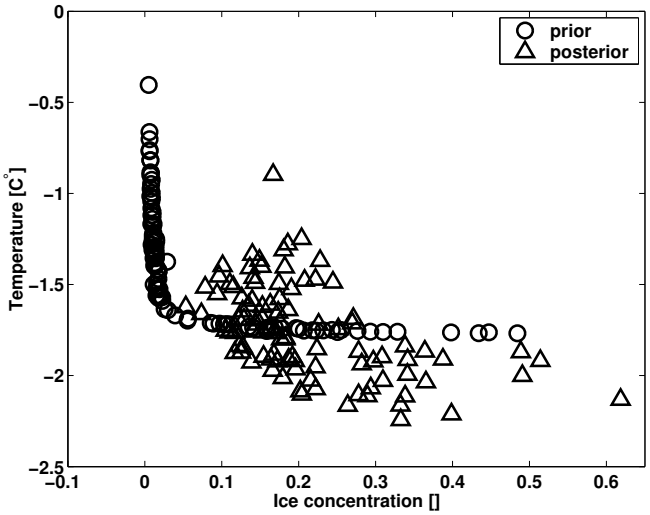


Figure 3.33: Sea surface temperature plotted against ice concentration for the point x_0=(58.85°W 66.05°N) on Julian day 162 in 1999. Shown are scatter plots for the situation prior to (circles) and after (triangles) the assimilation. Note the nonlinear relation prior to the analysis.

Note that the analysis equation (3.3) uses the covariances between surface temperatures and ice concentrations to update the model state. This assumes a linear relation between the variables, which is not the case for the relation between ice concentration and sea surface temperature in our model. This is due to the restriction mentioned above, where sea surface temperature is fixed to the freezing point when ice is present. To illustrate this, we have in Figure 3.33 plotted the ensemble values of sea surface temperature against ice concentration for one of the model grid points along section A (the point marked with 'O' in Figure 3.24(b)). For the case with an ensemble mixture of ice free and ice covered water in a grid cell the nonlinear ensemble relation between ice concentration and sea surface temperature can lead to unwanted effects. In Figure 3.33 this is apparent; prior to the analysis the sea surface temperature is fixed at the freezing point of sea water when ice is present, and the deviations from the freezing point only occur when ice is absent. After the analysis however, this is not the case, and there are cases with ice present and temperatures above the freezing point, and even cases with temperatures below the freezing point. These below freezing temperatures are corrected after the analysis is performed. The above freezing temperatures when $c > 0$ are not corrected, and will result in melting of some of the newly introduced ice in Figure 3.33. As mentioned earlier, the Gaussianity of the model ensemble could be improved by applying suitable transforms prior to the analysis *Bertino et al.* (submitted). This would also improve the nonlinear relation between ice concentration and sea surface temperature.

We note here that the requirement of having sea surface temperature fixed at the freezing point when ice is present is a quite common simplification in sea ice models. It is easy to change this requirement, and introduce more realistic thermal exchanges between the ocean and sea ice (*Holland and Jenkins*, 1999). Note however that more realistic thermal exchange formulas are very sensitive to the sea surface temperature elevation above the freezing point, and this leads to strong melting if ice is present. In the end the net result of such a model improvement might prove to give similar results for the ice concentration – sea surface temperature relation as illustrated in Figure 3.33.

The negative covariance between ice concentration and sea surface temperature is seen throughout the year, not only in the Hovmöller plot for the section in the Baffin Bay. A plot of the weighted average ice concentration – sea surface temperature correlation is shown in Figure 3.34. The average is weighted with the ice concentration variance, and shows how this average correlation has negative values throughout the experiment.

Ice concentration – SSS covariance For the covariance between ice concentration and sea surface salinity in Figure 3.32(b), we have a somewhat different situation than for the sea surface temperature–ice concentration covariance. With salinity we are no longer dealing with ice concentration covariance against a variable which is fixed to a lower

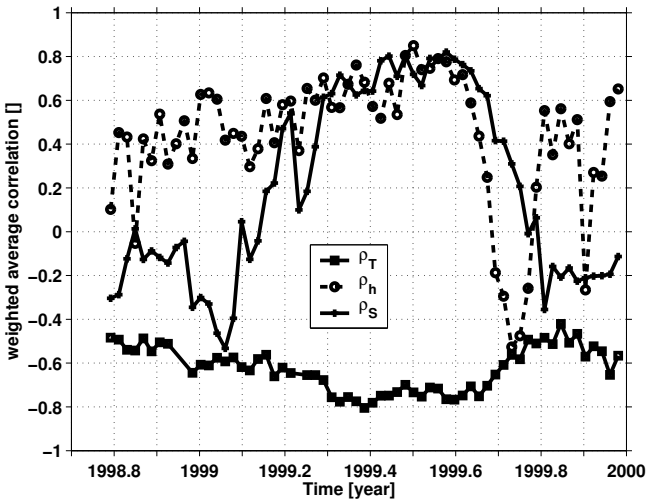


Figure 3.34: Weighted average correlation between ice concentration and sea surface salinity (solid line and star symbols), between ice concentration and sea surface temperature (solid line and squares), and between ice concentration and ice thickness (dashed and circles). The average has been weighted by the ice concentration variance.

bound when ice is present. As compared to covariance between ice concentration and sea surface temperature we also have a broader region with relatively high covariance absolute values. Note also the changes in sign for the covariance in the Hovmöller plot. As the ice edge starts to retreat, we see that the covariance changes sign along the ice edge from being mostly negative to being mostly positive (around time-axis point 99.2).

The cause for this change of sign in sea surface salinity–ice concentration covariance can be traced back to the different mechanisms controlling the ice concentration. Consider first the situation where freezing or melting changes the ice concentration. In this case an increased (reduced) ice concentration leads to increased (reduced) sea surface salinity because brine is released from the freezing ice, and melting ice gives a positive freshwater flux to the ocean surface layers. This results in a positive covariance between sea surface salinity and ice concentration. Note that brine released from newly frozen ice induce vertical mixing which could complicate this argument. If the water below the mixed layer is more saline and cold (typical of the Arctic Ocean ”Cold Halocline Layer”) then the positive covariance becomes even stronger. On the other hand, if the water below the mixed layer is relatively warm this would in turn affect the ice cover and decrease the magnitude of the negative salinity covariance.

As a second factor controlling the ice concentration, consider the case where ice is transported across a thermal boundary separating waters at the freezing point from warmer waters. As the ice reaches the warmer waters it melts, which leads to cooling and freshening of the surface water masses. If the heat content of the mixed layer is low enough, prolonged periods of ice transport help to cool the surface waters to the freezing point and further transport into the region leads to increased ice concentration. Concerning the covariance between sea surface salinity and ice concentration, the result of this transport situation is a reduced sea surface salinity being associated with increased ice concentration, that is, a negative covariance.

Returning to Figure 3.32(b), it seems like the model ensemble is able to describe the two different mechanisms mentioned above. The ensemble differences in ice transport is likely to be causing the mostly negative sea surface salinity/ice concentration covariance during the advance of the ice edge in the winter of 1999, and to some extent during the fall of 1999 and 2000. A strong support for this mechanism can also be found from covariances with the ice velocity normal to and seawards of the ice edge. Unfortunately we were not able to recover the entire record of ice velocity from our data files. The ice velocity data is missing for the spring and summer of 1999, leaving us with ice velocity data for the autumn and early winter of 1998 and 1999. At these times, and for the Baffin Bay section, the covariance between the ice velocity component normal to the ice edge and sea surface salinity (not shown) is consistently negative. This covariance do suggest that the negative ice concentration–sea surface salinity covariance

is a result of seawards ice transport, at least during the advance of the ice edge in fall and early winter. We believe that the transport mechanism persists throughout the year, while the change in sign of the ice concentration/sea surface salinity covariance in summer is due to ice–melt becoming the more dominant effect.

The positive ice concentration–sea surface salinity covariance in summer, seen in Figure 3.32(b), and at the start of the time series, indicates that the ice melt mechanism is dominating the covariance. Note however, that the sign of the covariance during the fall of 1999 changes frequently between positive and negative values. Mesoscale events occurring along the ice edge are important effects here. As the ice heat balance often changes from net melting to net freezing across the ice edge, events which bring warmer air in from the south can disturb this balance and cause melting well into the ice pack. As explained, this has the capability of changing the sign of the covariance between ice concentration and salinity.

The seasonally dependent covariance between ice concentration and sea surface salinity is not only important in the Baffin Bay section, were we have focused our attention. It can also be seen from the area and variance weighted average of the correlation between ice concentration and salinity. An illustration of this is given in Figure 3.34, where the seasonally changing average correlation between ice concentration and sea surface salinity is apparent. At the beginning of the experiment and through much of the winter, the correlation between ice concentration and salinity is negative, with values between -0.5 and 0. The low absolute values indicate that the effect of transport upon the correlation is relatively weak, probably due to spurious warm weather events occurring close to the ice edge. As the summer approaches, the impact of melting starts to dominate over transport along the ice margins, and the correlation between ice concentration and sea surface salinity changes to a positive sign and increases until it peaks at the highest values in summer. It then decreases, and towards the end of the fall freeze-up, negative correlation values are again seen.

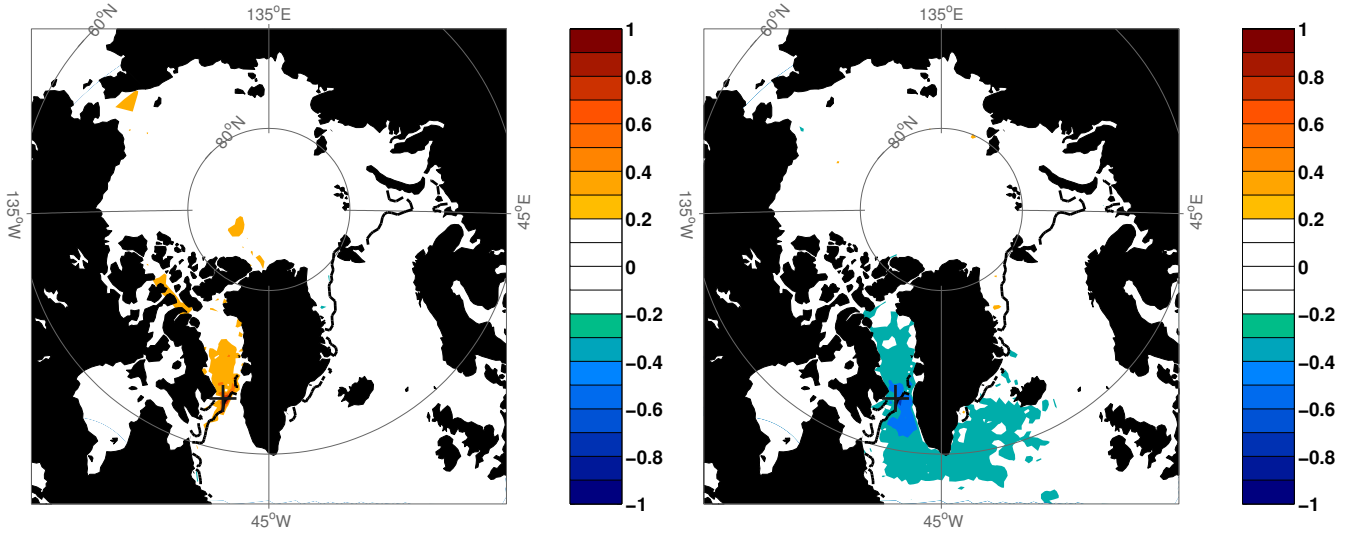
Ice concentration – ice thickness covariance The covariance between ice concentration and ice thickness along the same section as in 3.32(a) reveals a mostly positive covariance between ice concentration and ice thickness(not shown). This is consistent with the idea that “thicker ice lasts longer”, and can explained by reduced horizontal melting for thicker ice. One should note however, that situations arise in which there is a negative correlation, as is evident in Figure 3.34. During the fall freeze-up, a lot of thin newly frozen ice increases the ice concentration, and consequently reduces the thickness of existing ice. Starting from initially equal conditions, this means that the faster the ice cover freezes, the tinner the ice in the model will become. The result of this is a negative covariance between ice concentration and ice thickness.

3.8.3.3 Spatial correlation functions

Another important factor is how the ensemble covariance and correlation varies in the horizontal. We therefore focus the attention on so–called correlation functions. The correlation function used here is given by $\gamma\{F_1(\boldsymbol{x}), F_2(\boldsymbol{x}_0)\}$ where $F_1(\boldsymbol{x})$ and $F_2(\boldsymbol{x}_0)$ denote variables of interest. This gives the correlation between the ensemble members of variable F_2 located at $\boldsymbol{x}$ and variable F_1 located at $\boldsymbol{x}_0$. In the following we keep $\boldsymbol{x}_0$ fixed to a specified value for plotting purposes.

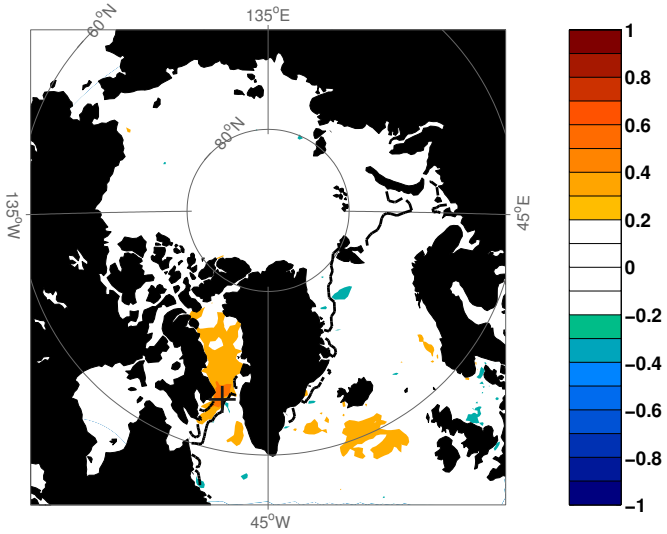
The ensemble correlation functions for ice concentration, temperature and salinity against the ice concentration of the fixed point $\boldsymbol{x}_0$=(58.85°W, 66.05°N) are shown in Figure 3.35. The correlation functions were acquired from the ensemble for day 162 in 1999. The fixed point $\boldsymbol{x}_0$ is identical to point O in Figure 3.24(b). It is located close to the ice edge at this time, at a location in the Davis Strait.

In general we see that the correlation functions reveal complex patterns that are dependent upon past flow history of the ice and ocean. If we look at the correlation function for ice concentration, Figure 3.35(a), we see a positive correlation that decay rather rapidly away from the fixed point $\boldsymbol{x}_0$. Ice concentration is not a conserved quantity in the ice model, and one would therefore expect that the past information from the flow is lost quickly, effects that would lead to such a fast decay of the correlation. The ice concentration correlation function is limited by the ice edges of the ensembles, and is close to zero seawards of the ensemble-average ice edge illustrated in Figure 3.35(a).



(a) $\gamma\{c(\boldsymbol{x}_0), c(\boldsymbol{x})\}$

(b) $\gamma\{c(\boldsymbol{x}_0), T(\boldsymbol{x}, 1)\}$



(c) $\gamma\{c(\boldsymbol{x}_0), S(\boldsymbol{x}, 1)\}$

Figure 3.35: Correlation functions for 1999 day 162, where the fixed point is located at $\boldsymbol{x}_0 = (58.85°\text{W}, 66.05°\text{N})$ and is marked with a plus sign in the figures. The correlation functions are for ice concentration in the fixed point against ice concentration 3.35(a), sea surface temperature 3.35(b) and sea surface salinity 3.35(c).

The correlation function for sea surface temperature is shown in Figure 3.35(b). As suggested by the covariance plots between ice concentration and sea surface temperature, Figure 3.32(a), we see that there is a negative correlation between ice concentration and sea surface temperature for $\boldsymbol{x} = \boldsymbol{x}_0$. This negative correlation also extends at some distance away from the fixed point, and relative to the correlation in $\boldsymbol{x} = \boldsymbol{x}_0$, the ice concentration–sea surface temperature correlations decay at a slower spatial rate than the ice concentration–ice concentration correlation. Also notice that the region with relatively high absolute values of temperature–ice concentration correlation extends beyond the ice edge and well into the Labrador Sea. The surface circulation around the southern tip of Greenland, Cape Farewell, generally follows the coastline in an anticyclonic sense, and water masses are transported northwards along

the western coast of Greenland (The West Greenland Current). These water masses help to maintain the roughly northeastwards direction of the ice edge from Newfoundland towards the West Coast of Greenland. South of Davis Strait a branch of the West Greenland Current crosses over from the West Coast of Greenland towards Baffin Island and finally join up with the Labrador Current (*Chapman and Beardsley*, 1989).

The point $\boldsymbol{x}_0$ used here is situated in the Davis Strait and the waters passing the point $\boldsymbol{x}_0$ should contain a component originating from the West Greenland Current. As water in the West Greenland Current is transported northwards, the ensemble differences in the upstream ocean surface temperature at an earlier time (e.g. at Cape Farewell) contribute to differences in ice concentration at the fixed point, where positive temperature anomalies lead to reduced ice concentration at the fixed point $\boldsymbol{x}_0$. This in turn leads to the negative covariance in Figure 3.35(b) between ice concentration in $\boldsymbol{x}_0$ and ocean surface temperatures in the oceanic upstream direction. These negative correlations can be traced back into the Irminger Sea, although with decaying absolute value with upstream distance. Also note that part of the water that circulate within the Labrador Basin show a relatively high correlation with temperature at the fixed point, probably because it originates in the same region as the water that reaches the ice edge at ($58.85°$W,$66.05°$N).

In Figure 3.35(b) we note that there is some correlation appearing in the Baffin Bay. At this time of the year, there still is a high fractional ice cover polewards of the point $\boldsymbol{x}_0$ in the Baffin Bay. Since the model restricts the sea surface temperature to the freezing point when ice is present, this correlation must come from the freezing point dependence upon sea surface salinity. The correlation function for sea surface salinity is shown in Figure 3.35(c), and close comparison of the ice concentration–temperature and ice concentration–salinity correlations reveals the exact same pattern (with different sign) in the Northern Baffin Bay. This is to be expected, as the model has a linear relationship between freezing point temperature and sea surface salinity.

For salinity we see the strongest correlation with ice concentration from the ice edge and northwards into the ice pack. Roughly speaking this can be attributed to the effect of thicker ice leading to a more persistent ice cover close to the ice edge. As thick ice is transported southwards it is less likely to completely melt away as with thinner ice. Thicker ice from enhanced freezing is again connected to the presence of more saline water in the upper water column leading to the positive correlation pattern polewards of the ice edge. As compared to the temperature correlation function, the correlation with sea surface salinity is much weaker over open water.

Finally we note that spurious correlations occur at regions far from the point of interest, of which some arise due to the spurious correlations in the random forcing fields. This highlights the importance of using local analysis and a radius of influence in the analysis scheme, at least for small ensemble sizes (*Mitchell et al.*, 2002).

3.8.4 Ice concentration assimilation – Summary

In this study we have demonstrated an implementation of the Ensemble Kalman Filter with a coupled sea ice and ocean model. The ocean model used is HYCOM (*Bleck*, 2002), while the ice model uses the Elastic–Viscous–Plastic rheology of *Hunke and Dukowicz* (1997). The observations used in the assimilation were sea ice concentration data, which were calculated using remotely sensed passive microwave data. The ice concentration was assimilated every 7th day for a period longer than a year, allowing us to draw conclusions on the impact of the assimilation over a seasonal cycle. The focus of the study has been on the corrections performed by the EnKF, and on the model ensemble statistics. The statistics of the model ensemble determines the impact of the ice concentration, and a good understanding of this is important in a multivariate ensemble system.

The assimilation experiment was first compared with the results of a model run without assimilation. As expected the difference between modeled and observed ice concentration was smaller in the case with assimilation, than in the case without. This effect was strongest in summer, as can be seen from Figure 3.25(a). In winter we found that the impact of the assimilation was weaker. This is in part due to the ECMWF forcing fields, which depends on SSM/I derived ice concentrations. In winter the strong atmospheric temperature gradient across the ice edge prescribed in the ECMWF forcing to some extent determines the location of the modeled ice edge. As the ECMWF forcing depends on SSM/I data, it will act to reduce the distance between observed (SSM/I–derived) and modeled (ECMWF–forced)

ice concentration. In summer the effect of using SSM/I data in the ECMWF forcing is not as important, because the atmospheric temperature gradient across the ice edge is smaller, and because solar absorption by open water becomes important.

The effect of the assimilation was also seen in the Arctic ice thickness distribution. The main effect of the assimilation was to retain a larger fraction of thick, multi–year ice in summer. The effect of this was seen on the total volume of ice in the Arctic, which was larger at the end of summer in the assimilation experiment than in the experiment with no assimilation. In addition, the effects of individual analysis updates on ice concentration, sea surface temperature and sea surface salinity were illustrated for typical summer and winter situations, Figure 3.28. The winter and summer situations reveal that the strongest analysis updates take place close to the ice edge.

The quality of the EnKF model error estimate was investigated by looking at the innovation RMS–sequence of the experiment. It was seen that the error estimate of the model was too low when compared with a theoretical estimate. This effect was strongest at the beginning of the experiment, while the most realistic estimate of the model errors was seen for the summer season. A lot of the model error underestimate in winter is connected to a collapse of the ice concentration towards a single value, an effect which was most clearly seen in the Central Arctic Ocean. The reason for this collapse can be traced back to a too simple lead parametrization in the ice model. The effect of this is an underestimated ice concentration variance, and too low model error estimates in the EnKF. Throughout the time period examined in the assimilation experiment, the largest model ice concentration variance, and probably the most realistic error estimates, were found close to the ice edge.

