



Integrated Arctic Observation System

Research and Innovation Action under EC Horizon2020
Grant Agreement no. 727890

Project coordinator:
Nansen Environmental and Remote Sensing Center, Norway

Deliverable 2.12


Observational gaps revealed by model sensitivity to observations

Start date of project:	01 December 2016	Duration:	60 months
Due date of deliverable:	30 November 2018	Actual submission date:	3 December 2018
Lead beneficiary for preparing the deliverable:	UHAM		
Person-months used to produce deliverable:	28.3 pm		

Authors: Detlef Stammer (UHAM), Guokun Lyu (UHAM), Roberta Pirazzini (FMI), Tuomas Naakka (FMI), Tiina Nygård (FMI), Timo Vihma (FMI), Martijn Pallandt(MPG), Mathias Göckede (MPG), Friedemann Reum (MPG).

Reviewed by: Michael Tjernström (MISU)

Version	DATE	CHANGE RECORDS	LEAD AUTHOR
1.0	24/10/2018	Template	Roberta Pirazzini
1.1	15/11/2018	1 st Draft	Guokun Lyu
1.2	22/11/2018	2 nd Draft	Guokun Lyu
1.3	23/11/2018	Minor modifications	Guokun Lyu
1.4	29/11/2018	Review	Michael Tjernström
2.0	02/12/2018	Minor modifications based on review	Guokun Lyu
2.0	03/12/2108	Minor edits of the lay-ou	Kjetil Lygre

Approval X	Date: 3 December 2018	Sign.  Coordinator
-----------------------------	--------------------------	--

USED PERSON-MONTHS FOR THIS DELIVERABLE					
No	Beneficiary	PM	No	Beneficiary	PM
1	NERSC		24	TDUE	
2	UiB		25	GINR	
3	IMR		26	UNEXE	
4	MISU		27	NIVA	
5	AWI		28	CNRS	
6	IOPAN		29	U Helsinki	
7	DTU		30	GFZ	
8	AU		31	ARMINE	
9	GEUS		32	IGPAN	
10	FMI	3.3	33	U SLASKI	
11	UNIS		34	BSC	
12	NORDECO		35	DNV GL	
13	SMHI		36	RIHMI-WDC	
14	USFD		37	NIERSC	
15	NUIM		38	WHOI	
16	IFREMER		39	SIO	
17	MPG	10	40	UAF	
18	EUROGOOS		41	U Laval	
19	EUROCEAN		42	ONC	
20	UPM		43	NMEFC	
21	UB		44	RADI	
22	UHAM	15	45	KOPRI	
23	NORUT		46	NIPR	
			47	PRIC	

DISSEMINATION LEVEL		
PU	Public, fully open	X
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

EXECUTIVE SUMMARY

To understand the quality of the existing observing system in the Arctic to capture important elements of change over the Arctic we performed a gap analysis with respect to the Arctic Ocean, the Arctic atmosphere and the high-latitude carbon-monitoring network. The main points of the findings are:

1) The ocean observing system: The satellite altimeter system is a critical system to monitor the high-frequency variability. Due to the presence of sea ice in winter time, most of the area can be observed only every 5-10 days, leading to large observing gaps. Closing the gap can be done with new arrays of bottom pressure sensors such as tide gauges or moorings in the ocean bottom. In addition, high-frequency transport measurements are required in the Fram, Davis Straights, the Barents Sea Opening, and north of the Laptev Sea. On the seasonal cycle, bottom pressure observations from GRACE are required to monitor the mass related variability and sea-ice observations are crucial for monitoring the halosteric related variability. On decadal time scales, it is important to have a sufficient hydrographic observing component capable of capturing temperature and salinity changes over the entire Arctic Ocean from the surface to the bottom. New algorithms that can recover sea level from sea ice covered areas may help to improve current satellite altimeter systems, and to improve the ability to monitor the Beaufort Gyre.

2) The atmosphere observing system: The density of the existing radiosonde observation network is not the most critical factor for the quality of T850 forecast. Instead, the results pointed out that stations on small islands in the middle of the Atlantic Ocean are critical for the quality of analysis. The Central Arctic Ocean and the Northern North-Atlantic would probably benefit most from new sounding stations. Efforts to improve the quality of radiosonde observations, especially in Russia, would be very beneficial for the quality of T850 forecasts in the Arctic and sub-Arctic. Current data assimilation systems are probably not adequate to optimally exploit the information from the existing observational network.

3) GHG fluxes observing system: The existing network of pan-Arctic atmospheric monitoring sites provides continuous, well-calibrated observations on atmospheric greenhouse gas mixing ratios, generating basic information to quantify surface-atmosphere greenhouse gas exchange processes for most regions in Canada, Europe, and Western Russia; also the Arctic Ocean receives good overall data coverage. Regions showing limited data coverage include the Russian Far East, Western Alaska, and the Eastern Canadian Provinces. Areas where footprint coverage gaps exist seasonally include parts of Western Russia and Central Siberia. Investments in observational infrastructure in any of these areas would be beneficial to increase the overall coverage of the pan-Arctic atmospheric network for greenhouse gases.

Table of Contents

Table of Contents	2
1. Introduction	3
1.1. The Arctic Ocean circulation and observing system (UHAM)	3
1.2. The Arctic atmosphere observing system (FMI)	4
1.3. The Arctic carbon observing system (MPG)	5
2. Experiment designs and observations	6
2.1. Ocean experiments and observations (UHAM)	6
2.1.1 ATL model simulations.....	6
2.1.2 The ocean observing systems.....	7
2.1.3 Relation of sea level with mass and steric contribution.....	10
2.1.4 The adjoint method.....	11
2.2. Atmospheric experiments and observations (FMI)	12
2.2.1 Sounding data.....	13
2.2.2 ECMWF data assimilation	13
2.2.3 Trajectory analyses.....	14
2.2.4 SOM analyses.....	14
2.3. Evaluating the pan-Arctic atmospheric greenhouse gas monitoring network (MPG) 15	
3. Results	18
3.1. Evaluation of the ocean observing system (UHAM)	18
3.1.1 Statistics of Sea Surface Height and Bottom Pressure.....	18
3.1.2 Sea level variability.....	21
3.1.3 Relation of sea level variability to mass and steric contributions.....	22
3.1.4 Existing observational system and gaps revealed by adjoint sensitivity	28
3.2. Evaluation of the atmosphere observing system (FMI)	33
3.2.1 Differences between analyses and forecasts.....	33
3.2.2 Differences between soundings and forecasts	35
3.2.3 Weight of soundings.....	37
3.2.4 Effects of air-mass origin.....	41
3.2.5 Effects of synoptic-scale circulation patterns.....	43
3.3. Evaluation of the pan-Arctic atmospheric greenhouse gas monitoring network (MPG) 45	
3.3.1. Single site footprint coverage: an example of the Ambarchik monitoring site.....	45
3.3.2. Network footprint coverage: Seasonality of pan-Arctic footprints.....	46
4. Summary of the identified gaps and recommendations	49
4.1. The ocean observing system (UHAM)	49
4.2. The atmosphere observing system (FMI)	50
4.3. The GHG fluxes observing system (MPG)	51
List of references.....	52

1. Introduction

1.1. The Arctic Ocean circulation and observing system (UHAM)

The Arctic Ocean plays a vital role in the climate system due to the high albedo of the sea ice and massive storage of freshwater. Due to global warming, the Arctic is experiencing a rapid loss of sea ice and an increase in freshwater content which may have far-reaching consequences (Notz & Stroeve, 2016).