In the EnKF the analysis is given as a variance–minimizing analysis, but often one finds that the maximum–likelihood estimate is a better estimate. If the sample distribution of ice concentration is Gaussian, one will find that the variance–minimizing and maximum–likelihood estimates are equal, so this is a property we would like the sample distributions to have. A box–whisker plot of the sample distribution of ice concentration revealed that the distribution showed signs of skewness, and having a larger range than what is expected from a Gaussian distribution. Evidence of non–Gaussian sample distributions were also found in the covariance between ice concentration and sea surface temperature. One should keep in mind that variables such as ice concentration always have non–Gaussian distribution since they are limited to the interval $[0, 1]$. An approach which is worthwhile considering is using suitable transforms, and then do the analysis on a variable which is closer to being Gaussian (*Bertino et al.*, submitted), but this approach has not been further considered here.

The covariance between ice concentration and variables of the ocean model were also considered. The covariances described by the ensemble statistics appear realistic, and we tried to connect them to model behavior. The temperature–ice concentration covariance was generally positive throughout the experiment, while the salinity–ice concentration covariance changed sign depending on physical mechanisms involved. It was suggested that the change of covariance sign was due to different effects of ice transport and local sea ice formation/melting, where transport resulted in a negative covariance, and local ice melt/freezing resulted in a positive covariance. In winter the transport effect can be seen, while in summer the local melt effect is dominant. The spatial correlation functions displayed in section 3.8.3.3 revealed complex patterns emerging for the correlations between sea ice concentration in a fixed point in Davis Strait and surrounding model grid point temperature, salinity and ice concentration.

The present implementation with a coupled ice–ocean model reveals some of the strengths of the Ensemble Kalman Filter. It is relatively easy to implement, only requiring a number of ensemble members to be run in parallel, and there is no need for computing the adjoint equations. This makes it especially suited in the coupled model context, where the different models can have different temporal and spatial scales. The computational demands of the EnKF can be large, but in the current application, it is felt that the complexity of the ensemble statistics favors the EnKF in place of simpler, less computationally expensive methods.

While this study focuses on the effect of assimilating ice concentration, it is understood that it will be beneficial to include data sets which directly describe the ocean surface, such as sea surface temperature and sea level anomalies (*Haugen and Evensen*, 2002). Although the correlation function between ice concentration and sea surface temperature extended to some distance beyond the ice edge, Figure 3.35(b), the ice concentration “influence” is still limited

when compared with the total model domain. The inclusion of sea surface temperature in the assimilation will produce a more direct control of the ocean properties, and cover a larger area than what can be achieved with assimilation of ice concentration alone. Care must be taken, however, as the ice concentration and sea surface properties are likely to be dependent close to the ice edge. This was clearly demonstrated for sea surface temperature in Figure 3.35(b).

To our knowledge there are no published results of studies using the Ensemble Kalman Filter to assimilate ice concentration in coupled sea ice/ocean models. Our main conclusion is that assimilation of ice concentration is a viable way of controlling the ice cover in a coupled ice–ocean model. In the current implementation the ice model is not able to correct the generally underestimated ice thickness of the model, but it gives a consistent update of the properties of the modeled ocean surface when ice is introduced or removed.

3.9 WP3200 – Ice Thickness assimilation

3.9.1 Ice Thickness Assimilation – Experiment Setup

The state vector $\boldsymbol{\psi}$ was chosen to include both ocean variables and ice variables as these are closely related, particularly for the ocean surface. For each of the 22 layers of the ocean model we included salinity, temperature, layer thickness and velocity. In addition the barotropic velocity and barotropic pressure components were included along with ice concentration, ice thickness and ice velocity from the ice model.

This gives a total of 116 variables for each grid cell. The analysis (3.3) can be calculated for the full model state vector $\boldsymbol{\psi}$. For this application the dimension for the state vector $\boldsymbol{\psi}$ is

$$n = \dim \boldsymbol{\psi} = 140 \times 130 \times (22 \times 5 + 5) \approx 2 \times 10^6 \,, \tag{3.19}$$

which can result in problems if we only have 100 ensemble members forming a basis for this vector space (the analysis update is in reality a linear combination of the ensemble members $\boldsymbol{\psi}_i$). A common practice in data assimilation for high–dimensional systems is therefore to look at the problem locally, meaning that each grid cell value is updated using observation values in a radius of influence r_0 around the grid cell. In this way the 100 ensemble members will better describe the vector space (now with dimension 115). Thus the local analysis should make the problem better behaved and was chosen for this study. Note that the inclusion of variables other than ice thickness will also force an update of these variables because they can be negatively or positively correlated with ice thickness through the ensemble covariance matrix.

During the integration of the ensemble members there is a need to incorporate the effect model errors have on the evolution of the ensemble members. How to best do this is a study in its own right. One approach could be to add random fields to the ocean layer interfaces and ice thickness in the model at the analysis time. The approach used here is to add pseudo–random fields, with a prescribed length and time scale, to the ECMWF forcing fields. This procedure is a crude way of incorporating the effect of model errors $d\boldsymbol{\beta}_i$ in equation (3.1) into the ensemble, and is the same as assuming that the dominant errors are in the forcing data. Furthermore, this approach does not disturb model dynamics in an unphysical manner. See the appendix in *Evensen* (1994) for a description of the pseudo–random fields added to the forcing, and Table 3.4 for a description of the random forcing component parameters.

The initial ensemble was generated from a realistic model state, to which we added random fields. For each ensemble member, new ice concentration and ocean model layer thickness fields were generated by adding smooth pseudo–random fields (*Evensen*, 1994) to the original model state. All other variables of the individual ensemble members were identical to the original model state. The random fields added to the ocean model layers were vertically correlated and had a horizontal decorrelation length of $\sim$ 400 km. The standard deviation of the random fields were set to 10% of the original ocean layer thickness, and the ocean layers were finally adjusted so that no barotropic waves were generated by this procedure. In a similar manner, the random fields added to the ice concentration had the same decorrelation length, and the standard deviation were set to 10% of the original ice concentration. The ensemble of model states generated this way has a distribution which through equation (3.2) describes our confidence in the initial

Table 3.4: The table shows some of the parameters used in the assimilation experiment.

Parameter	Explanation	value
N	Number of ensemble members	100
r_0	Observation radius of influence	200 km
σ_c	Obs. error standard deviation	variable
r_c	Obs. error decorrelation length scale	80 km
r_β	Decorr. length for random forcing / $P^\star$	~ 1000 km
τ_β	Atm. stress standard deviation	0.015 N/m^2
T_β	Atm. temp. standard deviation	3.0 K
$\frac{P^\star_\beta}{P^\star}$	$P^\star$ standard deviation (exp. SR3 only)	2000 N/m^2
	$P^\star$ mean	27000 N/m^2

ensemble average.

The ensemble of model states was integrated for two months before starting the experiment. The experiment covers the period from January 1990 in up to January 1991. Two different ensemble integrations were performed, differing in only one detail. The first experiment, designated ”SR2” used the setup already described, with random forcing as specified above.

In the second experiment, designated ”SR3”, we used random fields to disturb an important parameter in the ice dynamics model. In the ice model, the ”strength” of the ice, i.e. its resistance to ice convergence and shear, depends on a crucial parameter, $P^\star$. This ice strength is given as:

$$P = P^\star c h_i \exp(-c^\star(1-c)) \,, \tag{3.20}$$

where $P^\star$ and $c^\star$ are constants, c is ice concentration and h_i is ice thickness. In the original VP and EVP ice rheology models, the parameter $P^\star$ is a constant, while we here will set it as a random field with similar spatial and temporal characteristics as the random forcing fields above. The mean and standard deviation of $P^\star$ is given in Table 3.4. The sole purpose of this approach was to increase the error variance of the ensemble. It should be noted that this approach will have some side-effects on the ice dynamics. For instance, it is possible, due to gradients of $P^\star$, that the ice dynamics model will give resistance to divergence. This is an effect which will not be present in the original formulation of the EVP model, and neither should it. We note here that this heuristic approach does not seem to produce unstability in the model, or any clearly unphysical ice thickness fields.

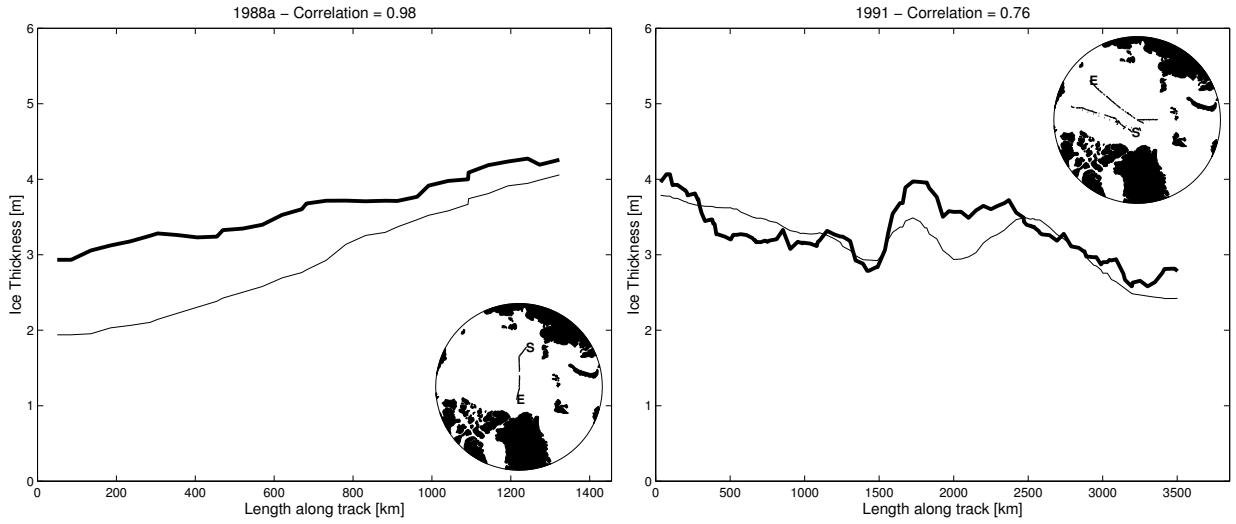
Parallel to the ensemble run a free-run model was also integrated for comparison. This model run used the standard ECMWF forcing fields with no pseudo–random forcing fields or $P^\star$ fields, and has an initial state equal to the initial ensemble mean. Some of the relevant parameters for the assimilation experiment are given in Table 3.4.

3.9.2 Synthetic ice thickness data

The ice thickness data used in this assimilation study, is intended to be a realistic representation of ice thickness products which will be available from the Cryosat mission of the European Space Agency (ESA). It is created using sea ice model data, which gives estimates of the ice thickness, ice concentration, surface temperature and snow thickness. The surface state is then used in a model which simulates the actual orbit of the Cryosat satellite, and the instrument and geophysical errors introduced by the sea ice thickness retrieval.

3.9.2.1 Modeled ice thickness data

The surface state was generated from a run of a coupled sea–ice/ocean model. The model consists of the Miami Isopycnic Coordinat Ocean Model (MICOM; *Bleck and Smith*, 1990) coupled to a dynamic–thermodynamic sea ice



(a) 1988 (b) 1991

Figure 3.36: The figures show comparisons between ice thickness from the model used to generate synthetic ice thickness (thin line), compared with ice thickness estimates from submarine sonars (thick line).

model. The sea ice model uses the Viscous–Plastic rheology of *Hibler* (1979), as implemented by *Harder* (1996). The thermodynamic module is described in *Drange and Bleck* (1996). The main differences between the model used to generate synthetic ice thickness data and the one used in the assimilation experiment later on is in the ocean model (MICOM/HYCOM) and the ice dynamics rheology (VP/EVP). It should also be noted that the numerical grids are slightly different, although they both include the Arctic and North Atlantic Oceans. The model forcing is also different, as the model used to generate synthetic ice thickness data utilizes NCEP forcing, while the assimilation model uses ECMWF forcing.

The model used to generate synthetic ice thickness was integrated for a 40–year period, from 1958 to 1998. Comparisons with ice thickness from submarine sonars have shown that the model appears to give a good description of the actual ice thickness. Two such comparisons are shown in figure 3.36. The synthetic ice thickness data used here were generated for the year 1990.

3.9.2.2 Generation of cryosat–like observations

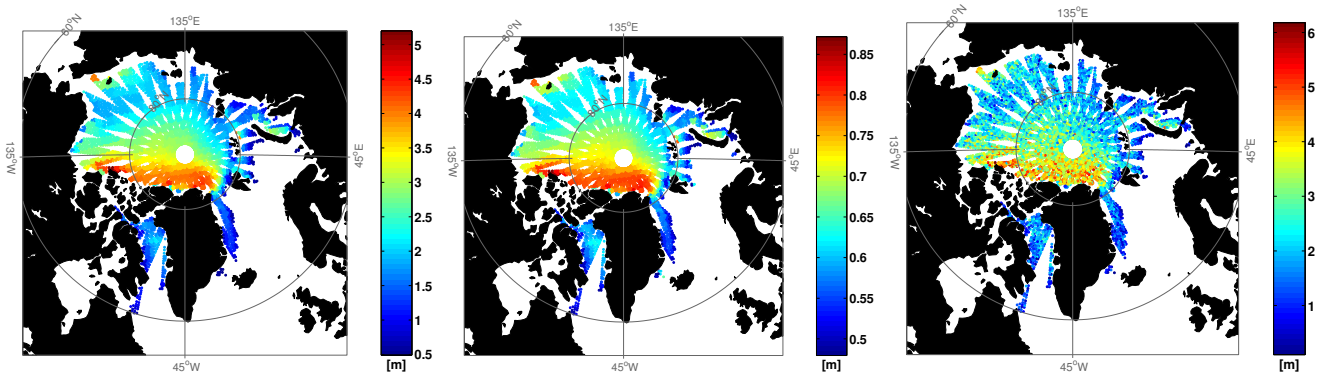
After having obtained the ice thickness estimates from the coupled ice/ocean model, errors were added to the thickness estimates to simulate the errors of the cryosat sensor. As the Cryosat satellite sensor measures the freeboard f, a first step is to convert model ice and snow thickness to freeboard height,

$$f = \frac{\rho_w - \rho_i}{\rho_w} h_i + \frac{\rho_s}{\rho_w} h_s \,, \tag{3.21}$$

where h and ρ are thickness and densities of the snow and ice (subscript i or s), and ρ_w is the density of seawater. The synthetic freeboard estimate is then calculated by adding a random error w_f to the model freeboard estimate:

$$f' = f + w_f \qquad\qquad \sigma_f^2 = \sigma_{Eice}^2 + \sigma_{Eocn}^2 \tag{3.22}$$

The random error w_f has error variance σ_f, which is a function of the ice elevation error variance σ_{Eice}^2 and the ocean elevation error variance σ_{Eocn}^2. The error variances of ice and ocean elevation depend on estimated formulas given in *Laxon* (2001). These errors are functions of the surface state, depending on ice concentration, ice thickness,



(a) Original ice thickness h_i (b) Ice thickness error variance $\sigma_{h_i'}$ (c) Synthetic observations $h_i + w_{h_i}$

Figure 3.37: The original ice thickness fields from the MICOM/VP run is shown in Figure 3.37(a). The estimated error variance of the synthetic observations, calculated from equation (3.23), is shown in Figure (3.37(b)). The final synthetic observations, calculated by adding a random error w_{h_i} to the original observations, is shown in Figure 3.37(c). The figures show the data generation for day 35 in 1990.

surface temperature, ice concentration and backscatter contrast between ocean and water. The ice concentration has the strongest impact on the error variance estimates, while the ice thickness and backscatter coefficients have a smaller impact on the error variance estimate.

The data generation also depends on the surface temperature. Experience from ERS indicate that as the surface temperature reaches the melting point, the location of the surface reflection becomes ambigous. If the surface temperature falls below $-5°\mathrm{C}$, the synthetic data gives a "null" value. Furthermore, due to the strong dependence of ice elevation error on the ice concentration, a ice concentration treshold has been introduced. If the model ice concentration is below 0.7, a "null" value is generated.

As the synthetic freeboard value is generated from equation (3.22), the synthetic ice thickness h_i' can be obtained from equation (3.21). Likewise, the ice thickness error variance of the synthetic data can be obtained by using equations (3.21) and equation (3.22):

$$\sigma^2_{h_i'} = \frac{\rho_w^2\left(\sigma^2_{Eocn} + \sigma^2_{Eice}\right) + \rho_s^2\sigma^2_{h_s}}{\left(\rho_w^2 - \rho_i^2\right)}\,, \tag{3.23}$$

where a snow thickness error variance has been included as well. An example of the original model ice thickness together with the synthetic ice thickness error variance and the final synthetic ice thickness is shown in figure (3.37).

3.9.3 Analysis updates

The analysis changes the ice concentration and ice thickness of the ice model, as well as fields in the ocean model. The individual updates depend on the statistics of the observation and model errors, as well as on the distance between observations and model. In this section we intend to show how the assimilation of ice thickness affects the properties of the ice cover and the properties of the ocean model.

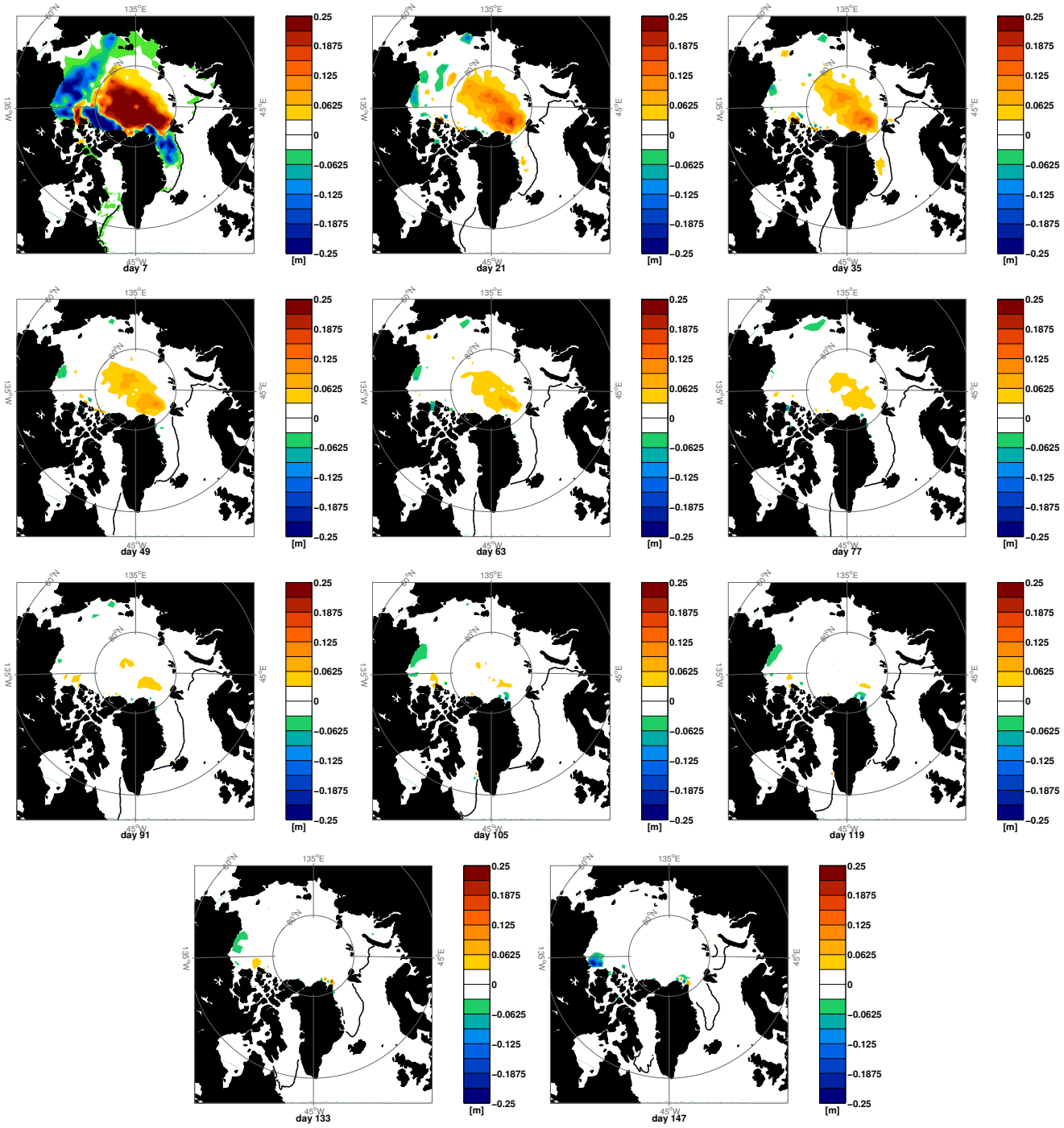


Figure 3.38.a: The ice thickness update for January–May in 1990. The day-of-year is indicated below each figure, where January 1st is day 0. The scales are identical for all figures. The black contour line denotes the 0.1 ice concentration contour.