In the vertical, the Arctic ocean can be separated into three primary layers (Aagaard et al., 1985) with different circulation patterns and driving mechanisms. The upper layer, which includes the pycnocline, stores more than 70000 km³ freshwater (Giles et al., 2012) and has been the subject of many studies. Based on model simulations, Proshutinsky and Johnson (1997) suggest two possible wind-driven circulation regimes in the central Arctic Ocean: the anticyclonic winter ocean circulation and the cyclonic summer ocean circulation. The two circulation regimes also show decadal variability. The wind forcing controls the accumulation and release of the freshwater through Ekman pumping. Freshwater from the Pacific Ocean, river run-off, ice melting, and precipitation is stored in the surface layer under anticyclonic wind forcing (Proshutinsky et al., 2002), increasing the sea surface height in the Canadian Basin. When the wind is weaker or cyclonic, the sea level gradient between the Beaufort Gyre and the Atlantic Ocean drives the freshwater into deep water formation region in the Atlantic Ocean through Fram Strait and passages of Canadian Arctic Archipelago (CAA), which may slow global thermocline circulation. (Häkkinen & Proshutinsky, 2004). Köhl and Serra (2014) propose that it is the sea level gradient of the Arctic periphery to the Atlantic Ocean, caused by weaker anticyclonic wind or cyclonic wind, that leads to transport of fresh water to the Atlantic Ocean. Observations in the Arctic Ocean are mostly limited to the summer season. Based on thirteen years of observations, the liquid freshwater content in summer increased 5410 km³ from 2003 to 2010, decreased a bit in 2011-2014, but in 2015 reached its absolute maximum of 22,600 km³ i.e., 5600 km³ over the climatology¹. Besides the wind-forced Ekman pumping mechanism, other factors, such as ice melting (McPhee et al., 1998) and river run-off (Macdonald et al., 1999), also contribute to the accumulation of freshwater in the Canadian basin, especially in the Beaufort Gyre. Based on observations, Giles et al. (2012) propose that declining and deformation of the sea ice increases momentum transfer to the ocean and accelerates the Beaufort Gyre, accumulating more freshwater. Morison et al. (2012) argue that redistribution of river-run from the Eurasian basin contributes to the increased freshwater content in the Canadian Basin.

The middle layer of the Arctic Ocean is occupied mainly by Atlantic Water with a temperature larger than 0 °C. The warm and salty Atlantic Water enters into the Arctic by two inflows through the Fram Strait and the Barents Sea Opening. The Fram Strait branch sinks beneath the fresher and colder surface water, isolated from the surface layer by the halocline, and flows following the bathymetry of the Eurasian Basin. The Barents Sea branch undergoes considerable modification due to heat loss to the atmosphere and exits the Barents Sea through the St. Anna Trough (Schauer et al., 2002). These two branches of Atlantic Water merge north of the Kara Sea and sink to depth ~500m. The merged Atlantic Water flows along the Eurasian Basin in a cyclonic sense trapped by the topography and splits near the Lomonosov Ridge. One branch follows the Lomonosov Ridge, flows northward and exits the Arctic to the Nordic Seas through Fram Strait. The other part enters the Canadian Basin. Fol-

¹ <http://www.whoi.edu/page.do?pid=153276>

lowing the topography, the Atlantic Water generates two cyclonic circulations: one in the Makarov Basin, and one in the Beaufort Gyre. Substantial variability was observed over the past decades. Shifts in the atmospheric circulation pattern have resulted in increased transport of Atlantic Water into the Arctic via Fram Strait (Quadfasel et al., 1991). Mooring observations in the Nansen Basin also show spreading of the Atlantic Water along the Eurasian Basin (Dmitrenko et al., 2008). Due to limited observations, causes and consequence of the Arctic Atlantic water variability are not very well understood.

In the third layer, the cold low-salinity deep-water in the Greenland Sea enters into the Arctic Ocean west of Spitsbergen and the relative warm saline water exits the Arctic Ocean along Greenland slope. The circulation is limited to the Eurasian Basin by the Lomonosov Ridge.

The Arctic ocean observing system, including fixed moorings, tide gauges, satellite altimetry, temperature/salinity profiles observed by shipboard equipment, are very sparse in space and time. Several mooring systems are deployed in the straits connecting the Arctic Ocean with the Pacific Ocean, and the Atlantic Ocean. Satellite altimetry observations, which provide better spatial coverage, are limited by sea ice. The tide gauges are distributed along the coasts and provide the longest lasting sea-level observations. Since the existing observing systems are deployed in different geographic locations with different observing frequency and variables, they may observe the Arctic variability over different timescale and may also reflect variability propagated from upstream. Therefore, it is critical to evaluate the potential effects of existing observing system on monitoring the Arctic ocean variability.

In this study, we will evaluate the potential effects of satellite altimetry and mooring observing systems on monitoring the Arctic changes using a suite of forward model simulations and adjoint model simulations. First, we compare the model simulations with tide gauges and bottom pressure records to identify the dynamic processes that the model can simulate. Second, based on model simulations we identify regions with high and low sea level variability as a function of timescale, which points out key regions and observing frequency required. Contributions of halo/thermohaline effects (salinity/temperature changes) and mass effects on sea level variability are analyzed, which gives alternative observing options if sea surface height cannot be observed. Then, five adjoint model simulation are performed: 1) Two adjoint model runs are performed to demonstrate the importance of observing upstream variability for monitoring the high-frequency sea-level variability in marginal seas. 2) In a third adjoint model run, we analyzed the potential effect of sea surface height from satellite altimeter on monitoring the Beaufort Gyre decadal variability. 3) Based on the last two adjoint model run, we analyzed the potential effect of observed freshwater/heat transport by the mooring system on monitoring the Arctic circulation.