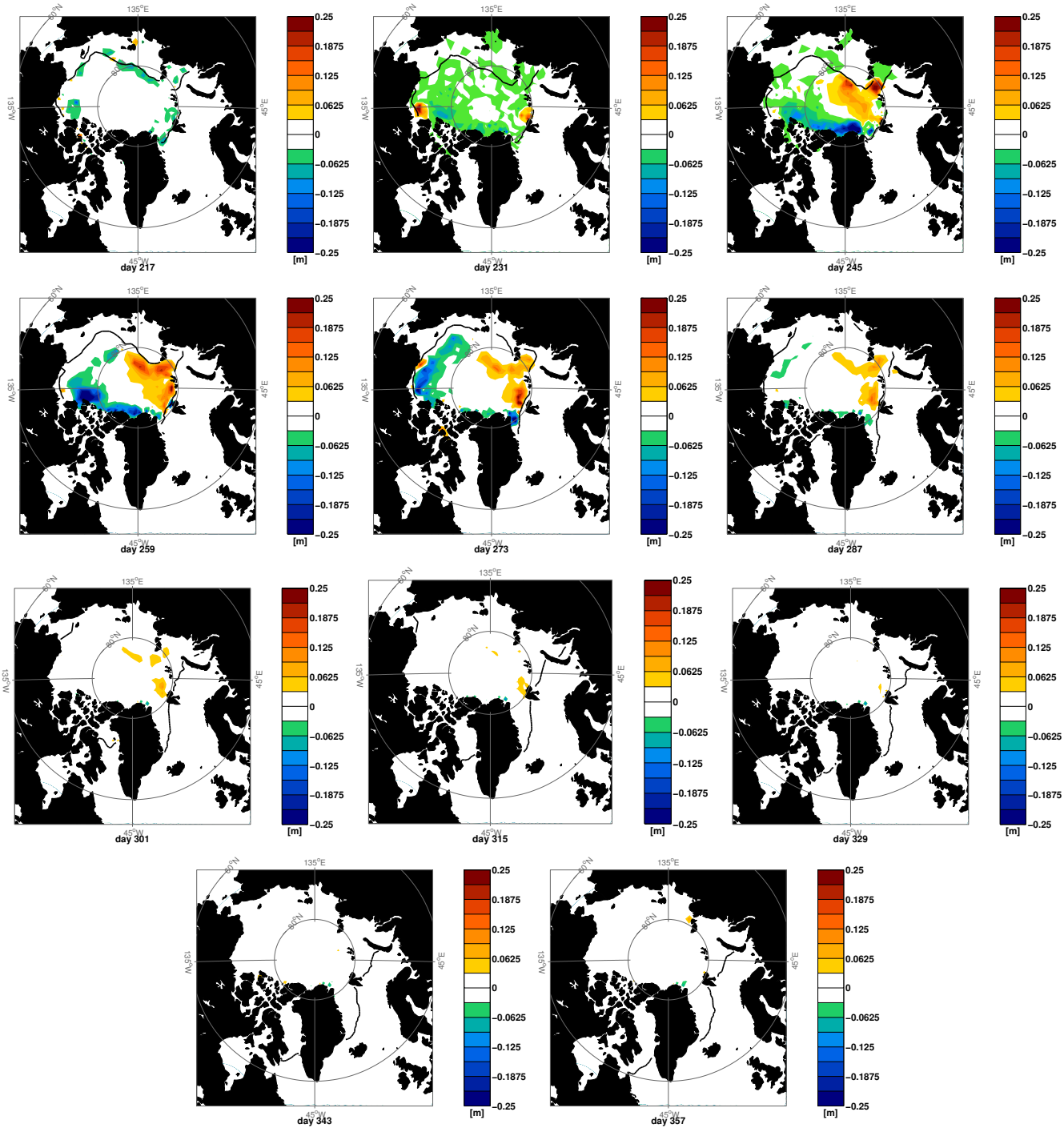
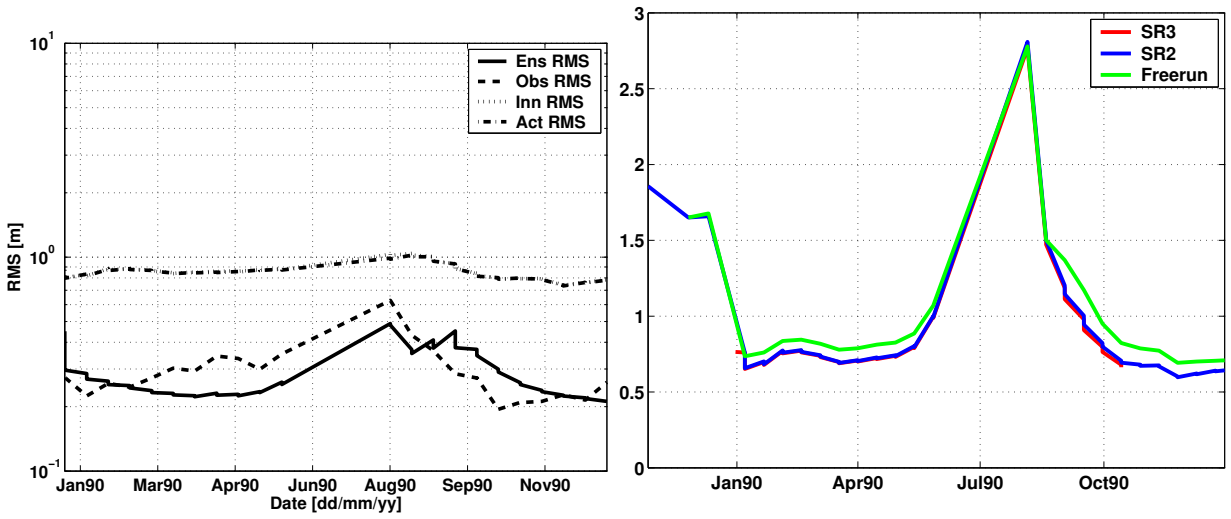


Figure 3.38.b: The ice thickness update for August–December in 1990. The day-of-year is indicated below each figure, where January 1st is day 0. The scales are identical for all figures. The black contour line denotes the 0.1 ice concentration contour.



(a) (b)

Figure 3.39: Figure 3.39(a): RMS values of the observation, model ensemble, innovation and actual error variance of experiment "SR2". The RMS-values are obtained only over regions where either the model or observation mean is greater than 0.1 m. Figure 3.39(b): The mean innovation for the three experiments.

only shows the ice thickness updates for the period January–May and August–December in 1990.

For the period January–May of 1990, we note that the updates are strongest at the beginning of the time period, especially at the first analysis time, day 7. The strong update is a combination of a large distance between observations and measurements, and high model error variance at this time. As the first analysis is performed, some of the distance between model and observations will be reduced, and the model error variance will also be reduced. Initially, there are some differences between the model fields and the synthetic observations. Roughly speaking, we can say that the model has too thin ice in the Central Arctic Ocean, whereas it generally has too thick ice in the surrounding seas; The Beaufort Sea, the East Siberian Sea, the Laptev Sea and the Greenland Sea, see day 7 in Figure 3.38.a.

As the model ensemble evolves, it is noticeable that the magnitude of the analysis updates steadily decreases. This is a result of decreased distance between observations and model, but mostly it is a result of reduced error variance in the ice model ensemble. This is clearly illustrated in Figure 3.39(a), where the RMS-values of the observation and ensemble variances are shown. The RMS-values of the observation variance has been calculated by using the variance of the observations within the radius of influence r_0 used in the EnKF. If the error statistics of the Ensemble Kalman Filter are correct for all grid points, then we should have similar values for the innovation RMS and the sum of RMS values for the ensemble and observations. From Figure 3.39(a), this is seen not to be the case, as the innovation RMS, especially in winter, is much higher than the sum of model and observation RMS-values. Also, the ensemble variance is seen to be somewhat lower than the observation variance in winter, giving increased faith in the model results in the EnKF scheme. It should also be noted that the largest updates take place close to the North Pole where the data density is highest.

The conclusions of experiment "SR3" are also similar to the conclusions of ecperiment "SR2". There is an increased ensemble spread in the "SR3" experiment relative to experiment "SR2", due to the random fields used in the ice dynamics. The increased ensemble spread is not enough, however to give any major differences in the conclusions reached for experiment "SR2", and so we will focus mainly on experiment "SR2" in the following. Note that there is some, albeit small improvement in the experiments with assimilation, relative to the experiment without assimilation, experiment "SRF", as can be seen in Figure 3.39.

The ice edges shown as the black contour line in Figure 3.38.a are seen to extend well into the Greenland Sea and Labrador Seas. Compared to SSMI-observations, these extents seem to excessive, and is probably connected to too low vertical mixing by the KPP scheme in the ocean model. It should be noted that the model data on which the synthetic observations are based, do not display such an excessive ice edge in the Labrador and Greenland Seas.

As summer sets in, the temperature of the ice/snow surface is too high to obtain reliabe ice freeboard measurements. As a consequence, the ensemble spread will increase throughout summer, as there are no restriction imposed by the assimilations. The assimilations recommence on day 217, and all the analysis updates for the period August–December 1990 are seen in Figure 3.38.b.

Beginning day 217, the first analysises are performed, and the result can be seen close to the ice edge marked by the 0.1 ice concentration contour. It is somewhat surprising that these updates occur so close to the ice edge, as the majority of the observations is expected to be in the interior of the ice pack at this time, due to the expected lower ice surface temperatures there. The reason is due to high ice thickness variance close to the ice edge, increasing the model uncertainty in the analysis. This high ice thickness variance is, incidentally, connected to ice concentration variance along the ice edge, which will be demonstrated in the next section.

At the beginning of the time period August–December, the analysis updates generally reduce the ice thickness, but after a while the updates seem to take on a ”dipole” shape, with ice thickness decrease along the Greenland Sea, Arctic Archipelago Coast and Beaufort Sea. Ice thickness increase can be seen in the Eurasian Basin and in regions close to Svalbard, Franz Josefs Land and Sevarnaya Zemlya. Such a structure could also be seen in the first analysis performed in January,

The reason for this dipolar structure is probably related to various differences between the model generating the synthetic observations, and the model used here. Differences exist in the thermodynamic, dynamic and atmospheric forcing applied to these model, and also in the amount of North Atlantic Water entering the Arctic Basin. We have not performed a breakdown of the exact nature of these differences.

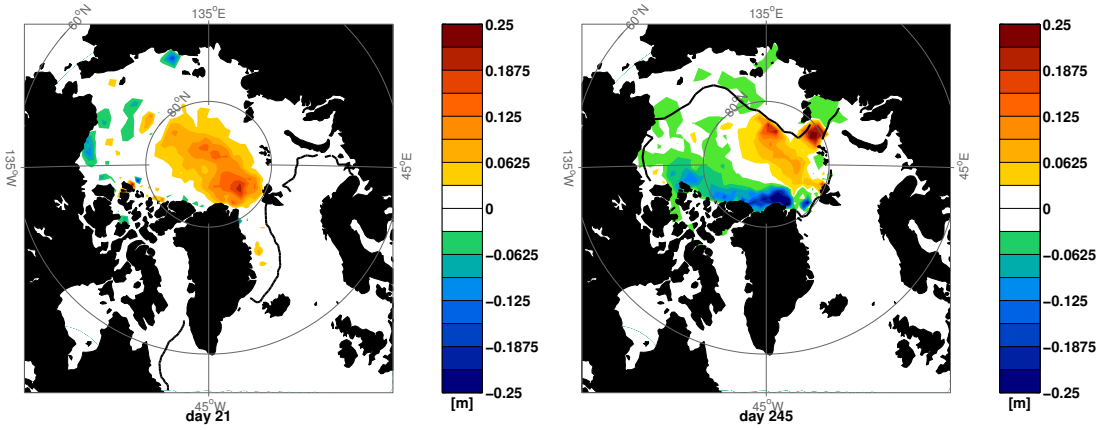
As the model evolves through the autumn of 1990, we again note that the updates are strongest at the beginning of the time period, while they steadily reduce towards the end of the time period. Again, this is the result of the reduced model–observation innovations and reduced model variance, leading to greater faith in model results in the EnKF scheme used here. As in winter, the effect of reduced model ensemble variance dominates, which can clearly be seen from Figure 3.39(a).

3.9.3.2 Salinity and ice concentration updates

The analysis updates of surface layer salinity and ice concentration are shown together with the ice thickness update in Figure 3.40 for two different times of the year, Winter (day 21) and Late Summer (day 245), for experiment SR2. These two times are at the beginning of the Winter and Autumn assimilation periods respectively, when the ice thickness variance is at its highest, see Figure 3.39(a). Consequently, the analysis updates are at their largest at this time for ice thickness. It turns out that such is the case for ice concentration and surface layer salinity as well. In the following we will use the word ”significant” to denote analysis updates which are of high enough absolute value to be colored in Figure 3.40. This subjective judgement should not be confused with statistical significance in this context.

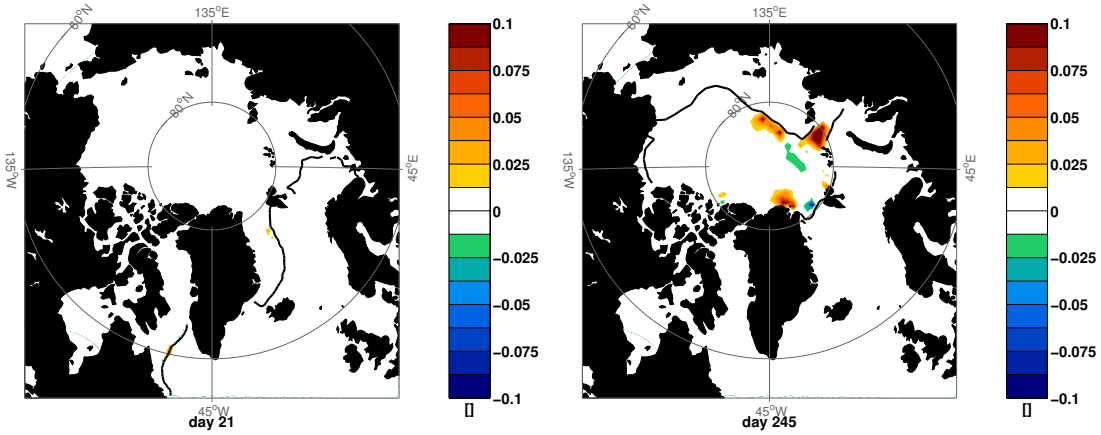
If we consider the 22nd January salinity, ice concentration and ice thickness update in som detail, we find differences in spatial distribution, and some interesting details. First of all, for the winter situation, we see that compared to the ice thickness update, the ice concentration update has a relatively small spatial extent, with some minor updates close to the ice edge. As was seen in section 3.8.3.1 , the ice concentration variance within the ice pack is close to zero in winter, resulting in minor updates there.

Examining the wintertime salinity update within the $80^\circ N$ latitude circle, we see that it is of a different sign compared to the ice thickness update, revealing that increased ice thickness is not necessarily corresponding to increased salinity.



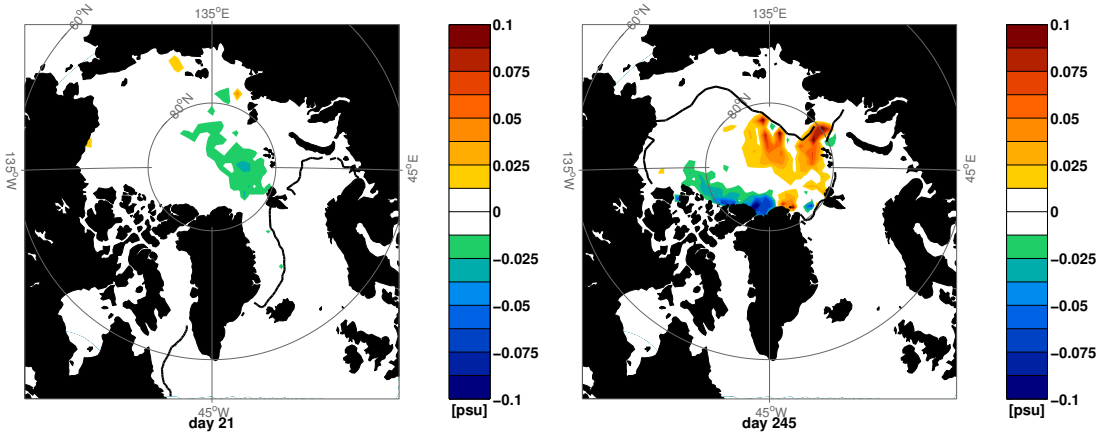
(a) Ice Thickness Update day 21

(d) Ice Thickness Update day 245



(b) Ice Concentration Update day 21

(e) Ice Concentration Update day 245



(c) Salinity Update day 21

(c) Salinity Update day 245

Figure 3.40: Selected model updates of Salinity, ice thickness and ice concentration from two times, day 21 and day 245 in 1990, for experiment ”SR2” . The thick line denotes the 0.1 ice concentration contour.

A lot of this is due to ice production over open water, which tend to reduce the overall ice thickness (by increasing ice concentration), but increases the total ice production. The amount of salt released into the surface layer is positively correlated with the total ice production. If we consider the spatial coverage of the salinity updates, we see that it is somewhat smaller than the region with significant ice thickness updates. This is probably connected to different ensemble error statistics for the region with thick ridged ice close to the Northern coast of Greenland, versus the generally thinner ice in the central parts of the Arctic Basin.

On day 245, there is a larger region with significant ice concentration updates relative to day 21. The majority of the ice concentration updates are located close to the ice edge, and we see that some of the regions with highest magnitude of ice thickness updates also correspond to regions with highest magnitude of ice concentration updates. The sign of the ice thickness updates do not show such a direct correspondence, however. For instance, the strong ice thickness update east of Zevarnaya Zemlya has the same sign as the ice concentration update there. North of Greenland, however, we see that a region of high ice concentration increase corresponds to a region with strong ice thickness decrease.

Complex underlying statistics on day 245 can also be inferred from the differences between ice thickness and surface layer salinity updates in Figure 3.40. The sign of the salinity updates are seen to follow the updates of ice concentration close to the ice edge. This behaviour was also seen in Section 3.8.3.2 , where summertime ice concentration was seen to be strongly linked to the summertime surface layer salinity through a positive correlation. In the updates of sea surface salinity, we also see a region with strong salinity reduction close to the Northern Coast of Greenland, in the same region with strong ice thickness reduction.

3.9.4 Ensemble Covariances

The second order statistics of the model ensemble are important for the analysis, and they are described by the matrix $\boldsymbol{P}^f$, present in the analysis equation (3.3). In this section we are going to describe the characteristics of the ensemble covariances, and how it varies in space and time. We will also compare the statistics of the EnKF scheme with that of an Ensemble OI scheme.

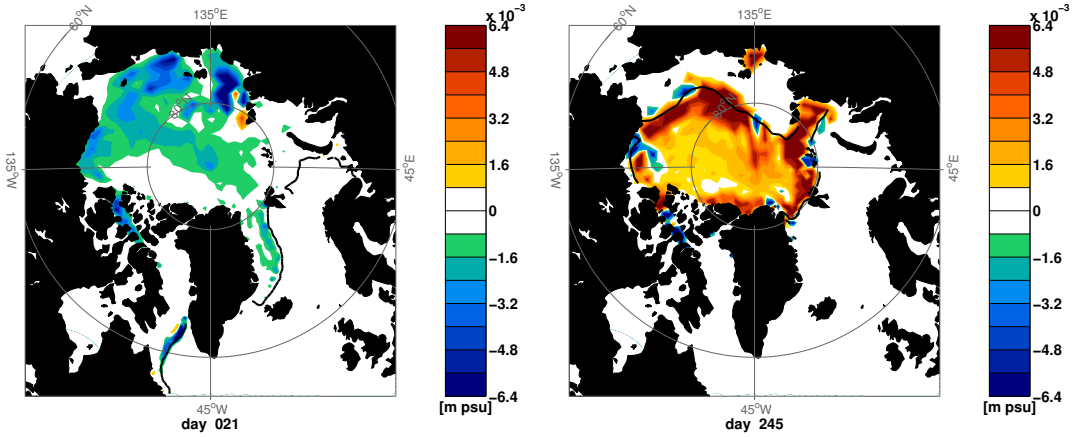
3.9.4.1 Covariance fields

We here focus on the covariance between different variables in the same grid cell, e.g. between ice thickness and the salinity in the first layer of the ocean model. We will write this type of covariance as $\mathrm{cov}\left\{c(\boldsymbol{x}), S(\boldsymbol{x}, 1)\right\}$. In Figure 3.41 we show the covariance field between ice thickness and ice concentration, ocean surface layer salinity and ocean surface layer temperature at two times of the year. These two times corresponf to the two different analysis times given in Figure 3.40.

The region with high magnitudes of ice thickness–salinity covariance is seen to include both the regions along the ice edge, as well as within the ice pack. This is in contrast to the covariance between ice thickness and ice concentration and the covariance between ice thickness and sea surface temperature, which have their highest values along the ice edge. There is a marked difference between the salinity–ice thickness covariance on day 21 and that of day 245, Figure 3.41(a) and 3.41(d) respectively . In winter, we generally see a negative covariance between ice thickness and sea surface salinity, whereas in summer, the covariance is mostly positive. This behaviour is connected to the different mechanisms which create the covariance in the model.

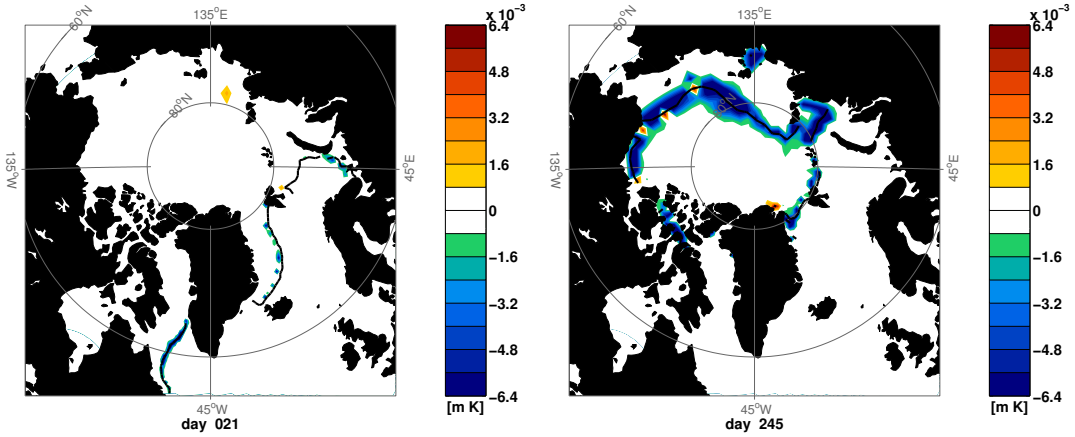
In winter, open leads in the ice model is an important contributor to the total mass of ice created in the model. Neglecting the vertical mixing induced by brine released during freezing of ice, we can say that increased ice production will lead to increased salinity in the ocean surface layer. In winter, the heat loss over open water can be one to two orders of magnitude larger than the heat loss over ice covered water, and this means that the more open water is created, the more ice will be produced. When open water is created during ice divergence, there is a net transport of thick ice out

cc



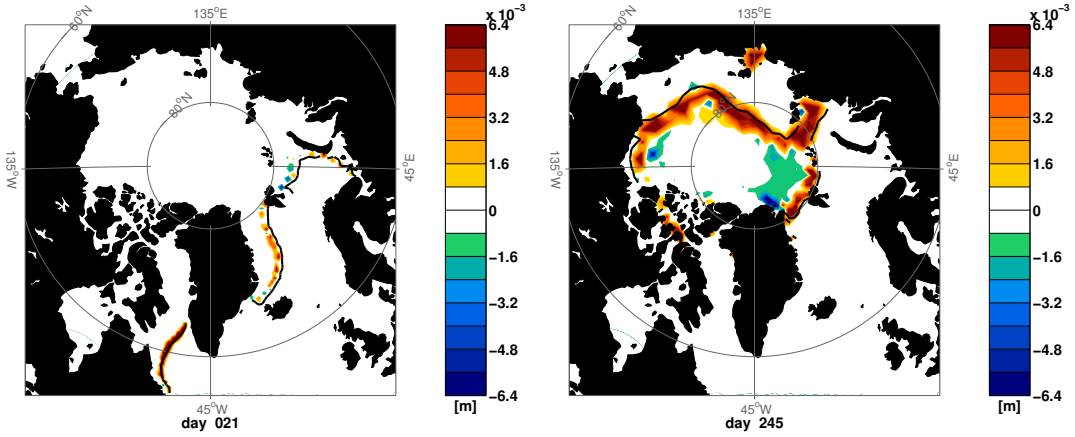
(a) cov $\{h_i(\boldsymbol{x}), S(\boldsymbol{x}, 1)\}$, day 21

(d) cov $\{h_i(\boldsymbol{x}), S(\boldsymbol{x}, 1)\}$, day 245



(b) cov $\{h_i(\boldsymbol{x}), T(\boldsymbol{x}, 1)\}$, day 21

(e) cov $\{h_i(\boldsymbol{x}), T(\boldsymbol{x}, 1)\}$, day 245



(c) cov $\{h_i(\boldsymbol{x}), c(\boldsymbol{x})\}$, day 21

(f) cov $\{h_i(\boldsymbol{x}), c(\boldsymbol{x})\}$, day 245

Figure 3.41: Covariance fields of ice thickness versus surface layer salinity, surface layer temperature and ice concentration at day 21 ans day 245 in 1990. The thick line denotes the 0.1 ice concentration contour of the model forecast.

of the area, which is replaced by the thin ice created in leads. This means that the average thickness of ice in the area will be reduced while the ice production and amount of salt released to the ocean surface layer will increase. The amount of ice divergence is different between different ensemble members, and the result of this process will be a negative ice thickness–surface layer salinity covariance in the ensemble. In accordance with this mechanism, regions with net convergence should not see large magnitudes of ice thickness–surface salinity covariance. This is probably the reason for the low magnitudes of the covariance in the region North of the Canadian Archipelago and the north coast of Greenland.

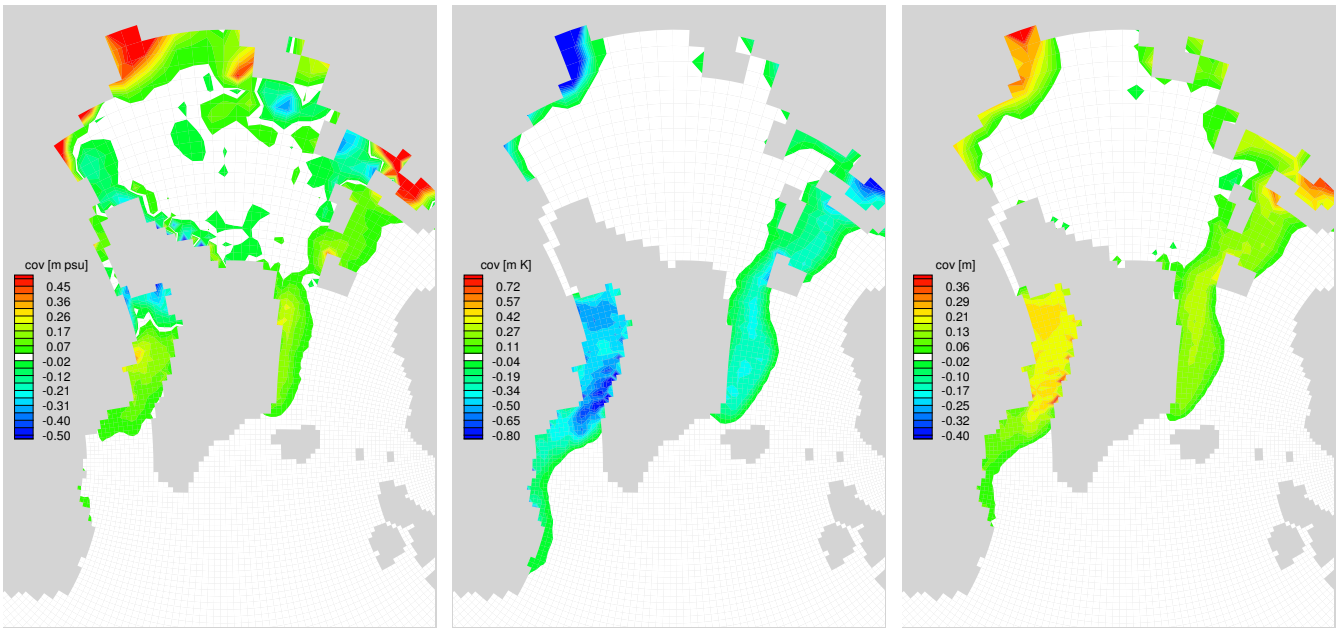
It should be noted that the above mentioned mechanism relies on the assumption of no vertical mixing in the ocean model. If we include the vertical mixing in the argumentation, the results will be different. For the central Arctic Ocean, however, the presence of the Cold Halocline Layer (CHL) will not alter the main conclusion of a negative covariance. Since the underlying waters in the CHL have higher salinity than the surface waters and are at the freezing point, vertical mixing will still result in a negative covariance in the ensemble. For other regions in the Arctic, such behaviour might not be the case. In Figure 3.41(a), positive covariances are seen close to Davis Strait and Sevarnaya Zemlya in the Laptev Sea. Also, the advection of ensemble anomalies by the mean current of the ensemble is an important effect, which complicates the above mechanism. It should also be noted that the ice transport mechanism commented in Section 3.8.3.2 is important in creating the negative ice thickness–surface salinity covariance found along the ice edge. This effect was seen to dominate in winter, along the ice edge.

For the summer situation, Figure 3.41(d), the contrast between heat fluxes over ocean and ice is lower than in winter. The differences are still noticeable, though, due to the higher albedo of ice relative to water. The ice model used here has a parametrization of horizontal ice melt. In contrast to the horizontal freezing of ice in leads during winter, the horizontal melt of does not modify the ice thickness. Therefore the vertical melt of ice is of greater importance in summer. When ice melts vertically, the ice thickness is reduced, and the salinity of the surface layer is reduced, leading to a positive ice thickness–surface salinity covariance in the ensemble. The major random forcing components for the summer situation is the atmosphere temperature and the variance of the atmospheric surface drag. The temperature component of the random forcing leads to different vertical melt rates, while the surface drag leads to differences in ice divergence in the ensemble. The ice divergence leads to heating of the mixed layer which in turn melts the ice from below. Both these effects contribute to the positive ice thickness–surface salinity covariance.

The high magnitudes of ice thickness–surface temperature variance is seen to be mainly located close to the ice edge for the entire period, Figures 3.41(b) and 3.41(e). The model does not allow ice to exist in a grid cell if the surface temperature to be higher than the freezing point when ice is present, meaning that the surface temperature within the ice edge is only dependent upon the surface salinity. The only region with high thickness–surface temperature magnitudes within the ice pack during winter is located in the Laptev Sea, a region which also has high ice thickness–surface salinity covariance, Figure 3.41(a). The majority of the covariance is therefore located close to the ice edge, both in summer and winter. The covariance with surface temperature is mainly negative throughout the year close to the ice edge. This is connected to differences in sea surface temperature when ice is present or not.

Within the ice pack in winter, the ice concentration in the model tends to reach the highest permitted value, c_{max}, for all ensemble members, thus giving a zero covariance within the ice pack. This can be seen in the covariance field for ice concentration–ice thickness in Figure 3.41(c), where there is zero covariance within the ice pack. In summer, Figure 3.41(f), the ice concentration is often below c_{max}, and covariance between ice thickness and ice concentration can develop. An important mechanism here is lateral freezing of ice at the beginning of fall, which tends to reduce the ice thickness while increasing the ice concentration. Different degrees of lateral ice freezing will lead to negative ice thickness–ice concentration variance. At day 245, this can be seen in a region north of Greenland. This is an important effect which should be present in winter as well. Unfortunately, the tendency of the ensemble members to all reach the maximum allowed ice concentration prevents this covariance from developing in winter.

The ice thickness–ice concentration covariance, as the sea surface covariance fields, has the highest magnitudes close to the ice edge, Figures 3.41(c) and 3.41(f). If we compare the ice thickness–ice concentration covariance fields with the ice thickness–surface temperature covariance, they have similar spatial patterns, leading us to the conclusion that



(a) cov $\{h_i(\boldsymbol{x}), S(\boldsymbol{x}, 1)\}$ (b) cov $\{h_i(\boldsymbol{x}), T(\boldsymbol{x}, 1)\}$ (c) cov $\{h_i(\boldsymbol{x}), c(\boldsymbol{x})\}$

Figure 3.42: Covariances from an EnOI covariance matrix. Panel 3.42(a) shows the covariance between ice thickness and ocean upper ocean layer salinity, panel 3.42(b) shows the covariances between ice thickness and upper ocean layer temperature and panel 3.42(c) shows covariances between ice thickness and ice concentration.

this covariance field is a result of ice being present or not in the ensemble. If ice is present the ice thickness and ice concentrations are both positive, while they are both zero when ice is absent, leading to a positive covariance.

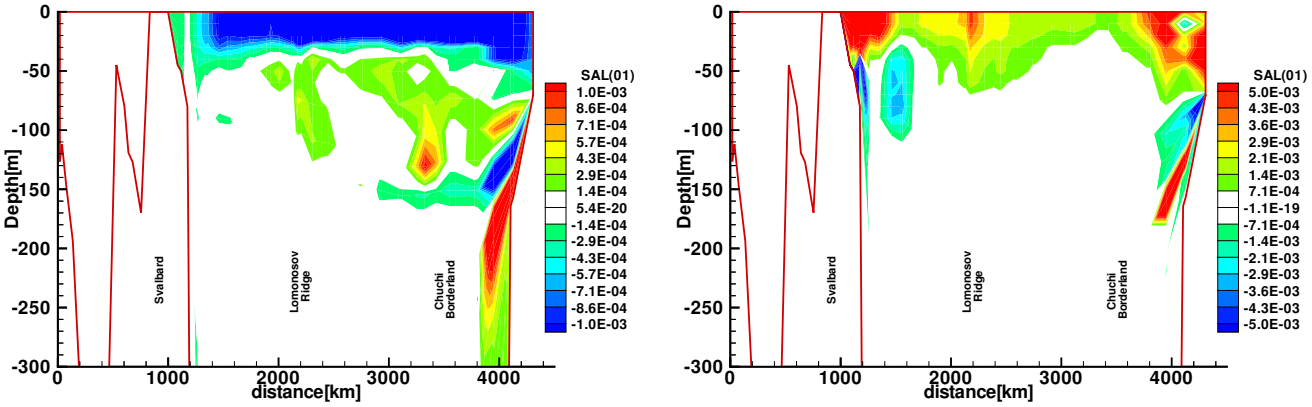
Comparison with EnOI statistics In a previous ESA project(ESA ESTEC Contract No. 13071/00/NL/DC) (*Lisæter*, 2001), several assimilation experiments were performed with a less resource–demanding assimilation scheme. The EnOI has a fixed covariance matrix used in the analyses throughout the year which is very different from whate we have seen regarding the EnKF covariances. It is interesting to see how the covariance fields of that experiment compares with the current experiment. The EnOI covariances of ice thickness against salinity, temperature and ice concentration are shown in Figure 3.42.

The EnOI covariance matrix was estimated from model samples. These samples were collected from a model run over one season. This is reflected in the ensemble statistics, as the seasonally changing position of the ice edge is seen as a continous region with relatively high magnitudes of covariances in both ice concentration, salinity and temperature. The EnOI ensemble can not capture the effects of a moving ice edge, which we have seen is an important element in the the EnKF covariances. In addition, the changing sign of the ensemble covariances over a season is not captured, this effect is probably of highest importance for the salinity update.

3.9.4.2 Covariance in depth

So far, we have concentrated on the analysis updates, and the covariances found at the ocean surface. The reason for this is the generally higher response at the surface than in deeper layers of the ocean, but a short demonstration of the covariance with depth is also called for. In Figure 3.43, the covariances between ice thickness and ocean salinity in deeper layers are demonstrated for a geodesic section from (20°E,70°N) in Norway, to (160°W,70°N) in Alaska. Two different times of the year are shown, to illustrate the different covariances in winter and summer conditions.

In winter, represented here by day 21 in Figure 3.43(a), the top layers of the ocean model have a negative covariance



(a) $\mathrm{cov}\,\{h_i(\boldsymbol{x}), S(\boldsymbol{x}, z)\}$, day 21

(b) $\mathrm{cov}\,\{h_i(\boldsymbol{x}), S(\boldsymbol{x}, z)\}$, day 245

Figure 3.43: Covariance between salinity and ice thickness for a section along the geodesic line from (20°E,70°N) in Norway, to (160°W,70°N) in Alaska. The covariance fields are shown for two different days, day 21, Figure 3.43(a) and day 245, Figure 3.43(b), in 1990.

between ice thickness and salinity. As mentioned earlier this is related to the ice production in open water leads, and different degrees of ice divergence in the ensemble. A detail in Figure 3.43(a) is the change of sign of the covariance with depth at many locations in the Central Arctic. As brine is released during ice production, turbulence will erode the stratification at the base of the oceanic mixed layer. In the presence of a Cold Halocline Layer, entrainment of saline water will increase the overall salinity of the mixed layer, but the lower part of the mixed layer will actually experience a decrease in salinity, compared to the initial stratified case. These different characteristics of the mixed layer are probably the reason for the change from negative to positive covariances with depth, for some of the regions along the section.

In summer, melting dominates, so the entrainment process as described above is less frequent. The sign of the covariance is positive close to the surface, Figure 3.43(a), and the sign does not change with depth in the central Arctic, e.g. the Lomonosov Ridge. Note however that close to the ice edge, such as in the proximity of Svalbard, there is a change in sign from positive to negative with depth.

Finally we note the region at around 400km distance into the section, close to the coast of Alaska. Here the covariance sign is seen to change sign several times when we go downward through the water coloumn, both in winter and summer. The exact mechanism for this is unclear, but we speculate that it is connected to processes related to inflow of water from the Bering Strait. The region is also close to the Chuckchi Borderlands, which also point to influences from shelf waters and waters of North Atlantic origin.

3.9.5 Ice thickness assimilation – summary

The ice thickness assimilation has been demonstrated, and shown to have an impact on both the ocean surface, and on the ice variables. The updates are, however, relatively small due to the properties of the ensemble.

The ice thickness variance of the ensemble is rapidly reduced by the ensemble updates, and not sufficiently increased by the random forcing applied between analysis steps. This could be seen by the large distance between the observations and the observed model state, relative to the ensemble and observation variances. This large distance indicates that there is significant bias in the model, an effect which cannot be properly adressed by the EnKF. Another exper-

iment, in which we varied properties of the ice dynamics showed some improvement, but not enough to overcome the majority of the bias. In the future, focus should be put on proper description of such biases in ice assimilation schemes. Although the impact of the assimilations are relatively small, there is a cumulative effect of them, and at the end of the assimilation experiment, the ensemble mean is seen to be closer to observations relative to a model run without assimilation.

In a previous work, a different assimilation scheme was used to assimilate ice thickness into a coupled ice–ocean model. This scheme, The Ensemble Optimal Interpolation, uses a fixed ensemble to represent the error statistics. In several experiments, it was demonstrated that the assimilation had a significant effect on the ensemble. As the error statistics of the EnOI are fixed, it is not subject to the problems the EnKF faces. We believe that when the biases of the EnKF is properly adressed in the EnKF scheme, we will see significant improvements in the assimilation updates.

One area where the EnKF is superior to the EnOI scheme, is in its description of the error covariances. Since the EnOI uses fixed covariances, it is not able to properly represent seasonally changing error statistics. As was demonstrated in the experiments performed here, the EnKF describes a seasonally changing error covariance matrix, with proper representation of covariances due to different physical mechanisms. We saw for instance that the covariance between surface salinity and ice thickness was very different for a situation in winter, relative to a situation at the end of summer. It was also seen that the ice edge region was a region were the covariances showed high magnitudes, while the interior ice pack had lower magnitudes. The proper representation of this is also an important attribuite of the EnKF, relative to the EnOI.

We conclude that the ice thickness assimilation gives physically consistent model updates, and that these updates are a large improvement over the updates from the much simpler EnOI scheme. The magnitude of the updates are small due to significant model bias, and to low EnKF error variances. Future work should focus on both describing this bias in the EnKF, and improving the model to reduce the bias. The CryoSat measurements, as described by the synthetic ice thickness measurements used here, should be of sufficient precision to improve model results when model bias is properly accounted for.

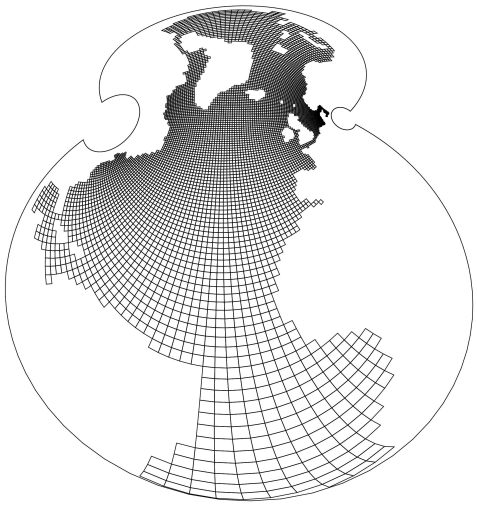


Figure 3.44: The orthogonal curvilinear 140×130 grid. Note the high resolution in the North Atlantic and that there is no singularity in the Arctic ocean.

3.10 Task 3300: Marine Ecosystem

This task presents an application of the ensemble Kalman filter for a marine biochemical model, utilizing ocean colour data from the SeaWiFS sensor. A hindcast experiment for April and May 1998 is summarized, i.e., including the early part of the North Atlantic spring bloom. We start by giving a short overview of the model setup and the data, then the results are presented and discussed, and finally a summary is given. For more details, please also refer to *Natvik and Evensen* (2003a,b).

3.10.1 Model setup

The model system consists basically of three sub-models. The physics of the ocean are described by an ocean circulation model - Miami Isopycnic Coordinate Ocean Model (MICOM), and the output fields are used to force the FDM-type biogeochemical model by *Drange* (1994a) describing a simple marine ecosystem. A dynamic and thermodynamic ice model is also included in the Arctic ocean. A general overview of the model system was given in the introduction of WP 3000. Note that this experiment uses a model system consisting of MICOM and FDM, and the coupling to HYCOM was done at a later stage, i.e., representing an improvement of the system used in the hindcast experiment.

In the model setup used here, we have 17 physical and 18 biochemical vertical layers, respectively, with two biochemical sublayers in the physical mixed layer. Even though a higher resolution of the biochemical layers would be desired for pure model studies or for operational systems, we decided to keep the coupled model as simple as possible in our assimilation “demonstration” experiments. This was partly also motivated by the amount of computer CPU-time, memory and disk space required for ensemble integrations and analyses of the coupled model system. The coupled model system was set up for the North Atlantic using the orthogonal curvilinear 140×130 grid shown in Figure 3.44.

3.10.2 The SeaWiFS data

The NASA operated Sea-viewing Wide Field-of-view Sensor (SeaWiFS), which is located on board the Orbview-2 satellite, is the second generation ocean colour sensor which has provided the observations used in the data assimilation experiments described below. The data have been processed at the Joint Research Centre (JRC); they were received from JRC as a global level 3 product of chlorophyll a binned (averaged) over ten days on a longitude/latitude grid with a resolution of $0.25°$.