1.2. The Arctic atmosphere observing system (FMI)

The radiosounding network is a critical component of the atmosphere observing system in the Arctic, consisting of 76 sounding stations located north of 60°N. The Arctic network is relatively denser in Northern Europe and in Western Russia, less dense in Eastern Russia and North America whereas no radiosonde observations are regularly made over the Arctic Ocean. Radio soundings, as well as many measurements from surface-stations and satellites, are assimilated into numerical weather prediction models, and are important for improving weather forecasts. Furthermore, radiosonde observations are used as a reference data for bias-correction of satellite soundings and aircraft data and also for forecast verification (Ingleby et al., 2016). It has been suggested that the relatively low skill of weather forecasts in the Arctic is, at least partly, due to the relatively sparse observational coverage (Jung et

al., 2016). According to observing system experiments (OSEs), additional Arctic observations (from new locations or with increased frequency) by radiosondes can substantially improve the forecasts, and contribute to a more accurate reproduction of the atmospheric circulation, both in the Arctic and on mid-latitudes (Inoue et al., 2015; Sato et al., 2017). Impact of additional observations has, however, been found to be flow-dependent (Inoue et al., 2015). On the other hand, the available observations have been found to usually be adequate for representing synoptic systems in the weather prediction models (Jung et al., 2016).

Recognizing that atmospheric observations in the Arctic, especially in the Central Arctic, are expensive and logistically challenging, and that new observations are not easily obtainable, it is important to evaluate the existing spatial coverage and critical gaps of the radiosonde network from the point of view of numerical weather prediction (NWP) in the Arctic. This can be done either by so-called data denial experiments using NWP models with different data sets assimilated in different experiments or by analyzing the existing model products and observations. Here we have applied the second approach. Our analyses are based on the operational weather forecasting model of European Centre of Medium Range Weather Forecast (ECMWF) and radiosonde observations from the 76 Arctic stations. By comparing the operational analyses, 12 h forecasts valid at the analysis time, and radiosonde observations, we aim to identify the geographical areas where additional radiosonde observations could potentially improve analyses and forecasts. In addition, we address spatial differences in the impacts of soundings in analyses and in short forecasts and identify the most important sounding stations in the Arctic. We also analyze the effects of synoptic-scale circulation patterns to differences between forecasts and soundings. Based on these results, we make suggestions for the future development of the Arctic radiosonde observing network.

1.3. The Arctic carbon observing system (MPG)

Atmospheric monitoring of carbon species mixing ratios can integrate signals from carbon cycle processes covering very large source areas (Desai et al., 2015), thus are ideally suited to support large-scale and long-term monitoring programs to constrain net budgets of carbon exchange between surface and atmosphere. This approach to data-based quantification of greenhouse gases is of particular interest in regions such as the Arctic, where a combination of lacking infrastructure and harsh climate severely limits the application of direct flux measurements such as, e.g. eddy-covariance towers, or flux chambers (Goodrich et al., 2016; Kittler et al., 2017). However, since high quality standards are required to facilitate flux constraints based on atmospheric mixing-ratio monitoring (Miller et al., 2014), the number of these observations is limited in many areas (Thompson et al., 2017).

The most comprehensive tool to constrain surface-atmosphere exchange processes based on time series of atmospheric trace gas mixing ratios is called atmospheric inverse modeling (e.g. Gurney et al., 2002). Atmospheric inverse modeling approaches link the temporal variability of mixing ratio signals captured at the tower locations to distributed, time-varying flux fields at the surface through atmospheric transport modeling. Combined with additional data sources that e.g. specify how flux fields can be structured in both space and time, final posterior flux rates can be constrained through optimization approaches such as Bayesian optimization (Rödenbeck et al., 2009; 2018), or Kalman Filters (Peters et al., 2007). Further extensions of the method include elements of geo-statistical modeling (e.g. Göckede et al., 2010; Michalak et al., 2004), which replace rigid priors with information on how flux fields are correlated across time and space, allowing an unbiased flux estimate and analysis of links between environmental controls and carbon fluxes (e.g. Miller et al., 2014; Yadav et al., 2010). This approach allows merging the information contained in atmospheric observa-

tion programs with arbitrary ancillary information sources, such as satellite remote sensing fields, or bottom-up constraints on flux emissions (e.g. icebreaker datasets from shelf areas).

Depending on scope and target domain of a study, atmospheric transport modeling is used to simulate trajectories of air masses over 15 days or more backwards in time, therefore the fields of view (a.k.a. footprints) of atmospheric mixing ratio observations can cover distances of several thousands of kilometers. The relative influence of each surface patch declines with traveling time, therefore the observed signals are heavily influenced by the so-called 'near field' that surrounds the towers, while the 'far field' contributes much less to the collected information. Still, even remote regions may contribute sufficient information, particularly if they are frequently part of the 'far field' footprints of several tower sites. Since size and position of the tower footprints vary with the wind regime, the areas that are well sampled by a network of towers change with the seasons.

To understand the quality of the existing observing system in the Arctic to capture important elements of change over the Arctic we performed a gap analysis with respect to the Arctic Ocean, the Arctic atmosphere and the high-latitude carbon-monitoring network. This report summarizes the findings.

2. Experiment designs and observations

2.1. Ocean experiments and observations (UHAM)

2.1.1 ATL model simulations

This study relies on four Atlantic and Arctic model simulations with varying resolution. Both daily output and monthly output are used in the analysis. The main characteristics of the model simulations are listed in Table 1.

We use results from four integrations of the MIT general circulation model (Marshall et al., 1997) covering the entire Arctic Ocean north of the Bering Strait and Atlantic Ocean north of 33°S with different resolutions. In all cases, the model uses a curvilinear grid with two poles located over North American and Europe; these model simulations are called ATL03, ATL06, ATL12, and ATL24 with a horizontal resolution of about 32, 16, 8 and 4 km, respectively. All the simulations use z-coordinates; ATL03, ATL06, and ATL12 have 50 vertical levels with a resolution ranging from 10 m in the surface to 456 m in the deep ocean, while ATL24 has 100 vertical levels with resolutions varying from 5 m in the upper ocean to 185 m towards the bottom. Bottom topography is derived from the ETOPO 2-min database. ATL03, ATL06, and ATL12 are initialized with annual mean temperature and salinity from the World Ocean Atlas 2005 (Boyer et al., 2005), while ATL24 starts from initial conditions from the year 2002 of ATL12.

At the ocean surface, the model simulations are forced by momentum, heat and fresh-water fluxes computed using bulk formulae and either the 1948-2016 6-hourly atmospheric RA1 (Kalnay et al., 1996) reanalysis (ATL03, ATL06, and ATL12) or the 2002-2012 6-hourly ECMWF ERA-Interim (Dee et al., 2011) reanalysis (ATL24). Virtual salt flux is used in all simulations. At the open boundaries, the model simulations are forced by the monthly output from a GECCO2 (Köhl, 2015) global model configuration. The river run-off is applied at river mouths with a seasonal climatology. A dynamic-thermodynamic sea ice model (Zhang & Rothrock, 2000) is employed to model the sea ice parameters. Evaluation of the ATL06 simulation regarding overflows through the Denmark Strait can be found in Serra et al. (2010). Comparison of the model simulations with freshwater content observations in the Arctic

Ocean can be found in Köhl and Serra (2014) and salinity variability in the Atlantic Ocean in Sena Martins et al. (2015). Sea-level variability in the Arctic Ocean (Koldunov et al., 2014) and sea-level spectral content in the Atlantic Ocean (Biri et al. 2016) have also been studied with these simulations.

Table 1. Summary of model forward runs used in the present study.

Model Run	Horizontal Resolution	Vertical Grid	Daily Data	Monthly Data
ATL03	1/3, ~32 km	z-coordinate, 50 levels		1948-2009
ATL06	1/6, ~16 km	z-coordinate, 50 levels	04.01.1990-08.12.2002	1948-2009
ATL12	1/12, ~8 km	z-coordinate, 50 levels	05.01.2003-01.12.2010	1948-2016
ATL24	1/24, ~4 km	z-coordinate, 100 levels	01.01.2003-23.08.2012	2003-2011

2.1.2 The ocean observing systems

In this study, both hourly and monthly tide gauge records are used (see **Figure 1** for locations). The monthly tide gauge records are derived from the Permanent Service for Mean Sea Level (PSMSL Woodworth & Player, 2003). Tide gauge stations located near the mouth of rivers are rejected since mass changes caused by the discharge of river runoff can lead to sea level variability of 1 m, which cannot be simulated by the model with a virtual salt flux parameterization. 69 stations are used for validation, and only valid values are used for computing root mean square errors (RMSE), correlations and standard deviations. The tide gauge data are compared against model simulations on the closest model grid point. Only the tide gauge with the most extended period and least gaps are retained if several tide stations are mapped onto the same model grid location. The invert barometer (IB) effect is removed from the tide gauge data before comparing against the model simulations.

For comparing the high-frequency variability, six tide gauge records with hourly observation frequency are used, supplied by the University of Hawaii Sea Level Center (UHSLC). Three high-frequency bottom pressure records deployed in the Beaufort Gyre, Mooring A and Mooring B and from near the North Pole (Morison et al., 2007) are also used. Tidal signals in the high frequency observations are removed using the T_TIDE Matlab program (Pawlowicz et al. 2002). Since we intended to compare the high-frequency variability of the model simulation with those high-frequency observations, the de-tided data are then filtered using a high-pass filter with a cut-off period of 10 months.

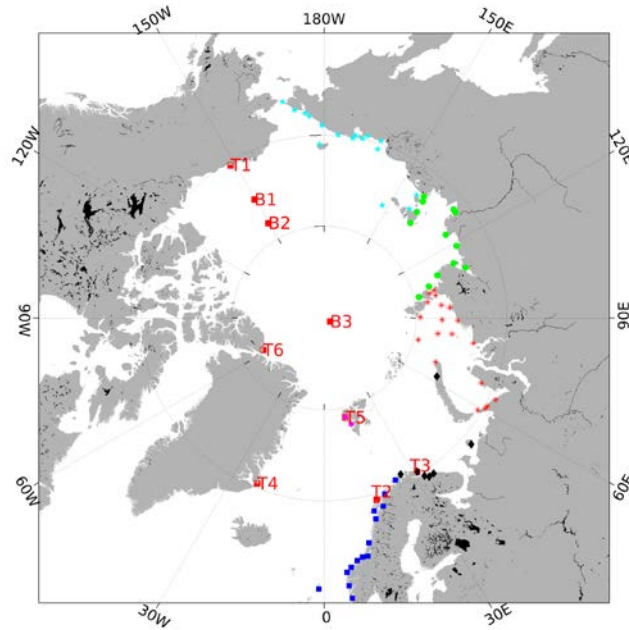


Figure 1. Map of the studied region presenting the locations of time series observations analyzed: from high-frequency tide gauges and moorings (red squares) and monthly tide gauges (blue square, red asterisks, green dots, and cyan dots).

Since the altimeter observational system and mooring systems are the most persistent observing systems, we evaluate the potential effect of the two systems on model parameters using adjoint sensitivity. The red dots and red pentagrams in Figure 2 mark the location of moorings that were deployed in the Arctic Ocean. The mooring system in Fram Strait (Fieig et al., 2010) has been deployed by AWI since 1997 and is still operational, but nowadays the system is shifted to the Greenland side (de Steur et al., 2009) for monitoring outflow from the Arctic. The Bering Strait system is operated by University of Washington (Woodgate et al., 2012) and has been monitoring the transport since 1998. The Davis Strait mooring system was deployed to monitor transport through the Davis strait from 2004 to 2005 (Curry et al., 2011). A series of moorings is also deployed in the Barents Sea Opening to observe the exchange between the Atlantic Ocean and the Barents Sea (Skagseth et al., 2008). In the Arctic Ocean, we include six moorings: the two moorings in the Eurasian basin are used to observe Atlantic inflow between 2002-2005 (Dmitrenko et al., 2008), the four mooring in Beaufort Gyre are used to monitoring freshwater content changes since 2003 to now. All the moorings observe salinity, temperature, and velocity, from which we can estimate the freshwater transport and heat transport.