It is necessary to perform an atmospheric correction of the data received at the sensor, since scattering from the atmosphere contributes with about 90% of the total signal (*Austin*, 1993). Originally, one main goal of the SeaWiFS project was to be able to estimate surface chlorophyll a within 35% accuracy in Case 1 waters (i.e., optically dominated by phytoplankton) over the range of $0.05 - 50.0$ mg m^{-3} (*Hooker et al.*, 1992). However, note that the NASA atmospheric correction algorithm breaks down in Case 2 waters, leading to higher concentrations of chlorophyll a. There is currently no doubt that even better accuracies than 35% can be obtained in Case 1 waters. In fact, *Keiner and Brown* (1999) showed that an artificial neural network algorithm (NN) was able to produce more accurate estimates of chlorophyll in oceanic waters (within $\sim$ 20%) than the operational SeaWiFS ocean chlorophyll 2 (OC2) algorithm. Similar improvements were reported by *Gross et al.* (2000), who also used an NN approach. Since the data we received from JRC did not include any error estimates, we assumed standard deviations of 35% in the main experiment, although other percentages were also tested. Assuming 35% errors seems reasonable, since the measurements were processed using a recalibrated version of the OC2 algorithm, and a similar algorithm as NASA for the atmospheric correction (*Bulgarelli and Melin*, 2000). However, the following comments should be noted: First, since the atmospheric correction scheme breaks down in Case 2 waters, the data are probably given too much weight in coastal zones, near rivers or in areas where human activity is important. Second, we use an empirical expression to convert the chlorophyll data to phytoplankton biomass in terms of nitrogen (see below), which obviously leads to larger errors in the converted observations. Thus, specifying 35% for the total errors is probably an underestimation for the current data set. However, near future observations will probably be more accurate, e.g., by using an NN algorithm similar to *Keiner and Brown* (1999). See also the sensitivity experiments. Although a better description of the data errors would be required for an operational system, the 35% hypothesis should be of sufficient quality for our "demonstration" data assimilation experiments.

To assimilate the chlorophyll data into our biochemical model, we have to be able to relate them to phytoplankton biomass in nitrogen units. For this purpose, we use a simple empirical (invertible) expression, an alternative would be to include chlorophyll as a separate explicit variable in the biochemical model.

3.10.3 Data assimilation experiments

A data assimilation experiment was performed for the months April and May in 1998, i.e. including the early part of the North Atlantic spring bloom. This section contains the results from the experiments, starting with a short description of the setup.

In our implementation of the ensemble Kalman filter analysis scheme, the analysis is calculated grid point by grid point (horizontally). Further, we define an influence radius for the observations, i.e., the update in each grid point is based on data within a certain distance. It can be expected that observations at long distance will only give a minor contribution to the analysis in some grid point, and excluding the contribution from far away data makes the analysis more effective. The influence radius is set to 100 km, representing the size of a typical mesoscale eddy. Within the influence radius, the correlation between the errors of any two data is exponentially related as $\exp(-r/r_0)$, where r is the distance between them and r_0 is a characteristic decorrelation scale set to 10 km.

After the biochemical model variables have been updated using the ensemble Kalman filter analysis, possible negative fields are simply set to zero. The species may become alive again during model integration.

	Number of ensemble members	Observation errors (std. dev.)
Main experiment (day 90-147)	100	35%
Sensitivity wrt. ensemble size (analysis at day 96)	10	35%
	20	35%
	40	35%
	60	35%
	80	35%
Sensitivity wrt. obs. errors (analysis at day 96)	100	1%
	100	10%
	100	60%
	100	300%

Table 3.5: Description of the data assimilation experiments.

In a main experiment, data were assimilated at six times; day 96, 106, 116, 126, 136 and 147 (see Table 3.5 and 3.6). Additionally, the sensitivity of the analysis with respect to the ensemble size and the observation errors was studied at the first assimilation time (i.e., at day 96). Note that the relatively short duration of the experiment does not allow for an investigation of long term trends of the ensemble statistics. The choice of performing a time limited experiment was motivated by the quite extensive numerical cost of integrating the ensemble (100 members) of 3-dimensional physical/biological states forward in time; a 10 day integration in parallel took about one day on 20 processors on a Cray Origin 2000 machine and the storage of the ensemble required approximately 4.5 gigabytes for each set of restart files. The effect of the assimilation on long term trends of the ensemble statistics should be investigated at a later stage, and the work presented here should only be regarded as an initial demonstration study. A short overview of the experiments is given in Table 3.5.

Although the sensitivity tests were performed at the first filter analysis time, we start by presenting the main experiment, which was performed with an ensemble size of 100 members and observation standard deviations of 35% (see Table 3.5). *Evensen* (1994) indicated that 100-500 ensemble members are sufficient to represent the error covariance matrices. We chose to use 100 members in the main experiment; this should be enough in our local (grid point by grid point) implementation of the analysis scheme and at the same time the numerical requirements do not become too large.

Even though the experiment was carried out for the domain shown in Figure 3.44, we focus on the results from the North Atlantic domain from $10^{\circ}\mathrm{S}$ to $70^{\circ}\mathrm{N}$. Further north there are few ocean colour data to update the estimate (because of the low inclination of the satellite/sensor), while the domain close to the South Atlantic boundary is less important in our study.

Table 3.6 shows some important characteristics of the experiment. The first column contains the times where the ensemble Kalman filter analysis is carried out, while the second column shows the time binning period of the chlorophyll data. In the third column, the percentage of bad data (mainly due to cloud cover and sun glitter) is given for wet points. A data point is regarded as wet only if all four surrounding model grid points have depths above 10 m. The fourth column shows the number of grid points which are skipped (i.e., keeping the forecast value) during the analysis. A point is skipped if it is considered as dry, i.e. whenever the depth is less than 10 m, or if there are no observations within the influence radius to update the estimate. There are $140 \times 130 = 18200$ grid points altogether, so the variables are influenced by the data in a majority of grid points (well above 90%). The last column shows that some grid points are influenced by many observations. (Our results are of course a direct consequence of the choice for the data influence radius). We will now focus on the results from the first and last assimilation time, i.e. at day 96 and 147.

The impact of the EnKF analysis on the observed variable (surface phytoplankton) at day 96 and 147 can be studied in Figure 3.45; the ocean colour data (in terms of $\mathrm{mmol\ N\ m^{-3}}$) binned over the days 91–100 and 141–151 are

Time of data assimilation	Time binning period of the chlorophyll data	Amount of bad data	Number of skipped grid points	Max number of data within the radius of influence
1998 day 96	day 91–100	35.3%	992	117
1998 day 106	day 101–110	35.9%	886	195
1998 day 116	day 111–120	36.4%	850	155
1998 day 126	day 121–130	38.3%	729	156
1998 day 136	day 131–140	39.3%	725	150
1998 day 147	day 141–151	40.2%	598	158

Table 3.6: Some characteristics of the main data assimilation experiment. Note that bad data are counted at wet points only, i.e., where the depths of the surrounding grid points are greater than 10 m. A particular grid point is skipped (not updated) if it is dry ($\mathrm{depth} < 10\ \mathrm{m}$) or if there are no observations available within the radius of influence.

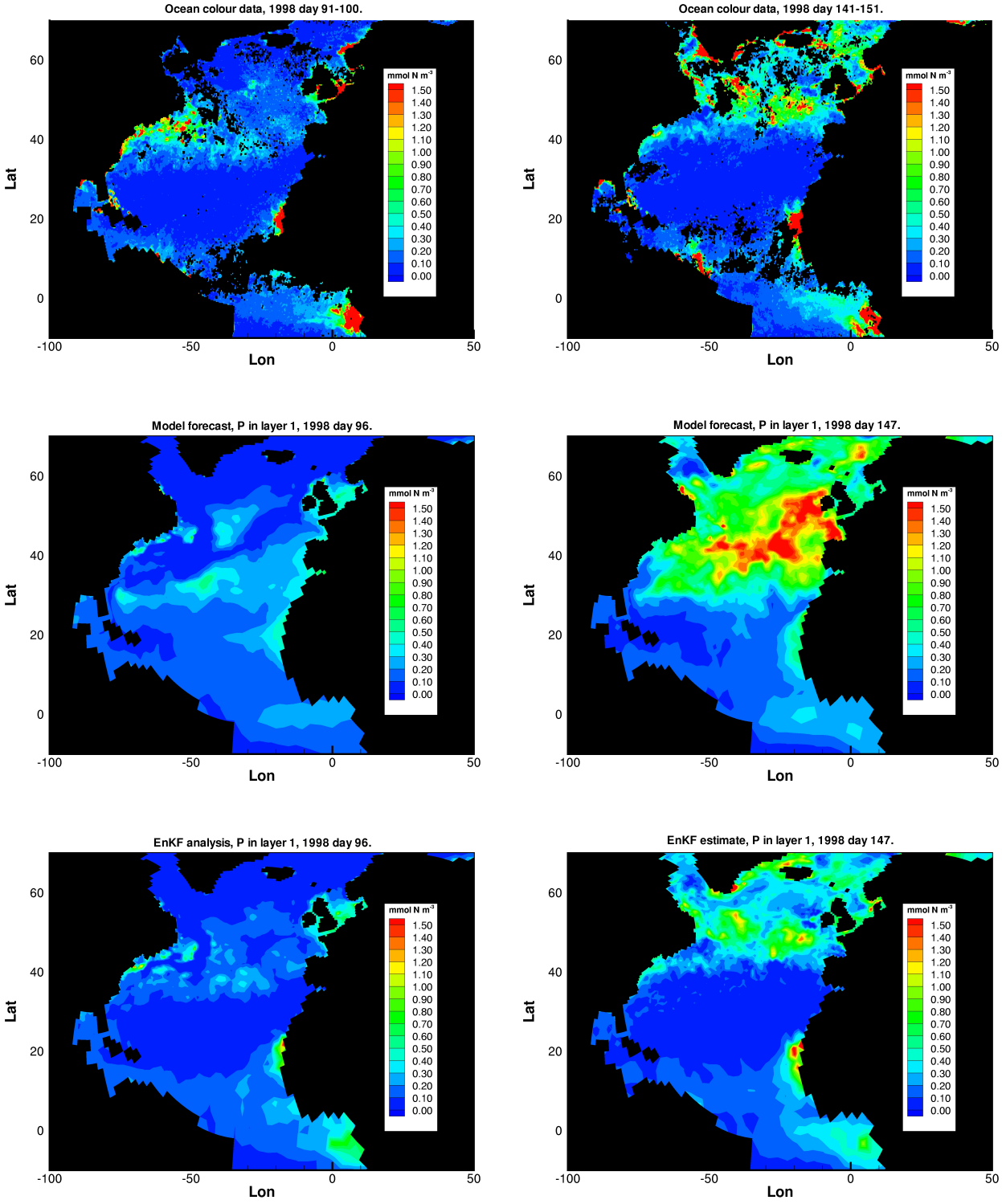


Figure 3.45: Results from the assimilation at day 96 (left) and day 147 (right). The observations, i.e. the ocean colour data binned over the days 91–100 and 141–151, after the conversion to phytoplankton biomass in terms of nitrogen, are shown in the top plots. Further, the ensemble mean of phytoplankton for the forecast and analysis ensembles are displayed in the middle and bottom plots, respectively.

displayed in the top figures, while the forecast and analyzed ensemble mean of surface phytoplankton are shown in the middle and bottom plots, respectively.

Even though the model produces phytoplankton blooms in several domains at day 96, the forecast estimate is quite different from the spatial distribution viewed by the SeaWiFS sensor. While the model seems to develop a large bloom extending from the eastern boundary of the North Atlantic basin, the observations show high concentrations at the western boundary. This undesired property of the model may have many causes. For example, we have only two biochemical layers in the physical mixed layer, which may be insufficient for resolving the ecosystem properly. Also, since the ecosystem is very sensitive to the supply of nutrients, it will also be very sensitive to the amount of turbulent mixing.

Figure 3.45 shows that the EnKF is capable of making successful updates during an assimilation, even for large differences between the model forecast and the observations. For example, by following the 0.10 mmol N m^{-3} contour line starting at the west coast of Portugal, it is seen that the analyzed estimate is quite close to the observations along this contour. Further, the positions of the phytoplankton blooms have certainly become more consistent with the data after the analysis, e.g., the two blooms in the ensemble Kalman filter estimate at the West African coast, the large concentrations near the western boundary of the North Atlantic basin, and the removal of the bloom that covers a large area of the eastern part of the basin in the model forecast. Similar improvements can be observed in the day 147 assimilation; the large bloom at mid latitudes seen in the model forecast is completely removed in the analysis, and the spatial distribution of phytoplankton has become consistent with the data.

Data-forecast and data-analysis residuals (interpolated onto the data grid) are shown in Figure 3.46 (top). Although the large negative residuals in the forecast are removed in the analysis mean estimate (i.e., consistent with the right plots of Figure 3.45), it is evident that the numerical peak values of the analysis are much lower than the corresponding values of the observations. Note that we have assumed observation standard deviations of 35% of the actual data values. Thus, the data errors are assumed to be large within a phytoplankton bloom (the observed maximum at day 147 is about 70 mmol N m^{-3}), i.e. the lower peak values in the analysis are consistent with our specifications of the error statistics. See also the sensitivity experiments below.

Figure 3.46 (bottom) shows the ensemble variance of phytoplankton for the forecast and for the analysis at day 147. It is seen that the variance is significantly reduced during the analysis. A long term experiment (i.e., a year or more) would be important to investigate whether the reduction of the variance also would result in a reduction of the long term variance in the model.

The ensemble Kalman filter is a multivariate data assimilation methodology. Thus, all the ecosystem components are updated during an analysis. The quality of the analyzed variables may be studied in so-called twin experiments, where observations are generated by adding noise to a "reference solution", which should be closely recovered by the analyzed estimates. The ensemble Kalman filter has already been investigated in twin experiments by *Breuillin et al.* (2000) and *Eknes and Evensen* (2001), who used 1-dimensional extensions of the *Fasham et al.* (1990) and *Evans and Parslow* (1985) ecosystem models, respectively. The results from these studies indicate that the information contained in a set of surface observations of phytoplankton is sufficient to provide successful updates also for other model variables. In our setup, we use real satellite ocean colour measurements, and a "reference solution" is of course not available. However, some desired properties of the multivariate analysis can still be investigated. Figure 3.47 shows the results at day 147 for nitrate (top) and zooplankton (bottom) for the upper 300 m of the water column along a section at $50^{\circ}\mathrm{N}$. The ensemble mean forecast estimates are shown in the left plots, while the residuals between the analysis and forecast ensemble mean estimates are shown in the right plots, respectively. As discussed earlier (in relation to Figure 3.45), the model predicted a too large surface phytoplankton bloom at mid latidudes. However, this large bloom was removed in the analysis, i.e., to become consistent with the data. In Figure 3.47 (top, right), it can be seen that there are large positive residuals at the surface from about $30^{\circ}\mathrm{W}$ to about $10^{\circ}\mathrm{W}$ between the analysis estimate and the forecast estimate of nitrate. Thus, while the too large forecast bloom of phytoplankton has become much smaller in the analysis, larger concentrations of surface nitrate can be seen in the analysis for the same domain. This seems reasonable, since phytoplankton is growing on nitrate and since a negative correction of phytoplankton therefore should lead to a positive correction of nitrate in the multivariate analysis. Further, large negative residuals between the analysis and forecast mean estimates of zooplankton for the same domain can be seen in Figure 3.47 (bottom, right). Again, this seems reasonable, since zooplankton is feeding on phytoplankton and a

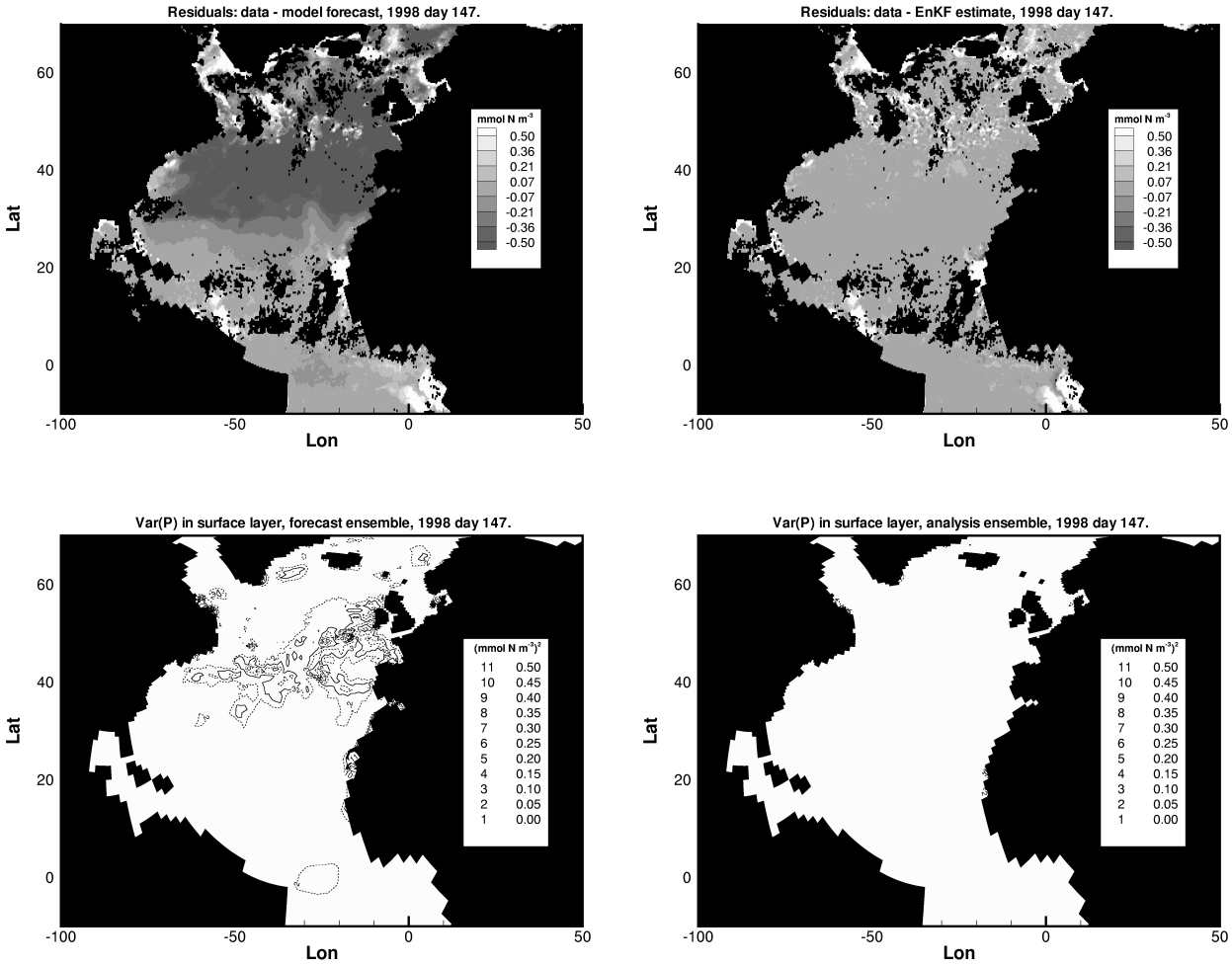


Figure 3.46: Residuals between the ocean colour observations and the ensemble mean estimate interpolated onto the data grid (day 147) for the forecast ensemble (top, left) and the analysis ensemble (top, right); and the variance of phytoplankton at day 147 for the forecast ensemble (bottom, left) and for the analysis ensemble (bottom, right).

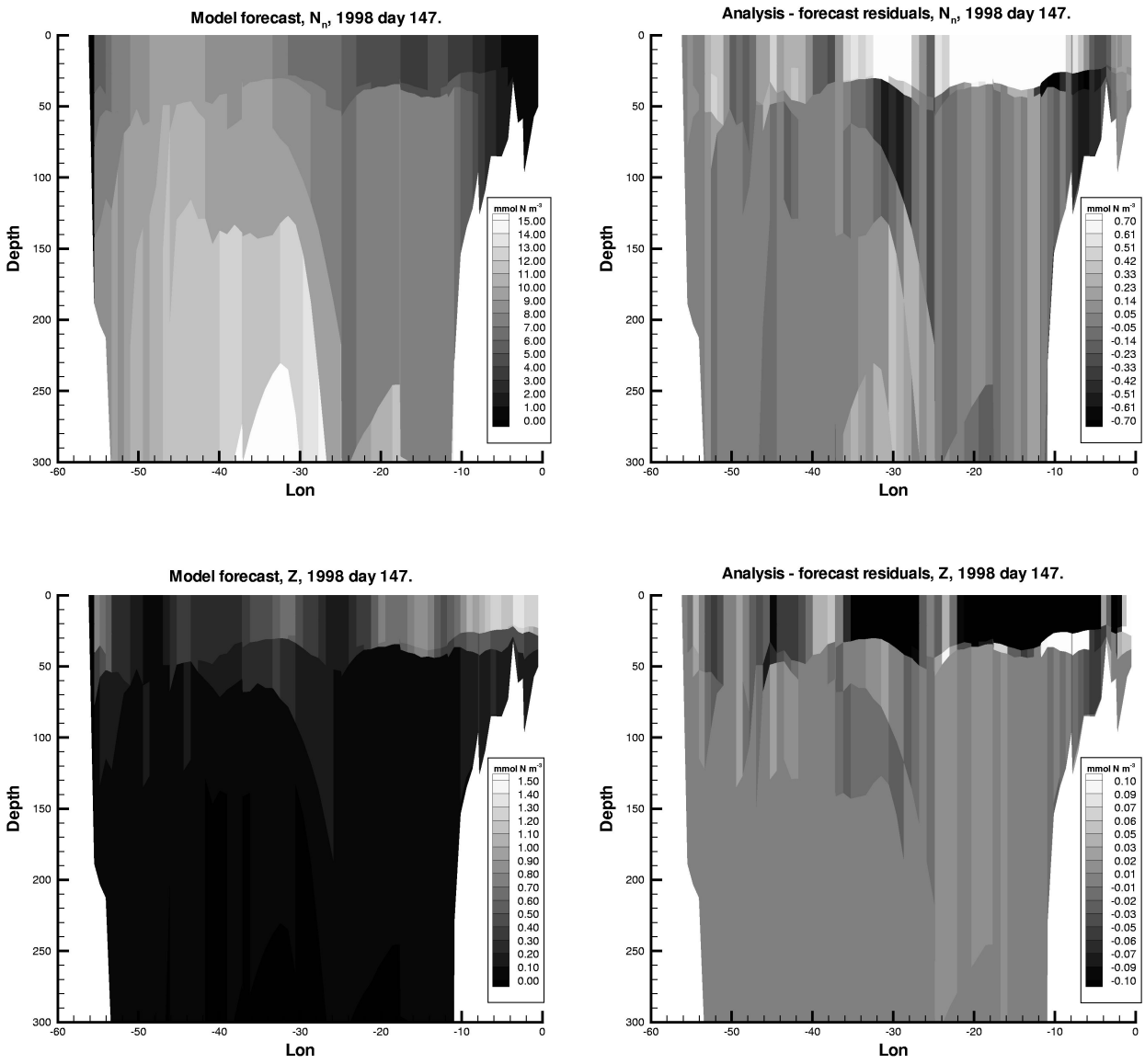


Figure 3.47: Depth profiles of the upper 300 m of the water column along a section at $50^{\circ}\mathrm{N}$ (day 147). The left plots show the forecast (ensemble) mean estimate of nitrate (top) and zooplankton (bottom), while the right plots show the residuals between the analysis (ensemble) mean estimate and forecast (ensemble) mean estimate for the same variables.