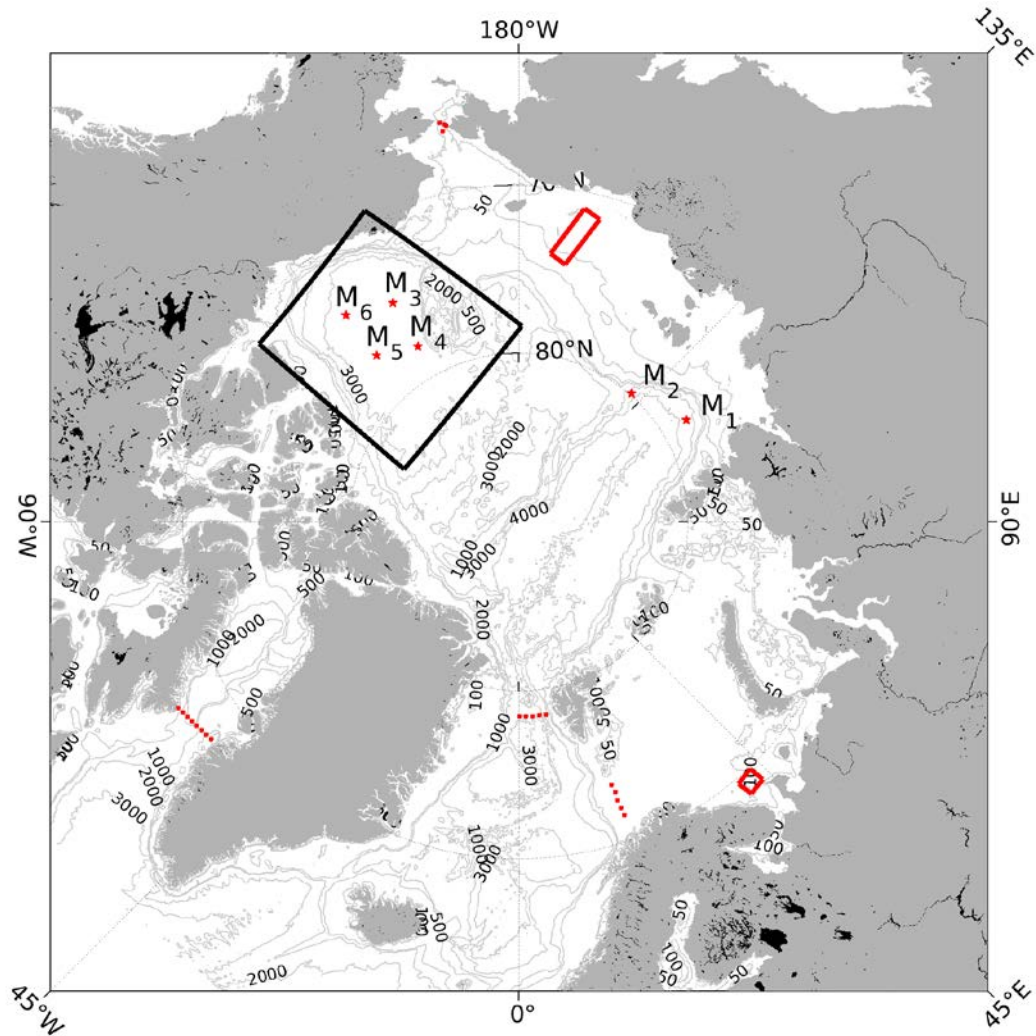


Figure 2. Map of the studied region presenting the locations of the mooring observing system deployed in the Barents Sea opening, the Fram Strait, the Bering Strait, and Davis Strait (red dots). The red pentagrams show locations of single moorings (M1~M6). The black box and red boxes define regions that we compute mean sea-level in the adjoint simulations.

Satellite altimeter system can continuously provide the sea level information over large areas of the Arctic Ocean. Figure 3 displays the observation frequency of each 1×1 box by a past altimeter system (ERS-2 satellite) for winter time (a) and summer time (b), and by a current altimeter system (Sentinel-3B, CFOsat, and Swot) for winter time (c) and summer time (d).

For the past altimeter system, only the Greenland, Iceland and Norwegian (GIN) seas and the Barents Sea were observed. The observation frequency is around $0.5-1 \text{ day}^{-1}$ (each point can be observed every 1-2 days) in the summer time and around 0.1 in the winter time. Current altimeter radar systems improve spatial coverage and can monitor nearly all marginal seas. However, the observation frequency does not improve. The central Arctic Ocean is not observed for both the systems because in the presence of sea ice, conventional data processing fails and specialized data processing is required to extract sea level information. Armitage et al. (2016) applied an algorithm to derive monthly sea surface height up to latitudes of 81°N . To evaluate the effect of the altimeter observing system on model parameters, we performed three adjoint model integrations. The first two experiments used area-averaged sea level data in the Barents Sea and the East Siberian Sea (red boxes in Fig-

ure 2). The model is integrated for 15 days, and the mean sea level is averaged over the 15th day. With these two experiments, we intent to identify high-frequency signals and causes. The third experiments used area-averaged sea levels in the Beaufort Gyre (black box in Figure 2 with water depth larger than 500 m). The model is integrated for four years, and the sea level is averaged for the last two years. We use the sensitivity to investigate the effect of sea level observation on salinity.

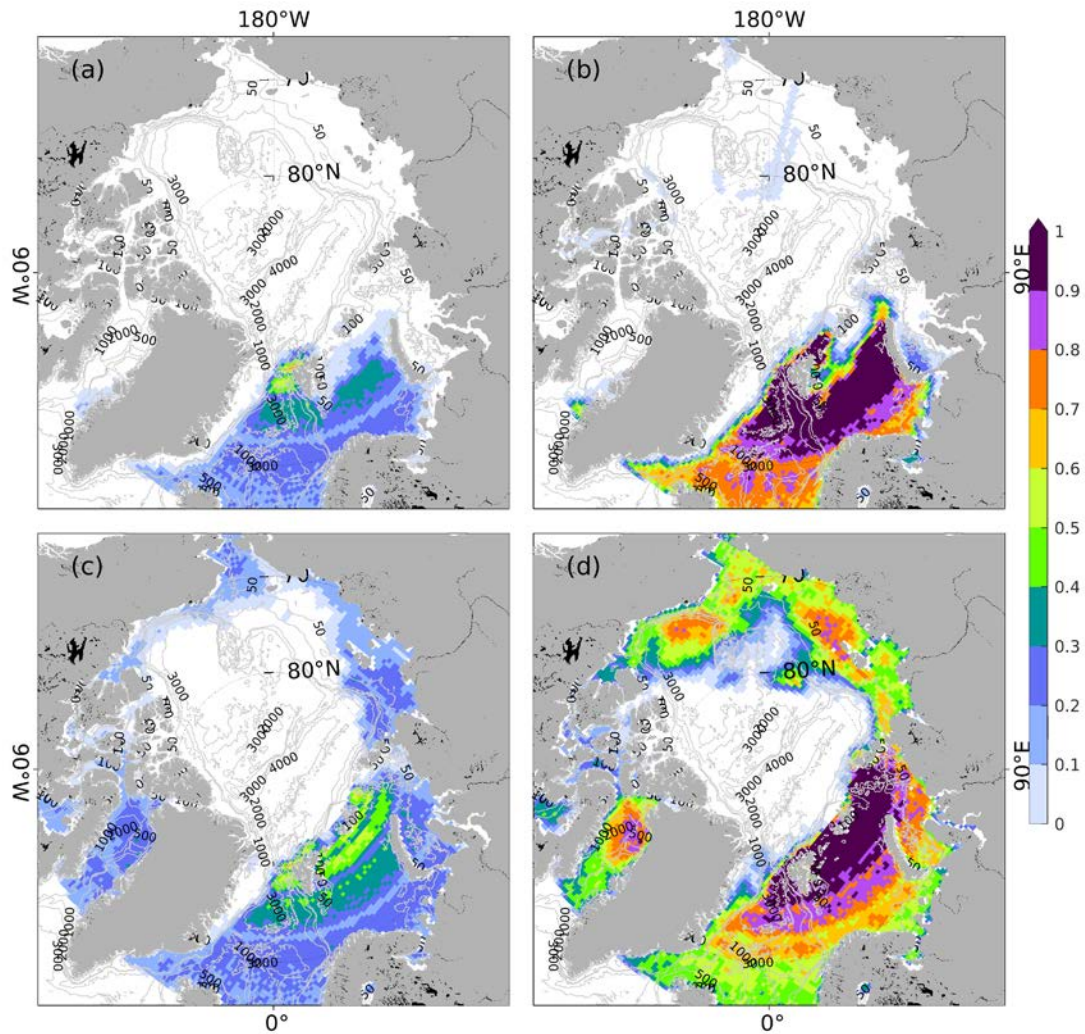


Figure 3. Altimeter observation frequency ($1/T$, day^{-1}) in the past (a,b) and current (c,d). The left column is for winter time (DJF), and the right column is for summer time (JJA). The observation frequency is mapped on 1×1 degree.

2.1.3 Relation of sea level with mass and steric contribution

The altimetry system provides persistent observations over the Arctic, especially in the marginal seas. Sea surface height is an integral indicator which reflects changing ocean conditions due to ocean dynamics, atmosphere forcing and terrestrial process (Stammer et al., 2013). Sea surface height changes can be used to monitor changes of other model state, such as circulation and freshwater content (Armitage et al., 2016). Vice Versa, the altimetry observing system may also be complemented by observing related parameters. Therefore, it

is crucial to understand causes of sea level variability, and its relation to mass component and steric component.

Following Ponte (1999) and Calafat et al. (2013), sea surface height changes can be separated into contributions from density changes (steric effect), inverse barometer (IB) effect and mass effect:

$$\eta' = -\frac{1}{\rho_0} \int_{-H}^0 \rho' dz + \frac{1}{\rho_0 g} (\bar{P}'_a - P'_a) + \frac{1}{\rho_0 g} (P'_b - \bar{P}'_a) \quad (1)$$

where g is the gravitational acceleration and η' represents sea surface height anomaly. The first term on the right-hand side represents steric effects, with ρ' being the density anomaly (ρ_0 is the reference density of 1025 kg/m³). The second term is the IB effect. \bar{P}'_a and P'_a represent spatial air pressure anomalies over the global oceans and local air pressure anomalies, respectively. The last term defines the mass effect, related to bottom pressure.

The steric effect can be further decomposed into thermosteric (due to temperature anomalies) and halosteric (due to salinity anomalies) effects:

$$\eta'_{ts} = \int_{-H}^0 \alpha T' dz \quad (2)$$

$$\eta'_{hs} = - \int_{-H}^0 \beta S' dz \quad (3)$$

where α and β represent the thermal expansion and saline contraction coefficients, respectively, and T' and S' are potential temperature and salinity anomalies.

2.1.4 The adjoint method

Denote a numerical model operator by M . For simplicity, we assume that the model variables X_n , at timestep n , only depends on initial conditions X_0 , and both have a size of $L \times 1$. Then, a target function J can be defined as:

$$J = F(X_n) = F(M_n \cdot M_{n-1} \cdots M_0 \cdot X_0) \quad (4)$$

where F maps X_n onto a scalar value.

Based on tangent linear approximation, changes of the target function by perturbing the initial condition can be written as

$$\frac{\partial J}{\partial X_0} = 1 \cdot \frac{\partial F}{\partial X_n} \cdot \frac{\partial X_n}{\partial X_0} \cdot I = \frac{\partial F}{\partial X_n} \cdot M'_n \cdot M'_{n-1} \cdots M'_0 \cdot I \quad (5)$$

where M' represents the tangent linear model operator and I is the unit matrix, both with a size of $L \times L$. The sizes of $\frac{\partial F}{\partial X_n}$ and equation (5) is $1 \times L$. Practically, we don't compute the matrix $M'_{0 \rightarrow t}$ explicitly. We use a tangent linear model, differentiated from the nonlinear model, to compute the error propagation. Therefore, we need to integrate the tangent linear model L times (each integration for one column of I) to evaluate the sensitivity in equation (5).

By taking the transpose of equation (5), we get the adjoint sensitivity as

$$\left(\frac{\partial J}{\partial X_0} \right)^T = M_0'^T \cdots M_{n-1}'^T \cdot M_n'^T \cdot \left(\frac{\partial F}{\partial X_n} \right)^T \cdot 1 \quad (6)$$

where superscript T means transpose and M'^T is the adjoint model operator. Based on Equation (6), the sensitivity of target function J with respect to initial condition X_0 is computed by integrating the adjoint model backward (from n to 0) once. The sensitivity can be explained as: by perturbing the initial condition X_0 with sensitivity from equation (6) and integrating the model from 0 to n, we will increase the target function by 1.

The adjoint sensitivity determines parameter perturbations (e.g., regions of forcing perturbations or initial conditions perturbations) that most efficiently change the given target. Therefore, we can use the sensitivity to explore the potential effect of existing ocean observing system on model variables in upstream regions.

With MIT general circulation model and its adjoint, we performed a suite of adjoint model simulations to investigate the potential effects of existing observing system on monitoring the sea-level variability and on the circulation. Based on the moorings system in Figure 2, we performed two adjoint model runs to evaluate the effect of freshwater transport and heat transport on the model state. The freshwater transport J_f and heat transport J_h are as follows:

$$J_f = \frac{1}{T} \int \langle \iint \left[-v \frac{S-S_{ref}}{S_{ref}} dx dz \right]_{f_s, bso, davis, bs} + \left[-v \frac{S-S_{ref}}{S_{ref}} dx dz - u \frac{S-S_{ref}}{S_{ref}} dy dz \right]_{M1 \sim M6} \rangle dt \quad (7)$$

$$J_h = \frac{1}{T} \int \langle \iint [vT dx dz]_{f_s, bso, davis, bs} + [vT dx dz + uT dy dz]_{M1 \sim M6} \rangle dt \quad (8)$$

where S_{ref} equals 34.8. For the mooring systems in Fram Strait, Barents Sea opening, the Davis Strait and Bering Strait, the transport is defined as the transport through the straits. Moreover, the transport monitored by the single moorings (M1~M6) is the sum of U and V transport components. In each experiment, the adjoint model is integrated for four years, and the target function is computed as the four years mean transport.