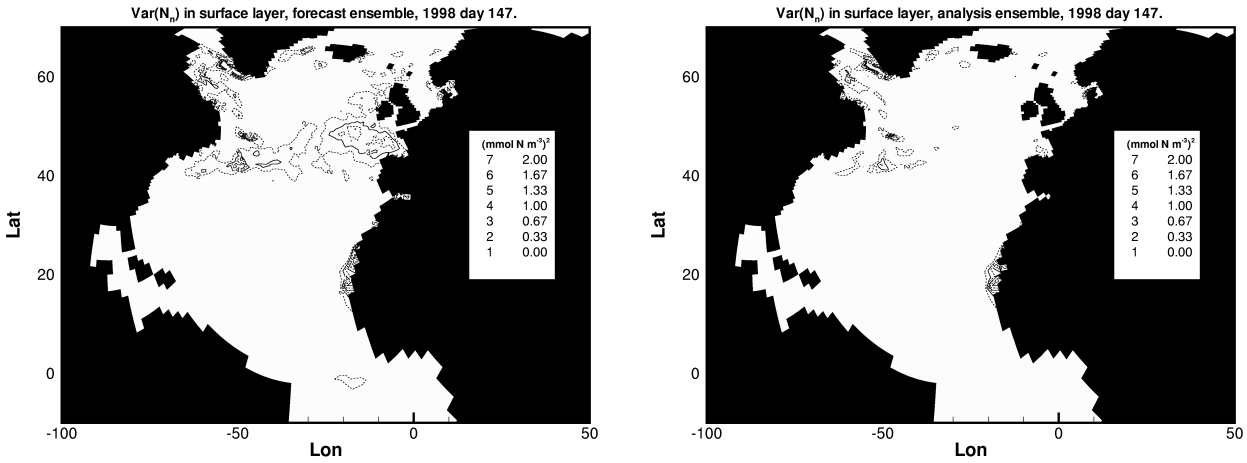


Figure 3.48: Variance of surface nitrate of the forecast ensemble (left) and the analysis ensemble (right) for day 147.

negative correction of phytoplankton therefore also should lead to a negative correction of zooplankton in the multivariate analysis. Note from Figure 3.47 that the analysis also affects sub-surface zooplankton and nitrate, although the analysis updates are weaker in the deep ocean (i.e., as expected, since we assimilate surface observations). To conclude, the multivariate assimilation seems to have a positive impact (qualitatively) on the coupled biochemical state.

Figure 3.48 shows the variance of surface nitrate at day 147 without assimilation (left) and with assimilation (right). It is clearly seen that the variance of nitrate is much smaller for the analyzed ensemble. As expected, this property was also seen for the other biochemical components.

The sensitivity experiments were carried out for the first analysis step, i.e. at day 96, basically to determine the setup for the main experiment. As said above, standard deviations of 35% were assumed for the observations in the main experiment. Other percentages were also used in the first analysis at day 96, as presented below.

A necessary requirement of the analysis would be the capability of almost matching the data for very small observation errors and almost matching the model forecast for very large errors, respectively. Figure 3.49 (top, left) shows the ensemble mean estimate of surface phytoplankton after an analysis using very accurate data; standard deviations of only 1% were specified. It is seen that the analyzed estimate is very close to the data (Figure 3.45; top, left), both the positions and the intensities of the phytoplankton blooms almost perfectly match those described by the observations. The analyzed mean estimate of surface phytoplankton in the case of very large measurement errors (300%) is shown in Figure 3.49 (bottom, right). As desired and expected, the estimate is close to the model forecast shown in Figure 3.45 (middle, left).

Figure 3.49 also shows the concentrations of surface phytoplankton assuming observation standard deviations of 10% and 60% in the analysis, respectively. As expected, the estimate in the former case is also quite close to the observations, both in terms of spatial distribution and bloom intensity. The main differences from the case with 1% data errors appear in the western North Atlantic basin, where several phytoplankton blooms are seen from the data. Note that the use of 60% standard deviations for the observations results in an analyzed estimate very close to the standard case with 35% errors displayed in Figure 3.45 (bottom, left). As expected, the peak values observed in the data are more in agreement with the latter analysis. However, for lower concentrations the spatial distributions of phytoplankton are very similar for the two cases.

We also looked at the other biochemical fields for the different cases described above (not shown here). As expected, an assumption of accurate observations led to quite large updates of the other variables, while larger errors in the

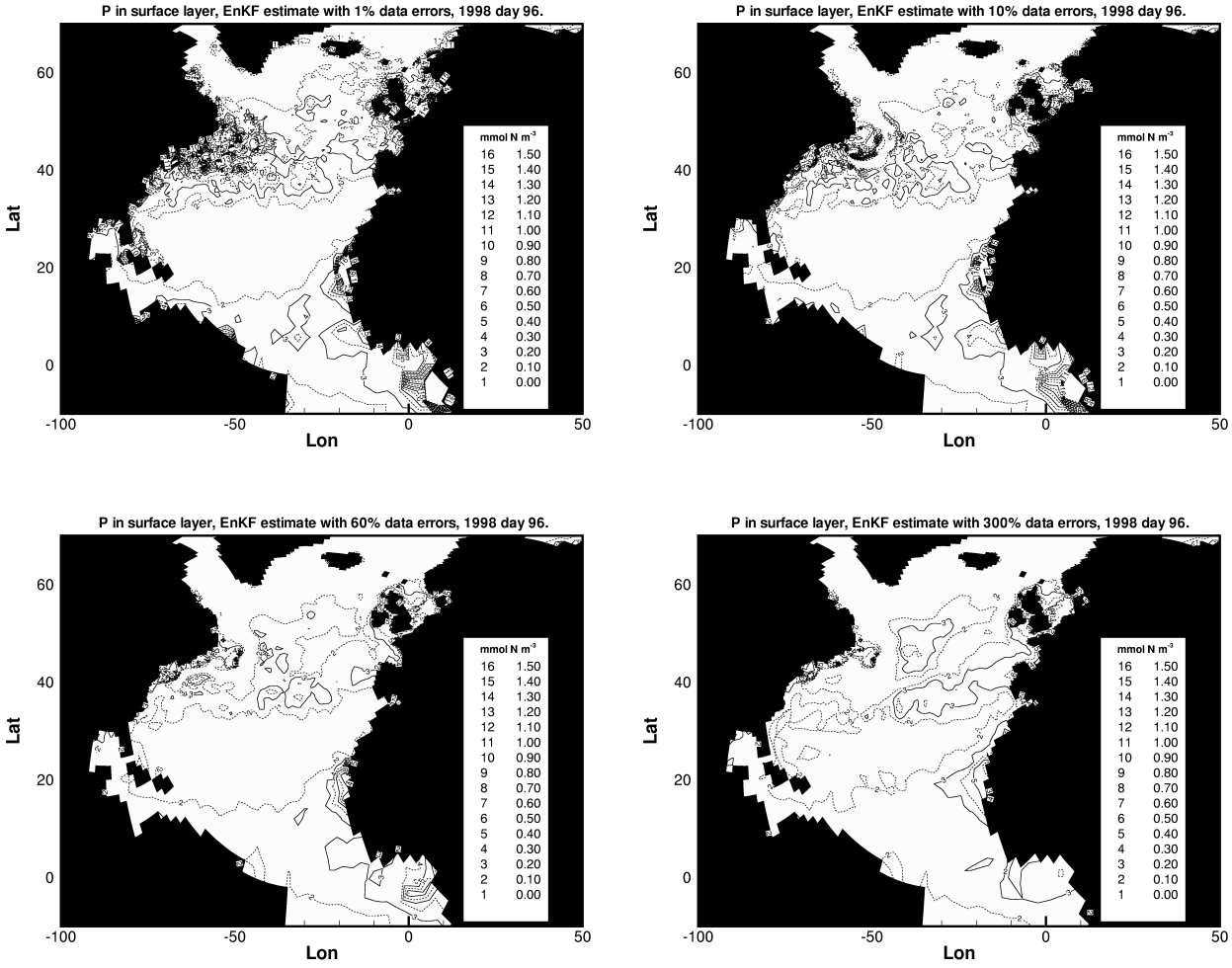


Figure 3.49: Ensemble mean of phytoplankton after an EnKF analysis with data errors of 1% (top, left), 10% (top, right), 60% (bottom, left) and 300% (bottom, right), respectively.

data resulted in weaker updates during the analysis. Further, more accurate observations led to lower values in the analyzed variance fields (not shown). This is also expected, since the updates of the members due to accurate data should bring the members closer together, i.e. the ensemble should experience a high degree of convergence.

3.10.4 Marine Ecosystem Summary

A data assimilation system has been developed for the biochemical part of a coupled physical-biochemical 3-dimensional model of the North Atlantic. The data assimilation technique is the ensemble Kalman filter, which has been implemented for the first time with a realistic coupled model, and the system utilizes real chlorophyll data from the SeaWiFS ocean colour sensor.

The results from the experiments were presented above. A first observation was that the model system as implemented here is not capable of producing a spatial distribution of phytoplankton consistent with the data. At day 96, the model estimated a bloom of phytoplankton covering a large domain of the eastern part of the North Atlantic, while the observations showed the largest concentrations near the western boundary of the basin. After the ensemble Kalman filter analysis, the phytoplankton field became consistent with the data, as a result of assuming unknown errors in the model. To the contrary, note that a pure parameter estimation problem relies on the model to be able to reproduce the data (within specified error bounds) for some set of parameters.

We studied the variance of each ecosystem variable during the assimilation. As desired, our results showed significant reductions in the variance fields for the different ecosystem components (only the phytoplankton and nitrate variance fields were shown here). Note that the short duration of our experiment does not allow for investigating any long term trends in the model. Normally (but not necessarily), one will experience a spread of the ensemble during model integration. Thus, it would be interesting to see if the ensemble Kalman filter is able to reduce the long term variance in the model. This should be investigated at a later stage.

In a twin experiment setup, where the measurements are generated from a "reference" solution, it is possible to investigate the quality of all ecosystem components, even if only (surface) phytoplankton is observed. The ensemble Kalman filter has already been investigated in twin experiments with simple 1-dimensional ecosystem models in *Breuillin et al.* (2000) and *Eknes and Evensen* (2001). The results from these works indicate that it is sufficient to observe surface phytoplankton to get reliable estimates of the other model components. In this paper, we use real measurements, and a reference solution is of course not available. However, some qualitative properties of the multivariate assimilation can still be investigated. A first observation was that the model predicted a too large bloom of phytoplankton at mid latitudes at day 147. This bloom was removed in the analysis, and the distribution of phytoplankton became consistent with the data. Further, the multivariate analysis led to a significant increase of surface nitrate in the same domain. This seems reasonable, since phytoplankton grows on nitrate and a reduction of phytoplankton therefore should lead to increased concentrations of nitrate in the multivariate assimilation. Further, zooplankton concentrations decreased significantly in the same domain in the analysis. Again, this seems reasonable, since zooplankton feed on phytoplankton and a decrease of the latter therefore also should lead to a decrease of the former in the multivariate analysis. We also verified that the analysis affected sub-surface nitrate and zooplankton concentrations, although the analysis updates are weaker in the deep ocean (as expected). Of course, our qualitative investigation should be followed up by a validation of the system against independent observations. Note that in the present setup of the system, the model errors were sampled by adding Gaussian noise to the atmospheric forcing fields. One important improvement would be to add noise to the biochemical model equations also, to sample the true errors of the system more correctly. Thus, validation against independent observations will be done at a later stage, when we have improved the sampling of the model noise.

We investigated the sensitivity with respect to the number of members in the ensemble (not shown here). Since the ensemble members are updated grid point by grid point (horizontally) in the analysis, it is important to have enough members to resolve the covariances locally within the radius of influence with respect to each point. From the sensitivity experiments, we conclude that at least 60-80 members are required to have reasonably well resolved

covariances. However, an ensemble of about 150 members would be desired for any operational application. The required number of ensemble members is related to the degrees of freedom of the system. Thus, the same size order can be expected for similar coupled physical-biochemical model systems.

The sensitivity with respect to the errors specified for the observations was also studied. It was seen that very accurate data led to an analyzed estimate of phytoplankton very close to the observations, while large errors in the data resulted in a corresponding estimate close to the model forecast, as desired. Since we defined the observation standard deviations as a certain percentage of the actual data value (35% in the standard case), the analyzed estimate showed much lower peak values of phytoplankton than those seen in the data. Thus, the response of the analysis is consistent with the specification of observation errors. A very positive result of the investigation was that even with 60% standard deviations, the analyzed estimate provides an important improvement of the poor model prediction. For the other model components, accurate data led to strong updates during the assimilation (and small variance fields), while large errors in the data resulted in weak updates (and larger variance fields), as expected.

One improvement of the model system used in the current experiments include a replacement of the MICOM model with HYCOM (HYbrid Coordinate Ocean Model) - see the introduction of WP 3000. The HYCOM model uses ordinary z-coordinates in the mixed layer (the desired number of layers and vertical grid spacing can be specified), and isopycnal layers below. This is an improvement for the ecosystem, since no biochemical sublayer splitting is needed in the physical mixed layer. Another important improvement of the system will be to seek a better way of relating the chlorophyll a data to the model. Currently, the data are simply converted to phytoplankton biomass using an empirical algorithm, and the converted data are assimilated as “phytoplankton measurements”.

3.11 WP3000 Summary

This workpackage has examined the impact of several satellite observation types in a state of the art ocean and ecosystem modelling and data assimilation system. The ocean model is the MICOM/HYCOM as described in Section 3.1, the assimilation sheme is the EnKF as described in Section 3.3, and the overall system used is named TOPAZ and is briefly described in Section 3.2.

The major conclusions from this workpackage are summarized in the following sections.

3.11.1 Assimilation of SST data

The assimilation of satellite observed SST data provides an important source of information which helps correct for errors and/or biases in the atmospheric heat fluxes used in the model. The major impact is in the upper mixed layer of the model. The data used in the current experiments consisted of very coarse data (Reynolds SST of one degree resolution). Thus, only the large scales could be impacted. We do expect that the new high resolution SST data being developed in different projects (i.e. GHRSST) will provide information which is useful to control mesoscale circulation in and coastal regimes as well.

3.11.2 Assimilation of SLA data

The assimilation of SLA data is complicated since the data are relative to a reference sea surface height, itself depending on the geoid which is not known accurately at scales below 500 km. Therefore we may expect a bias in the assimilation of SLA data depending on the choice of a reference sea surface height. In SIREOC, this reference sea surface height was constructed from observations and it has turned out to work well.

The conclusion from these experiments is that the SLA data contain excellent information about the mesoscale variability in the ocean. The assimilation of SLA data allows us to introduce and track eddies and rings very well. They also have an impact on the thermocline depth which is partly determining the upper ocean mesoscale variability. In addition there is a weak impact on the mixed layer temperature which also impacts the sea level height.

3.11.3 Assimilation of SSS data

The assimilation of SSS or brightness temperature T_{B} data has similar value as the SST assimilation and has been examined using simulated data. It introduces a correction to the errors from erroneous freshwater fluxes in the model. The large scale data as expected from SMOS will be very valuable since they together with SST data will ensure that the upper ocean water-mass characteristics will be correctly represented in the model. This will most likely completely remove any need for additional surface flux relaxation which is used in many ocean models today.

Note that we have proven the capability of assimilate the T_{B} data directly, without first deriving SSS values. This is possible by using a modelled relation between T_{B} and other variables where some are modelled variables, e.g. SST and SSS, and other are input parameters to the model, e.g. atmospheric wind speed. Thus, it is possible to compute the model predicted T_{B} and relate this to the observed T_{B}, and subsequently to compute the difference. The computation of the cross correlations, used in the EnKF analysis, is based on an ensemble of model states.

3.11.4 Assimilation of ice concentration data

The assimilation of ice concentration data is essential to ensure a correct ice extent in the model. The marginal ice zone is an area which is extremely sensitive to errors in water masses and surface fluxes and it is difficult to have a

good modeling of the real ice extent without the introduction of additional information from observations.

In this study we developed the assimilation capabilities for ice concentration data using the EnKF. The assimilation becomes multivariate in the sense that both the ice and ocean variables are updated consistently with each other. As an example, if the data introduces ice in an ice free region, then the mixed layer temperature is also set low enough to allow for the presence of ice. Likewise the introduction/removal of ice also implies introduction/removal of salt in the mixed layer.

In the interior of the ice pack there is little impact of the assimilation of ice concentration data. The first reason is that the ice concentration, at least during winter, is close to one with verly little variability. Secondly, the leads parameterization used in the model is not very good and the model under estimates the ice concentration within the icepack.

There is little impact on the total ice volume from the ice concentration data, and the overall ice thickness should be controlled directly with ice thickness data.

3.11.5 Assimilation of ice thickness data

The assimilation of ice thickness data has been examined using simulated data with properties similar to what we can expect from the Cryosat mission. The use of the EnKF again provided a consistent multivariate update of both the ice and ocean variables. The ice thickness data had a positive impact on the modelling of the total ice volume but was somewhat hampered by a too strong bias in the ice model, and associated under estimate of prediction variance. This problem should be addressed by improving the ice model and also the error model use for the ice model.

The use of the Cryosat type data will in combination with the ice concentration data ensure the simulation and prediction of realistic ice cover and allow for operational forecasting of ice parameters.

3.11.6 Assimilation of ocean colour data

Ocean colour data from SeaWiFS were assimilated into the coupled MICOM and ecosystem model and the properties of the assimilation system was examined in detail.

From the study with MICOM we found that the SeaWiFS data had a significat impact on the evolution of the marine ecosystem on the large scales. Thus, with the ecosystem model used and the correlations which exists between different ecosystem variables it is possible to control the evolution of the marine ecosystem on the large scale. This is a promising result related to the ongoing development of large scale and even global ocean and ecosystem data assimilation systems to be used for real time ocean monitoring and prediction.

Even though we recently completed the coupling of the ecosystem model with HYCOM we could not redo these experiments using HYCOM in time for the SIREOC report. HYCOM provides a better framework for coupling to an ecosystem model and we would therefore expect the assimilation results using HYCOM to be better than what could be obtained using MICOM. On the other hand, the actual assimilation precedure and the behaviour of the assimilation scheme is identical when used with MICOM and HYCOM and the general conclusions drawn using MICOM should be valid for HYCOM as well.

Further studies will involve more complex ecosystem models and a downscaling towards the coastal zones. For such applications it is not clear that ocean colour is sufficient to properly control the marine biology.

3.11.7 Overall conclusion

The data sets which have been evaluated in SIREOC for assimilation in operational oceanography applications, are all complementary. Thus, they provide information about different model variables or processes, and they are all important for the optimal modelling of different processes in different regions.

It should be noted that only satellite observed surface information has been used in this project. Thus, we have only been able to control the upper part of the ocean and its circulation. In operational systems, additional in situ data will be essential for ensuring a correct representation of the more slowly evolving deep water masses. The ARGO program supplies in situ vertical profiles of temperature and salinity from the global ocean. We have also developed the capability of assimilating these profiles and further impact studies on the simultaneous assimilation of all parameters will be continued in future studies.

3.11.8 Recommendations

Particular issues of concern involve the following:

1. Start using high resolution SST (GHRSST) data for assimilation and examine the impact in coastal regions.
2. Start using absolute dynamic topography and improved geoid information following the GOCE missions to improve the accuracy and impact of SLA assimilation.
3. Reformulate HYCOM to use SSH as a prognostic rather than a diagnostic variable. This will provide a more realistic representation of SSH in the model which is directly comparable with the observed dynamic topography (using an accurate GEOID).
4. Improve the modelled temperature brightness function and further examine the impact of errors in winds, etc., in the computation of temperature brightness and hence SSS for a better use of the data in the assimilation system.
5. Improve the ice model formulation used to remove identified biases for a better use of the ice data in the assimilation system.
6. Continue the proper estimation and modelling of dynamic model errors and error statistcs in the different kinds of assimilated observations.

Chapter 4

Summary and Recommendation

The main focus of this study has been to assess and quantify the relative impact of different Earth Observation data types for operational ocean prediction and seasonal to interannual environmental monitoring systems as experienced in the Nansen Center model suite (including the DIADEM/TOPAZ system and the Bergen Climate Model). The impact is examined in light of availability of satellite observations of physical oceanographic variables, sea ice variables and marine ecosystem variables for the North Atlantic and Nordic Seas. In so doing model simulations have been conducted with specific focus on assessing the impact of applying existing and planned (simulated) remotely sensed data sets to simulate the mean state and variability in a coupled ocean-sea ice-ecosystem model.

The Hybrid Coordinate Ocean Model (HYCOM) developed by Bleck and coworkers at the University in Miami (Bleck, 2002) has now been included in the data assimilation system. The model is an extension of the previous MICOM model where a generalized vertical coordinate is used. The vertical coordinate is isopycnal in the deep ocean below the thermocline but reverts to z-levels in the mixed layer near the surface. This allows for high vertical resolution in the upper part of the ocean which in turn ensure a better representation of the mixed layer dynamics and the biological processes which are typically active in the upper layers near the surface. In addition, the generalized vertical coordinate allows for the use of a sophisticated vertical mixing scheme based on the k-profile parameterization by Large et al (1994).

The HYCOM model has been operated in real time assimilation mode for three months (in the TOPAZ project). Preliminary analyses of the results lead us to conclude that the assimilation of SLA and SST data work as well with HYCOM as it did with MICOM. In particular the better representation of the upper layers are clearly revealed in comparison to profiles from the GTSPP data base. It also appears that assimilation of SLA provides a better advection of signal vertically into the upper ocean. This is not the case for SST.

The effect of assimilating ice variables into a coupled ice–ocean model system was investigated. Two different ice variables, ice concentration and ice thickness were assimilated in two different hindcast experiments. The results showed that assimilation of ice concentration provided a good mechanism for controlling the location of the ice edge in all seasons, while the impact in the central ice pack was virtually nonexistent in winter. The assimilation of ice thickness was hampered by significant model–induced bias in the EnKF scheme. The assimilation updates in the ice thickness assimilation were nonetheless seen to be realistic. The multivariate assimilation scheme used in the experiments demonstrate that the assimilation of ice concentration and ice thickness has a significant impact upon the properties of the ocean surface layers. To apply ice assimilation to the ice model unly, would clearly not give a realistic update of the coupled model, and should be avoided. Future development of the model system and assimilation scheme needs to adress the existence of bias in the model system, particularly when it comes to ice thickness assimilation.

The EnKF was also set up with a coupled MICOM-Ecosystem formulation and examined in detail in a hindcast experiment. In addition we assimilated SeaWiFS data in real time for three months in May-July 2001 using an

ensemble optimal interpolation (EnOI) scheme. The conclusion from this work was that the ocean colour data, and in particular the spatial distribution, were useful for controlling the evolution of the model chlorophyll and also had a multivariate impact on other biological variables.

The ecosystem model has also been coupled with the HYCOM code, and is currently being validated within TOPAZ. Additional assimilation experiments will thereafter be carried out with the new coupled system. Since the resolution in the thermodynamic mixed layer is improved in HYCOM, the coupling of the OGCM model and the ecosystem model is expected to improve the simulation of the biological processes. Further validation and impact analyses from the use of SeaWiFS ocean color data is also planned.

In the climate change monitoring and research the analysis have been focused on the role SSS play in the Subpolar region, especially on the deep water formation rate, and the subsequent link to the maximum strength of the Atlantic Thermohaline Circulation (THC) via meridional overtuning.

Four model integrations for the period 1948-2000 are analyzed, differing by widely different initial conditions and in horizontal model resolution. The obtained results are similar despite the differences between the model integrations:

- During the period 1948-2000, the THC shows a marked variability on decadal time scales, and also a significant (increasing) trend over the last three decades. Both the trend and the variability correlate, in general, to the NAO index.
- Very high correlations (0.7) are found between the Labrador and Irminger Seas deep convection rate and the maximum THC. It is also clear that the Labrador and Irminger Sea SSS play a pivotal role in the intensity of deep convection in this region, and thereby also on the actual strength of the THC. For instance, the salt anomalies in the early 70's and the early to mid 80's are followed by a rapid decline in the strength of the THC.

It is also found that the density of the convective water masses are of importance for the THC; whereas dense surface waters force the THC in a very direct way, less dense convective waters, although mixing to similar depths, do not necessarily set up a sufficient sub-surface pressure gradient to force the THC.

It is clear from the model simulations that better knowledge of the SSS would lead to much better description and characterization of the THC and its variability and relation as well as influence on the deep convection rate of the waters in the Greenland and Labrador Seas.

The major outcome of the SIREOC study project set in the context of satellite observations is summarized in Table 4.1 according to the two distinct workpackages on Climate change (WP 2000) and Operational oceanography (WP 3000).

The results summarized in Table 4.1 suggest that that the SMOS satellite will most likely detect the transition from one NAO state to the other (typically 3–5 years) as the difference in SSS between high and low NAO-years range between 0.6 (max) to 0.2 (min) psu. Similarly it is concluded that remotely sensed SSS from SMOS have the potential to be used to detect transitions associated with decadal-scale changes in the Atlantic Thermohaline Circulation (*Furevik et al.*, 2002)

The satellite data sets which have been evaluated in this WP 3000 on assimilation in operational oceanography applications, are all complementary; i.e. they provide valuable information about different model variables or processes, and they are all important for the optimal modelling of different processes in different regions.

It should be noted that only satellite observed surface information has been used in this project. Thus, we have only been able to control the upper part of the ocean and its circulation. In operational systems, additional in situ data will be essential for ensuring a correct representation of the more slowly evolving deep water masses. The ARGO program supplies in situ vertical profiles of temperature and, to some extent, salinity, from the global ocean. We have also developed the capability of assimilating these profiles and further impact studies on the simultaneous assimilation of all parameters will be continued in future studies.

WP 2000-Climate Change	Main Outcome
Sensitivity to CO2.	At a doubling of the atmospheric concentration of CO2 the most pronounced change is found in the Atlantic Ocean. Here the salinity increases by 0.1-0.5 psu, with the strongest increase in the two sub-tropical gyres, and smallest change at the high (northern and southern) latitudes and in the equatorial region.
Sensitivity to NAO-variability.	It is found that during high NAO-years, the SSS field are saltier than normal SSS in the North Sea (exceeding 0.3 psu), in the eastern part of the sub-tropical region (up to 0.1 psu), in the polar waters north of Newfoundland (up to 0.2 psu), in the Guyana Current region and south of the Hispaniola islands (about 0.1 psu), and the lower than normal SSS in the waters just north of the Gulf Stream (about 0.1 psu).
Sensitivity to the Atlantic thermohaline circulation.	It is concluded that there is a region of positive correlation between the SSS and the Atlantic THC in the sub-polar gyre and the band between 40-50 degrees North. Furthermore, regions of negative correlations are found in the Gulf of Mexico and over a large region in the South Atlantic sub-tropical gyre.

WP 3000-Operational Oceanography	Main Outcome
Assimilation of SST data	The assimilation of satellite observed SST data provides an important source of information which helps correct for errors and/or biases in the atmospheric heat fluxes used in the model. The major impact is in the upper mixed layer of the model
Assimilation of SLA data	The assimilation of SLA data contain excellent information about the mesoscale variability in the ocean. The assimilation of SLA data allows us to introduce and track eddies and rings very well. They also have an impact on the thermocline depth which is partly determining the upper ocean mesoscale variability. In addition there is a weak impact on the mixed layer temperature which also impacts the sea level height.
Assimilation of SSS data	The assimilation of SSS or brightness temperature TB data introduces a correction to the errors on freshwater fluxes in the model and ensure that the upper ocean water-mass characteristics will be correctly represented in the model.
Assimilation of ice concentration data	The assimilation of ice concentration data is essential to ensure a correct ice extent in the model. The marginal ice zone is an area which is extremely sensitive to errors in water masses and surface fluxes and it is difficult to have a good modeling of the real ice extent without the introduction of additional information from observations. In the interior of the ice pack there is little impact of the assimilation of ice concentration data.
Assimilation of ice thickness data	The assimilation of ice thickness data has a positive impact on the modelling of the total ice volume but was somewhat hampered by a too strong bias in the ice model, and associated under estimate of prediction variance.
Assimilation of ocean colour data	Assimilation of ocean colour data from SeaWiFS into the coupled MICOM and ecosystem model has a significat impact on the evolution of the marine ecosystem on the large scales. Thus, with the ecosystem model used and the correlations which exists between different ecosystem variables it is possible to control the evolution of the marine ecosystem on the large scale.

Table 4.1: Summary ofWP 200 and WP 3000 on the impact, improvement and skill assessment from availability of EO data in operational oceanography

System	Targeted Area	Key Input	Output: Hindcast-Nowcast-Forecast
TOPAZ	North Atlantic and Nordic Seas	Atmospheric forcing data. Remote sensing SLA, SST and ocean color. From spring 2002 also satellite sea ice, and Argo profiling floats.	Ocean currents, temperature, salinity, nutrients, chlorophyll-a, mixed layer depth
MERCATOR	North Atlantic, Azores, Mediterranean	Atmospheric forcing data. SLA, SST and ocean color., and Argo profiling floats.	Ocean currents, temperature, salinity, mixed layer depth
FOAM	Global, North Atlantic, Mediterranean Sea	Atmospheric forcing data. Remote sensing SLA, SST and sea ice, and Argo profiling floats, VOS XBTs.	Ocean currents, temperature, salinity, mixed layer depth
MFS	Mediterranean	Remote sensing SLA and SST, VOS XBTs, Buoy and Argo profiler data	Ocean circulation forecasts at basin scale and selected coastal areas.
MI–POM, NORWECOM, POSEIDON, BOOS, POL3DB, ERSEM	North Atlantic, North European shelf seas, Baltic, Greek Seas, Adriatic Sea.	Atmospheric forcing data, tides, buoy data information, river run-off, satellite SST, SSI and wind, boundary conditions from SYS 1 to SYS 4.	Sea level, storm surges, T; S; currents, sea ice, drift (oil, objects), transports of pollution, dispersion
National Monitoring Programs (NMP)	North Sea, Skagerrak, Mediterranean, Aegean Sea, Baltic	Atmospheric forcing data, River run-off, boundary conditions from SYS 1 to SYS 4. Ferry-box data	Ocean currents, temperature, salinity, mixed layer depth, eutrophication, transport, distribution of fish larvae, organic pollutions and nuclear waste.
WAM	North Atlantic and adjacent seas	Meteorological forcing, wave buoy information, satellite SAR, scatterometer and altimeter data	Wind wave, swell, significant waveheight, wave spectra, wave-period

Table 4.2: Overview of existing modelling and assimilation systems.

There is currently several projects undertaken in the context of development and implementation of operational oceanography system for global monitoring of environment and security (GMES). An underlying demand in this context is the security of long term continuity of EO data beyond Envisat and Jason. Those projects that to some extent capitalize on SIREOC are briefly reported below.

MERSEA Strand–1 (CEC GMES project; initial phase 2003–2004). The marine community has developed and operates both near real time distribution of satellite and in-situ observations and forecasting models, with innovative data assimilation tools as indicated in Table 4.2.

The modelling and data assimilation system can be subdivided into two main categories: a) global ocean systems producing assimilated analysis of the ocean state and forecasts; b) regional/coastal high resolution systems producing user-oriented products, increasing the quality of the global models at the regional/shelf level and extending to ecosystem modelling. The ecosystem modelling, available as first generation pre-operational systems, will be complementary to these existing operational oceanography systems.

A global assessment of these systems has not been carried out yet and it is timely to do so in view of the need for assessment of environmental stress upon the marine ecosystem. MERSEA Strand 1 will establish interfaces between the different ocean operational systems so that inter-comparison can be made, quality of products demonstrated and

further integration of satellite and in-situ observations can be evaluated. This will, in turn, ensure that expected regional characteristics (i.e. advantages/disadvantages) can be reported in the context of problems and lessons learnt for European monitoring for environment and security. Based on this we will subsequently address the procedure for:

- harmonising data gathering, utilisation and analyses,
- product generation and quality control,
- information dissemination, and
- standardised and consistent reporting.

The primary deliverables of MERSEA Strand 1, which can also be considered as the European contribution to GODAE, will be used to assess the current status of European capacity in European monitoring for environment and security. Based on this a set of key recommendations for targeted research and development will be specified and eventually considered for integration in the EU FP 6 programme, notably with the aim to have GMES in operation by 2008.

ROSES (ESA GMES project, initial phase 2003-2004). ROSES (standing for Real-Time Ocean Services for Environment and Security) is a multi-service chain dedicated to operational oceanography, that will be demonstrated for two initial services, notably a) Oil spill and b) Algae Bloom monitoring. These services will deliver information in real-time at ocean variability scale, i.e. hours to few days, resulting from Earth Observations, in-situ measurements, data assimilation and numerical ocean models.

ROSES addresses risks, either physical and biological, from local to much larger scale, where forecasting is almost not existing wrt biological events. Such events are regulated by international agreements, conventions or local bye-laws, the scale of the control being generally in line with the nature of the event. ROSES will provide management tools to manage this regulatory framework.

MERSEA Integrated Project (4 year project approved under FP6 Aeronautics and Space GMES call) will run from 2004 to 2007. MERSEA aims to develop a European system for operational monitoring and forecasting on global and regional scales of the ocean physics, biogeochemistry and ecosystems. The prediction time scales of interest extend from days to months. This integrated system will be the Ocean component of the future GMES system.

At the core of the system is the collection, validation and assimilation of remote sensed and in situ data into ocean circulation models that allow for the self consistent merging of the data types, interpolation in time and space for uniform coverage, nowcasting (i.e. data synthesis in real-time), forecasting, and hindcasting, and delivery of information products.

The project will develop Marine Applications addressing the needs of both intermediate and end-users, whether institutional or from the private sector, with the objective to:

1. Improve the safety and efficiency of maritime transport and naval operations;
2. Enable the sustainable exploitation and management of ocean resources (offshore oil and gas industry, fisheries);
3. More efficiently mitigate the effects of environmental hazards and pollution crisis (oil spills, harmful algal blooms);
4. Improve contribution to ocean climate variability studies and seasonal climate prediction and its effects on coastal populations;
5. Improve national security and reduce public health risks;

6. Advance marine research with the aim to better understand the global climate, the ocean and its ecosystems.

The project will lead to a single high-resolution global ocean forecasting system shared by European partners together with a coordinated network of regional systems for European waters which will provide the platform required for coastal forecasting systems. During the project the main pre-operational systems will be transitioned towards operational status ready for full GMES implementation by 2008.In so doing the possible structure and implementation of a European Centre for Operational Marine Forecasting (ECOMF) will also be addressed.

Bibliography

Aagaard, K., and E. C. Carmack, The role of sea ice and other fresh water in the arctic circulation, *J. Geophys. Res.*, *94*, 14,485–14,498, 1989.

Anderson, L. A., and A. R. Robinson, Physical and biological modeling in the Gulf Stream region: II. physical and biological processes, *Deep-Sea Res.*, *48*, 1139–1168, 2001.

Arakawa, A., and V. R. Lamb, Computational design of the basic processes of the UCLA General Circulation Model, *Methods Comput. Phys.*, *100*, 174–265, 1977.

Austin, R. W., Optical remote sensing of the oceans: Bc (before CSCS) and ac (after CZCS), in *Ocean Colour: Theory and Applications in a Decade of CZCS Experiance*, edited by V. Barale and P. M. Schlittenhardt, pp. 1–15, Kluwer Academic Publishers, 1993.

Baretta, J. W., J. G. Baretta-Bekker, and P. Ruardij, Data needs for ecosystem modelling, *ICES Journal of Marine Science*, *55*, 756–766, 1998.

Barnier, B., M. Bentsen, C. Boone, J.-P. Boulanger, B. Chapron, H. Drange, E. Dinnat, M. Espino, J. Font, T. Furevik, M. Garcia, C. Gommenginger, A. Guissard, J. Johannessen, D. L. G. Larnicol, N. Martin, J. Molines, P.-Y. L. Traon, E. Obligis, N. Reul, J. Schulz, S. Masson, P. Sobieski, M. Srokosz, and J. Verron, Scientific Requirements and Impact of Space Observations of Ocean Salinity for Modelling and Climate Studies, NERSC Technical Report no. 214, 2002, under Contract no 14273/00/NL/DC - ESA.

Belkin, I. M., S. Levitus, J. Antonov, and S.-A. Malmberg, Great salinity anomalies in the North Atlantic, *Progr. Oceanogr.*, *41*, 1–68, 1998.

Bentsen, M., and H. Drange, Parameterizing surface fluxes in ocean models using the NCEP/NCAR reanalysis data, in *RegClim, Regional Climate Development Under Global Warming, General Technical Report*, 4, pp. 149–157, Norwegian Institute for Air Research, Kjeller, Norway, 2000.

Bentsen, M., G. Evensen, H. Drange, and A. D. Jenkins, Coordinate transformation on a sphere using conformal mapping, *Mon. Weather Rev.*, *127*, 2733–2740, 1999.

Berger, M., A. Camps, J. Font, Y. Kerr, J. Miller, J. Johannessen, J. Boutin, M. Drinkwater, N. F. N. Skou, M. Rast, H. Rebhan, and E. Attema, Measuring ocean salinity with esa's smos mission, *esa bulletin*, *Vol*, 113–121, 2002.

Bertino, L., G. Evensen, and H. Wackernagel, Sequential data assimilation techniques in oceanography, *Int. Stat. Rev.*, submitted.

Bleck, R., An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modelling*, *4*, 55–88, 2002.

Bleck, R., and L. T. Smith, A wind-driven isopycnic coordinate model of the North and Equatorial Atlantic Ocean. 1. Model development and supporting experiments, *J. Geophys. Res.*, *95*, 3273–3285, 1990.

Bleck, R., C. Rooth, D. Hu, and L. T. Smith, Salinity-driven thermohaline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic, *J. Phys. Oceanogr., 22*, 1486–1515, 1992.

Bossuet, C., M. Déqué, and D. Cariolle, Impact of a simple parameterization of convective gravity-wave drag in a stratosphere-troposphere general circulation model and its sensitivity to vertical resolution, *Ann. Geophysicae, 16*, 238–249, 1998.

Bougeault, P., A Simple Parameterization of the Large-Scale Effects of Deep Cumulus Convection, *MWR, 113*, 2108–2121, 1985.

Breuillin, C., G. Evensen, and M. Eknes, Data assimilation methods applied to marine ecosystem models, *Tech. Rep. 187*, Nansen Environmental and Remote Sensing Center (NERSC), Edv. Griegs vei 3a, 5059 Bergen, Norway, 2000.

Browning, G. L., and H.-O. Kreiss, Initialization of the shallow water equations with open boundaries by the bounded derivative method, *Tellus, 34*, 334–351, 1982.

Browning, G. L., and H.-O. Kreiss, Scaling and computation of smooth atmospheric motions, *Tellus, Ser. A, 38*, 295–313, 1986.

Brusdal, K., J. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. J. van Leeuwen, E. Dombrowsky, and J. Verron, A demonstration of ensemble-based assimilation methods with a layered ogcm from the perspective of an operational ocean forecasting system, *J. Marine. Sys., 40 – 41*, 253 – 289, 2003, in print.

Bulgarelli, B., and F. Melin, *SeaWiFS data processing code REMBRANDT version 1.0 - REtrieval of Marine Biological Resources through ANalysis of ocean color DaTa - code elements*, Space Applications Institute, Marine Environment Unit and Joint Research centre, 2000, project SAI-05 COAST - EUR19514EN.

Burgers, G., P. J. van Leeuwen, and G. Evensen, On the analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev., 126*, 1719–1724, 1998.

Cariolle, D., and M. Déqué, Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model, *J. Geophys. Res., 91*, 10,825–10,846, 1986.

Cassou, C., P. Noyret, E. Sevault, O. Thual, L. Terray, D. Beaucourt, and M. Imbard, Distributed Ocean–Atmosphere Modelling and Sensitivity to the Coupling Flux Precision: The CATHODe Project, *Mon. Wea. Rev., 126*, 1035–1053, 1998.

Chapman, D. C., and R. C. Beardsley, On the origin of shelf water in the Middle Atlantic Bight, *J. Phys. Oceanogr., 19*, 384–391, 1989.

Coiffer, J., Y. Ernie, J. F. Geleyn, J. Clochard, J. Hoffman, and F. Dupont, The operational hemispheric model at the French Meteorological Service, in *J. Met. Soc. Japan, Special NWP Symposium Volume*, pp. 337–345, 1987.

Courtier, P., C. Freydier, J. F. Geleyn, F. Rabier, and M. Rochas, The ARPEGE project at Météo-France, in *Proceedings of ECMWF workshop on numerical methods in atmospheric modelling, 9-13 Sept. 1991, Vol. 2*, pp. 193–231, ECMWF, Shinfield Park, Reading, Berkshire, UK, 1991.

Curry, R., and M. S. McCartney, Ocean gyre circulation changes associated with the North Atlantic Oscillation, *J. Phys. Oceanogr., 31*, 3374–3400, 2001.

Curry, R. G., M. S. McCartney, and T. M. Joyce, Oceanic transport of subpolar climate signals to mid-depth subtropical waters, *Nature, 391*, 575–577, 1998.

Déqué, M., and J. P. Piedelievre, High resolution climate simulation over Europe, *Clim. Dyn., 11*, 321–339, 1995.

Déqué, M., C. Dreveton, A. Braun, and D. Cariolle, The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling , *Clim. Dyn., 10*, 249–266, 1994.

Déqué, M., P. Marquet, and R. G. Jones, Simulation of Climate change over Europe using a global variable resolution general circulation model, *Clim. Dyn., 14*, 173–189, 1998.

Dickson, R. R., From the Labrador Sea to global change, *Nature, 386*, 649–650, 1997.