As for sea-level, we also performed a suite of adjoint simulation to identify causes of the sea level variability and demonstrate potential upstream regions for the Beaufort Gyre. The target function is defined as region-averaged sea level:

$$J = \frac{\frac{1}{T} \iiint SSH dx dy dt}{\iint dx dy} \quad (9)$$

where T is the average period, and J is the mean region-averaged sea level over defined time. The target function is defined as region-averaged sea level in the Barents Sea and East Siberian Sea (red box in Figure 2) in two experiments and in the Beaufort Gyre (black box in Figure 2) in the third experiments. The results are analyzed in Section 3.1.4.

2.2. Atmospheric experiments and observations (FMI)

To evaluate the radiosonde observation network, we compared operational analyses and 12 h forecasts of the ECMWF atmospheric model, from times 00 UTC and 12 UTC, and radiosonde observations from the same hours. These radiosonde observations had been transmitted to the GTS (Global Telecommunication System) network and stored in the Integrated Global Radiosonde (IGRA) archive. The study period was from January 2016 to September 2018, except that a shorter period, from January 2016 to December 2017, was addressed applying trajectory analyses. The study region was the circumpolar Arctic north of 60°N. Analyses and forecast fields were interpolated to a 0.25×0.25 degree grid. In comparisons

with the sounding stations, the values of analyses and forecasts were averages of a 1×1 degree box around each sounding station.

2.2.1 Sounding data

Radiosonde observations were taken from the IGRA archive, which is a comprehensive, freely available global dataset of quality-assured radiosonde observations (Durre et al. 2006). Figure 4 shows the locations and the WMO number of the sounding stations north of 60°N . The temporal coverage of the radiosonde observations was mostly sufficient, except at Egilsstadir (station 04089) in Iceland, where most of the soundings were missing, and at Coral Harbour (station 71915) in the Canadian archipelago, which did not report soundings after August 2017. Shorter breaks in sounding time series occurred at several stations. Furthermore, at Ny Ålesund (station 01001) soundings were performed only at 12 UTC and at Luleå Kallax (station 02185) only at 00 UTC. Radiosonde types varied between the stations.

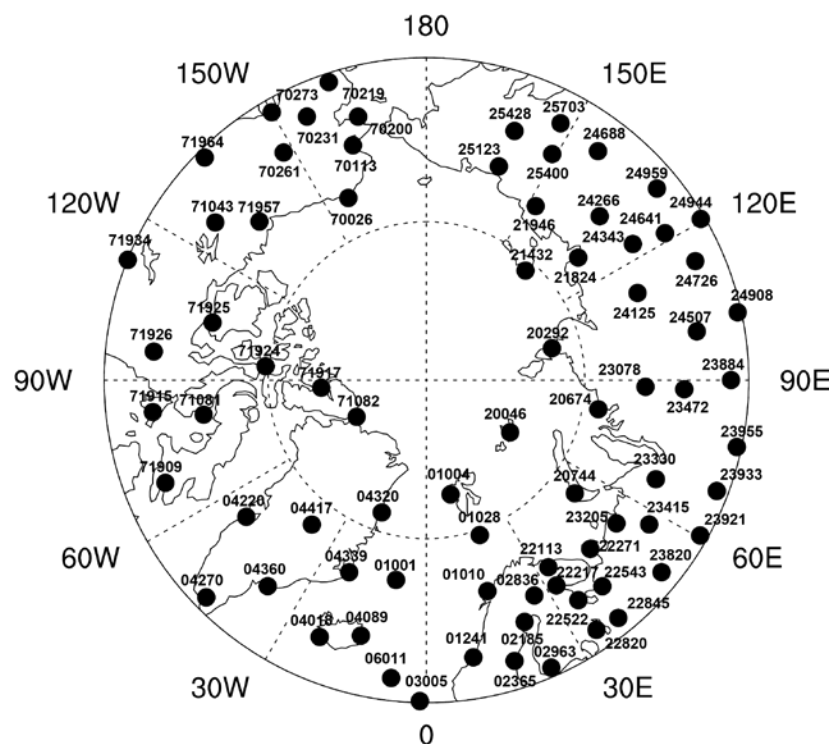


Figure 4. Sounding stations located north of 60°N .

2.2.2 ECMWF data assimilation

When comparing the analyses, forecast, and observations, it is essential to understand the main principles of the ECMWF operational data assimilation system. The ECMWF forecasting system uses 4-D variational data assimilation to produce analyses. The operational data assimilation consists of two procedures: 1) 4-D variational data assimilation with a long (12 h) window, and 2) 4-D variational data assimilation with a short (6 h) window. The long-window data assimilation provides initial conditions and background fields both for the next long-window data assimilation and for the next short-window assimilation, and is, thus, only used in assimilation procedures. The short-window assimilation, in turn, produces operational analyses which are used as the initial conditions for the actual forecast. These assimilation

lation cycles are repeated twice a day. It is important to note that the operational forecasts are not background fields for the next analysis cycle and, thus, the difference between 12 h forecast valid for the analysis time and analysis is not the same as the analysis increment (which is defined as the difference between the analysis and the background field based on the long window data assimilation). Compared to 12 h forecast, the background field includes observations from a longer window and it also includes the impact of observations which have been delayed from the cut-off of the short-window assimilation.

2.2.3 Trajectory analyses

To identify the geographical areas where new sounding stations would potentially improve analyses, we analyzed backward air mass trajectories at each sounding station. Trajectories were calculated 12 h (and 24 h) backwards from the locations of sounding stations to identify the air mass origin. Trajectory calculations were made using the Hysplit programme (Stein et al. 2015), and meteorological conditions for Hysplit were taken from the ERA-Interim reanalysis. Trajectories for each sounding station were calculated approximately at 850 hPa level.

Starting points of trajectories, i.e., points from where air mass had been advected to a sounding station, were organized into a regular grid. The resolution of this grid was 5° in the meridional direction. In the zonal direction, the resolution was 10° in the areas south of 65N, but it gradually decreased towards the pole to keep the area covered by a grid cell nearly constant. A single grid cell might contain starting points of trajectories ending up to several sounding stations. When statistics related to the starting points were calculated, the difference (between a forecast and a sounding) at the sounding station (where the air mass ended up) was identified and placed to the grid point of the air mass origin. From these values, averages and other statistics for a certain air mass origin could be calculated.

The trajectories revealed how different air mass origins affected differences between the soundings and the 12 h forecasts for the sounding stations. The method used is based on assumptions that errors in the forecast are, at least partly, being advected to the sounding station along the air mass, and errors in short forecast are more related to errors in initial conditions than in model physics. The hypothesis is that large differences between soundings and forecasts occur when the air mass is advected from an area where differences between soundings and initial conditions of forecasts are large.