Dickson, R. R., J. Meincke, S. A. Malmber, and A. J. Lee, The "great salinity anomaly" in th northern atlantic 1968–1982, *Progr. Oceanogr., 20*, 103–151, 1988.

Douville, H., J.-F. Royer, and J.-F. Mahfouf, A new snow parameterization for the Météo-France climate model. Part II: Validation in a 3D GCM experiment, *Clim. Dyn., 12*, 37–52, 1995.

Drange, H., An Isopycnic Coordinate Carbon Cycle Model for the North Atlantic, *Tech. Rep. 93*, Nansen Environmental and Remote Sensing Center, Edv. Griegsv. 3a, N–5059 Bergen, 1994a.

Drange, H., An Isopycnic Coordinate Carbon Cycle Model for the North Atlantic; and the Possibility of Disposing of Fossil Fuel CO_2 in the Ocean, Ph.D. thesis, Nansen Environm. Remote Sensing Centr. and Dep. of Math., Univ. of Bergen, Bergen, Norway, 1994b.

Drange, H., A 3-dimensional isopycnic coordinate model of the seasonal cycling of carbon and nitrogen in the Atlantic Ocean, *Phys. Chem. Earth, 21*, 503–509, 1996.

Drange, H., and R. Bleck, Multi-dimensional forward-in-time and upstream-in-space based differencing for fluids, *Mon. Weather Rev.*, 1996, in print.

Drange, H., and K. Simonsen, Formulation of air-sea fluxes in the ESOP2 version of MICOM, *Tech. Rep. 125*, Nansen Environmental and Remote Sensing Center, Edv. Griegsv. 3A, N-5037 Solheimsviken, Norway, 1996.

Drange, H., P. Schlussel, and M. Srokosz, Study of Critical Requirements for Ocean Salinity Retrieval Using a Low Frequency Microwave Radiometer, ESA Contract No. 13224/98/NL/MV, Final Report, Earth Sciences Division, ESA–ESTEC, 1999, keplerlaan 1, 2201 AZ Noordwijk, Netherlands (65 pp.). (Available from http://www.nrsc.no/ helge/rieber.html).

Ducet, N., P. L. Traon, and G. Reverdin, Global high resolution mapping of ocean circulation from the combination of topex/poseidon and ers-1/2, *J. Geophys. Res., C8*, 477–498, 2000.

Dümenil, L., K. Isele, H. J. Liebscher, U. Schröder, and K. Wilke, Discharge data from 50 selected rivers for GCM validation, *Tech. Rep. 100*, Max Planck Institute, 1993.

Eknes, M., and G. Evensen, An ensemble Kalman filter with a 1-D marine ecosystem model, *J. Mar. Sys.*, 2001, submitted.

Emery, W. J., C. Fowler, and J. Maslanik, Arctic sea ice concentration from special sensor microwave imager and advanced very high resolution radiometer satellite data, *J. Geophys. Res., 99*, 18,329–18,342, 1994.

Evans, G. T., and J. S. Parslow, A model of annual plankton cycles, *Biol. oceanogr., 3*, 327–347, 1985.

Evensen, G., Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model, *J. Geophys. Res., 97*, 17,905–17,924, 1992.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res., 99*, 10,143–10,162, 1994.

Evensen, G., Advanced data assimilation for strongly nonlinear dynamics, *Mon. Weather Rev., 125*, 1342–1354, 1997.

Evensen, G., and P. J. van Leeuwen, An Ensemble Kalman Smoother for nonlinear dynamics, *Mon. Weather Rev.*, *128*, 1852–1867, 2000.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment, *J. Geophys. Res.*, *101*, 3747–3764, 1996.

Fasham, M. J. R., Modelling the marine biota, in *The global carbon cycle*, edited by M. Heimann, pp. 457–504, Berlin, Springer-Verlag, 1993.

Fasham, M. J. R., H. W. Ducklow, and S. M. McKelvie, A nitrogen-based model of phytoplankton dynamics in the oceanic mixed layer, *J. Mar. Res.*, *48*, 591–639, 1990.

Fichefet, T., and P. Gaspar, A model study of upper ocean-sea ive interaction, *J. Phys. Oceanogr.*, *18*, 181–195, 1988.

Font, J., P. Gaspar, T. Guymer, J. Johannessen, G. van Kolff, C. L. Provost, A. Ratier, D. Williams, and N. Flemming, Operational ocean observations from space, *EuroGOOS Publication No. 16*, 2001.

Friedrich, H., and S. Levitus, An approximation to the equation of state for sea water, suitable for numerical ocean models, *J. Phys. Oceanogr.*, *2*, 514–517, 1972.

Furevik, T., M. Bentsen, H. Drange, J. Johannessen, and A. Korablev, Spatial and temporal variability of the sea surface salinity in the nordic seas, *J. Geophys. Res.*, *107*, 2002.

Furevik, T., M. Bentsen, H. Drange, I. K. T. Kindem, N. G. Kvamstø, and A. Sorteberg, Description and validation of the Bergen Climate Model: ARPEGE coupled with MICOM, *Clim. Dyn.*, 2003, in press.

Gargett, A. E., Vertical eddy diffusivity in the ocean interior, *J. Mar. Res.*, *42*, 359–393, 1984.

Gaspar, P., Y. Grégoris, and J.-M. Lefevre, A simple eddy kinetic model for simulations of the oceanic vertical mixing: Tests at Station Papa and Long-Term Upper Ocean Study Site, *J. Geophys. Res.*, *95*, 16,179–16,193, 1990.

Geleyn, J. F., Use of a modified Richardson number for parameterizing the effect of shallow convection, in *J. Met. Soc. Japan, Special NWP Symposium Volume*, pp. 141–147, 1987.

Geleyn, J. F., Interpolation of wind, temperature and humidity values from model levels to the height of measurement, *Tellus*, *40A*, 347–351, 1988.

Geleyn, J. F., and H. J. Preuss, A new dataset of satellite derived surface albedo values for operational use at ECMWF, *Arch. Met. Geoph. Biocl., Ser. A*, *32*, 353–359, 1983.

Gloersen, P., W. J. Campbell, D. J. Cavalieri, J. C. Comiso, C. L. Parkinson, and H. J. Zwally, *Arctic and Antarctic sea ice, 1978–1987: Satellite passive-microwave observations and analysis*, NASA Special Publication NASA SP-511, National Aeronautics and Space Administration, 1992.

GOCE group, GOCE Mission Requirments Document (MRD), ESA document EE–GOCHE version 2.0, 1998, available at URL: http://www-sv.cict.fr/cesbio/smos/Proposal.html.

Gross, L., S. Thiria, R. Frouin, and B. G. Mitchell, Artificial neural networks for modeling the transfer function between marine reflectance and phytoplankton pigment concentration, *J. Geophys. Res.*, *105*, 3483–3495, 2000.

Guilyardi, E., and G. Madec, Performance of the OPA/ARPEGE–T21 global ocean–atmosphere coupled model, *Clim. Dyn.*, *13*, 149–165, 1997.

Harder, M., Dynamik, rauhigkeit und alter des meereises in der arktis, Ph.D. thesis, Alfred-Wegner-Institut für Polar- und Meereisforschung, Bremerhaven, Germany, 1996.

Haugen, V. E. J., and G. Evensen, Assimilation of sla and sst data into an ogcm for the indian ocean, *Ocean Dynamics*, *52*, 133–151, 2002.

Hibler, W. D., III, A dynamic thermodynamic sea ice model, *J. Geophys. Res.*, *84*, 815–846, 1979.

Holland, D. M., and A. Jenkins, Modelling thermodynamic ice-ocean interactions at the base of an ice shelf, *J. Phys. Oceanogr.*, *29*, 1787–1800, 1999.

Holland, M. M., C. M. Bitz, M. Eby, and A. J. Weaver, The role of ice–ocean interactions in the variability of the north atlantic thermohaline circulation, *J. Clim.*, *14*, 656–675, 2001.

Holloway, G., and T. Sou, Has Arctic sea ice rapidly thinned?, *J. Clim.*, *15*, 1691–1701, 2002.

Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, An overview of SeaWiFS and ocean color, in *NASA Technical Memorandum 104566 - SeaWiFS Technical Report Series*, edited by S. B. Hooker and E. R. Firestone, vol. 1, NASA, 1992.

Hortal, M., Aspects of the numerics of the ECMWF model, in *Proc. Recent developments in numerical methods for atmospheric modelling, 7–11 September*, pp. 127–143, 1998.

Hortal, M., and A. J. Simmons, Use of reduced Gaussian grids in spectral models, *Mon. Wea. Rev.*, *119*, 1057–1074, 1991.

Houtekamer, P. L., and H. L. Mitchell, Data assimilation using an ensemble Kalman filter technique, *Mon. Wea. Rev.*, *126*, 796–811, 1998.

Hughes, C., and A. Ash, Eddy forcing of the mean flow in the southern ocean, *J. Geophys. Res.*, *C2*, 2713–2722, 2001.

Hunke, E. C., and J. K. Dukowicz, An elastic-viscous-plastic model for sea ice dynamics, *J. Phys. Oceanogr.*, *27*, 1849–1867, 1997.

J.A.Johannessen, C. le Provost, H. Drange, M. Srokosz, P. Woodworth, P. Schlussel, P. L. Grand, Y. Kerr, D. Wingham, and H. Rebhan, Observing the ocean from space - emerging capabilities in europe, in *Observing the ocean in the 21st century*, edited by C. Koblinsky and N. Smith, Bureau of Meteorology, Melbourne, Australia, 2001.

Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, San Diego, Calif., 1970.

Johannessen, J. A., G. Balmino, C. Le Provost, R. Rummel, H. S. R. Sabadini, C. Tscherning, P. Visser, C. P. Woodworth, W. Hughes, P. LeGrand, N. Sneeuw, F. Perosanz, M. Aguirre-Martinez, H. Rebhan, and M. Drinkwater, The european gravity field and steady-state ocean circulation explorer satellite mission: Impact in geophysics, *Sur. Geophys.*, 2003, in press.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, A. Leetmaa, R. Reynolds, and R. Jenne, The NCEP/NCAR reanalysis project, *Bull. Am. Met. Soc*, *77*, 437–471, 1996.

Keiner, L. E., and C. W. Brown, Estimating oceanic chlorophyll concentrations with neural networks, *Int. J. Remote Sens.*, *20*, 189–194, 1999.

Kerr, Y., SMOS: Proposal in answer to the call for earth explorer opportunity missions. (reference: Cop 16), Available at URL: http://www-sv.cict.fr/cesbio/smos/Proposal.html, November 1998.

Kwok, R., Sea ice concentration estimates from satellite passive microwave radiometry and openings from SAR ice motion, *Geophys. Research Letters*, *29*, 10.1029/2002GL014,787, 2002.

Large, W. G., J. C. McWilliams, and S. C. Doney, Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, *Rev. of Geophys.*, *32*, 363–403, 1994.

Laxon, S., Simulated ice thickness from radar altimeter (task 4.1), in *The quantification of the importance of the sea ice budget in the climate system*, Technical report no. 207, Nansen Environmental and Remote Sensing Center, 2001.

Le Traon, P., P. Gaspar, F. Bouyssel, and H. Makhmara, Using topex/poseidon data to enhance ERS-1 orbit, *J. Atmos. Ocean. Tech., Vol.*, 161–170, 1995.

Le Traon, P., F. Nadal, and N. Ducet, An improved mapping method of multi-satellite altimeter data., *J. Atmos. Ocean. Tech., Vol*, 522–534, 1998.

Legates, D., and C. Willmott, Mean seasonal and spatial variability in gauge–corrected global precipitation, *Int. J. Climatol., 10*, 110–127, 1990.

Levitus, S., and T. P. Boyer, *World Ocean Atlas 1994 Volume 4: Temperature*, NOAA Atlas NESDIS 4, Washington, D.C., 1994.

Levitus, S., R. Burgett, and T. P. Boyer, *World Ocean Atlas 1994 Volume 3: Salinity*, NOAA Atlas NESDIS 3, Washington, D.C., 1994.

Lisæter, K. A., Assimilation of simulated ice thickness data into a coupled ice–ocean model(task 4.2), in *The quantification of the importance of the sea ice budget in the climate system*, Technical report no. 207, Nansen Environmental and Remote Sensing Center, 2001.

Lisæter, K. A., J. Rosanova, and G. Evensen, Assimilation of ice concentration in a coupled ice-ocean model, using the ensemble kalman filter, *Ocean Dyn.*, 2003, in print.

Lott, F., Alleviation of Stationary Biases in a GCM through a Mountain Drag Parameterization Scheme and a Simple Representation of Mountain Lift Forces, *Mon. Wea. Rev., 125*, 788–801, 1999.

Lott, F., and M. J. Miller, A new subgrid-scale orographic drag parameterization: Its formulation and testing, *Q. J. R. Meteorol. Soc., 123*, 101–127, 1997.

Louis, J. F., A Parametric Model of Vertical Eddy Fluxes in the Atmosphere, *Boundary-Layer Meteor., 17*, 187–202, 1979.

Madec, G., and M. Imbard, A global ocean mesh to overcome the North Pole Singularity, *Clim. Dyn., 12*, 381–388, 1996.

Mahfouf, J.-F., A. O. Manzi, J. Noilhan, H. Giordani, and M. Déqué, The land surface scheme ISBA within the Météo-France climate model ARPÉGE. Part I: Implementation and preliminary results, *J. Climate, 8*, 2039–2057, 1995.

Masutani, S. M., C. M. Kinoshita, G. C. Nihous, H. Teng, L. A. Vega, and S. K. Sharma, Laboratory Experiments of CO_2 Injection into the Ocean, in *Direct Ocean Disposal of Carbon Dioxide*, edited by N. Handa and T. Ohsumi, pp. 123–141, Terra Scientific Publ. Comp., 1995.

McDougall, T. J., and W. K. Dewar, Vertical Mixing and Cabbeling in Layered Models, *J. Phys. Oceanogr., 28*, 1458–1480, 1998.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, Ensemble size, Balance, and Model–Error Representation in an Ensemble Kalman Filter, *Mon. Wea. Rev., 130*, 2791–2808, 2002.

Montgomery, R. B., Circulation in upper layers of southern North Atlantic deduced with use of isentropic analysis, in *Papers in Phys. Oceanogr. and Meteorol.*, vol. VI, p. 55 pp., Woods Hole Oceanogr. Inst., Woods Hole, Mass., 1938.

Morcrette, J. J., Radiation and cloud radiative properties in the European Centre for Medium Range Forecasts forecasting system, *J. Geophys. Res.*, *95*, 9121–9132, 1991.

Nagurnyi, A. P., V. G. Korostelev, and V. V. Ivanov, Multiyear variability of sea ice thickness in the Arctic basin measured by elastic-gravity waves on the ice surface, *Meteor. Hydrol.*, *3*, 72–78, 1994.

Natvik, L. J., and G. Evensen, Assimilation of ocean colour data into a biochemical model of the North Atlantic. Part 1. Data assimilation experiments, *J. Marine. Sys.*, *40–41*, 155–169, 2003a.

Natvik, L. J., and G. Evensen, Assimilation of ocean colour data into a biochemical model of the North Atlantic. Part 2. Statistical analysis, *J. Marine. Sys.*, *40–41*, 155–169, 2003b.

Nilsen, J. E. Ø., Y. Gao, H. Drange, T. Furevik, and M. Bentsen, Simulated North Atlantic-Nordic Seas water mass exchanges in an isopycnic coordinate OGCM, *Geophys. Res. Lett.*, 2003, in press.

Noilhan, J., and S. Planton, A Simple Parameterization of Land Surface Processes for Meteorological Models, *Mon. Wea. Rev.*, *117*, 536–549, 1989.

Parkinson, C. L., and W. M. Washington, A Large-Scale Numerical Model of Sea Ice, *J. Geophys. Res.*, *84*, 311–337, 1979.

Peng, T. H., T. Takahashi, W. S. Broecker, and J. Olafsson, Seasonal variability of carbon dioxide, nutrients and oxygen in the North Atlantic surface water: Observations and a model, *Tellus*, *39B*, 439–458, 1987.

Pham, D. T., J. Verron, and M. C. Roubaud, A singular evolutive extended Kalman filter for data assimilation in oceanography, *J. Marine. Sys.*, *16*, 323–340, 1998.

Pickart, R., and K. Lavender, Is Labrador Sea Water Formed in the Irminger Basin?, *WOCE Newsletter*, *39*, 6–8, 2000.

Pickart, R. S., F. Straneo, and G. W. K. Moore, Is Labrador Sea Water formed in the Irminger Basin?, *Deep Sea Res.*, 2002, submitted.

Provost, C., J. LeGrand, E. Dombrowsky, P. Traon, M. Losh, F. Ponchaut, J. Schröter, B. Sloyan, and N. Snnuw, Impact of GOCE for ocean circulation studies, ESA study contract 13175/98/NLGD, Final Report, 1999, eSA–ESTEC, Noordwijk, Netherlands.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, Hydrologic data assimilation with the Ensemble Kalman Filter, *Mon. Wea. Rev.*, *1*, 103–114, 2002.

Richard, J. L., and J. F. Royer, A statistical cloud scheme for use in an AGCM, *Ann. Geophysicae*, *11*, 1093–1115, 1993.

Rothrock, D. A., Y. Yu, and G. A. Maykut, Thinning of the Arctic sea-ice cover, *Geophys. Research Letters*, *23*, 3469–3472, 1999.

Rudels, B., H. J. Friedrich, and D. Quadfasel, The Arctic circumpolar boundary current, *Deep-Sea Res. II*, *46*, 1023–1062, 1999.

Sarmiento, J. L., R. D. Slater, M. J. R. Fasham, H. W. Ducklow, J. R. Toggweiler, and G. T. Evans, A seasonal three-dimensional ecosystem model of nitrogen cycling in the North Atlantic euphotic zone, *Global Biogeochem. Cycles*, *7*, 417–450, 1993.

Semtner, J., A. J., A model for the thermodynamic growth of sea ice in numerical investigations of climate, *J. Phys. Oceanogr.*, *6*, 379–389, 1976.

Simmons, A. J., and D. M. Burridge, An Energy and Angular Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinate, *Mon. Wea. Rev.*, *109*, 758–768, 1981.

Slutz, R., S. L. J. Hiscox, S. Woodruff, R. Jenne, D. Joseph, P. Steurer, and J. Elms, *Comprehensive Ocean-Atmosphere Data Data Set; Release 1,NTIS PB86-105723*, NOAA Environmental Research Laboratories, Climate Research Program, Boulder, Col., 1985.

Smolarkiewicz, P. K., A Fully Multidimensional Positive Definite Advection Transport Algorithm with Small Implicit Diffusion, *J. Comput. Phys.*, *54*, 325–362, 1984.

Smolarkiewicz, P. K., and T. L. Clark, The multidimensional positive definite advection transport algorithm: Further development and applications, *J. Comput. Phys.*, *67*, 396–438, 1986.

SMOS, SMOS mission objectives and scientific requirements (version 5), esa earth explorer, opportunity mission, ESA document EEOM-SMOS-MRD: http://www.esa.int/livingplanet/smos, 2001, eSA-ESTEC, Noordwijk, Netherlands.

SP-1196(1), Report for assessment: Gravity Field and Steady-State Ocean circulation Mission, ESA Special Publication, SP-1196(1), 1996, eSA Publication Division, ESTEC, Noordwijk, Netherlands.

SP-1227, The earth explorer: The science and research element of esa's future earth observation programme, ESA Special Publication, SP-1227, 1998, eSA Publication Division, ESTEC, Noordwijk, Netherlands.

SP-1233(1), Report for Mission Selection: Gravity Field and Steady-State Ocean circulation Mission, ESA Special Publication, SP-1233(1), 1999, eSA Publication Division, ESTEC, Noordwijk, Netherlands.

Sünkel, H., R. Rummel, R. Klees, K. Ilk, R. Sabadini, C. Tscherning, and F. Sanso, From Eotvös to mGal, ESA study contract, ESTEC,No. 13392/98/NL/GD, Final Report., 2000.

Svendsen, E., K. Kloster, B. Farrelly, O. M. Johannessen, J. A. Johannessen, W. J. Campbell, P. Gloersen, D. Cavalieri, and C. Mätzler, Norwegian remote sensing experiment: Evaluation of the nimbus 7 scanning multichannel microwave radiometer for sea ice research, *J. Geophys. Res.*, *88*, 2781–2791, 1983.

Terray, L., and O. Thual, Oasis: le couplage océan–atmosphère, *La Météorologie*, *10*, 50–61, 1995.

Terray, L., O. Thual, S. Belamari, M. Déqué, P. Dandin, C. Lévy, and P. Delecluse, Climatology and interannual variability simulated by the arpege–opa model, *Clim. Dyn.*, *11*, 487–505, 1995.

Terray, L., S. Valcke, and A. Piacentini, OASIS 2.2 Ocean Atmosphere Sea Ice Soil. User's Guide and Reference Manual, *Tech. rep.*, CERFACS, 1998.

Thorndike, A. S., and R. Colony, Sea ice motion in response to geostrophic winds, *J. Geophys. Res.*, *87*, 5845–5852, 1982.

Thorndike, A. S., D. A. Rothrock, G. A. Maykut, and R. Colony, The thickness distribution of sea ice, *J. Geophys. Res.*, *80*, 4501–4513, 1975.

Whitaker, J. S., and T. M. Hamill, Ensemble data assimilation without perturbed observations, *Mon. Wea. Rev.*, *130*, 1913–1924, 2002.

Wingham, D., CryoSat Mission Requirements Document (MRD), ESA document., 1999, ESA-ESTEC, Noordwijk, Netherlands.

Zalesak, S., Fully multidimensional flux-corrected transport algorithms for fluids, *J. Comp. Physics*, *31*, 335–362, 1979.