2.2.4 SOM analyses

We further investigated how differences between soundings and forecasts vary related to synoptic-scale circulation patterns. The motivation for this stems from previous studies, which have indicated that the impacts of radiosonde observations on model performance are flow-dependent (e.g., Inoue et al. 2015). To investigate the connections to synoptic-scale atmospheric circulation patterns, we applied the Self-Organized Map (SOM) method. The SOM method has been developed by Kohonen (2001), and it can be characterized as a non-linear mapping of a high-dimensional input data onto a two-dimensional array of reference vectors, i.e., nodes. Using the SOM method, we generalized the atmospheric circulation patterns at each time step to a small number of states. The SOM method was applied to the mean sea-level pressure fields of ECMWF operational analyses and as a result, a 4 × 5 arrays (20 nodes) of characteristic atmospheric circulation patterns was obtained. Statistics on differences between forecasts and radiosonde observations could then be calculated and analyzed for each atmospheric circulation regime separately.

2.3. Evaluating the pan-Arctic atmospheric greenhouse gas monitoring network (MPG)

To determine whether the existing atmospheric tall tower network collects sufficient information to reliably quantify carbon exchange processes within a chosen target region, we conducted an integrated footprint analyses for the entire monitoring network in the region, covering the entire annual course to also account for seasonal variability of the atmospheric transport processes.

Atmospheric transport was computed using the setup presented in Henderson et al. (2015). This model is based on the 'Stochastic Time-Inverted Lagrangian Transport model STILT (Michalak et al., 2004) coupled to the mesoscale model WRF (Skamarock et al., 2008) driven by MERRA reanalyses (Rienecker et al., 2011). The domain setup (Figure 5) followed a 3-way nesting structure with 41 vertical levels introduced by Henderson et al. (2015), with horizontal grid resolutions ranging from 3.3 km in the innermost domain focusing on Alaska to 30 km for the full domain. For each observation, trajectories of 500 particles were computed backwards in time over a period 15 days. Simulations cover the period July 2014 to December 2015, with daily particle release times restricted to the early afternoon (2-4 pm local time), i.e., focusing on times when the probability for well-mixed boundary layer conditions is highest. The surface influence of a particle was calculated as the time spent in the lower half of the atmospheric boundary layer. From these back-trajectories, which were provided by John Henderson (AER), the surface influence based on all particles ("footprint") was calculated.

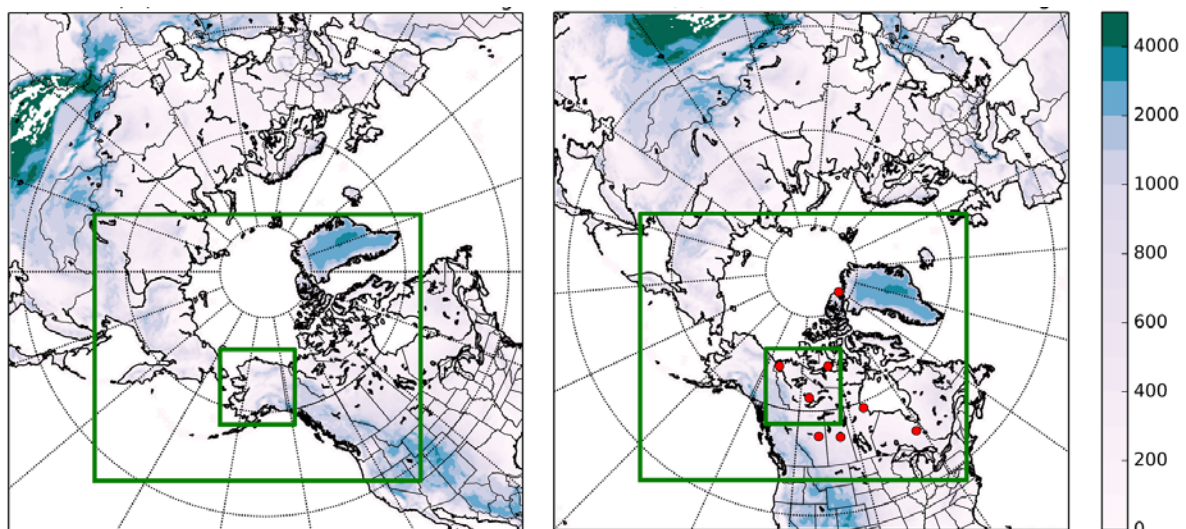


Figure 5. Structure of the 3-way nested domains used for the atmospheric transport modeling, with background colors indicating elevation above sea level [m]. The full domain has a horizontal resolution of 30 km, the 2nd domain (big green box) uses 10 km, while the innermost domain (small green box) has 3.3 km grid spacing. (left) Alaska domain (Figure taken from Henderson et al., 2015) used for the Alaskan and East Siberian sites (AMB, CHS), and (right) Canadian domain used for the remaining sites (John Henderson, personal communication).

Selection criteria for monitoring sites to be included into this network evaluation study were that (i) the site provides continuous observations (i.e., year-round measurements at averaging intervals of 30-60 minutes) of well-calibrated atmospheric greenhouse gas mixing ratios, and (ii) the site is located north of 50°N. A threshold that far south of the Arctic was selected since on the one hand sites in the boreal zone are important to define and/or evaluate the boundary conditions for Arctic domains, and on the other hand, their footprints

often extend far into the Arctic domain. The latter constraint was loosened once for Fraserdale (located at 49.9°N), which filled an important gap in Eastern Canada. Conducting a survey among site PIs and online data repositories, we identified a total of 29 observation sites that fulfilled these criteria, with details on site locations given in Table 2 and Figure 6. Site distribution can be split into four major domains:

- **Alaska:** This domain features only two towers with continuous measurements, but one of them (Barrow) provides one of the longest data records in the Arctic (since 1971).
- **Canada:** All nine sites included herein for the Canadian domain belong to a regional network operated by Environment Canada (network PI: Douglas Worthy). While most of these sites were established about a decade ago (2009-2014), the longest running tower (Alert) features a 32-year observational record.
- **European Arctic:** Within this domain, we picked six sites that provide the most relevant data for Arctic monitoring, either because of location (high latitude) or the length of the data record. Site operators include, e.g., the Finnish Meteorological Institute (FMI), the Norwegian Institute for Air Research (NILU) or the ICOS network. For this domain, in principle more sites would be available, but they are located quite close to the ones selected here, and thus would not add much information at pan-Arctic scales.
- **Russia:** The 12 Russian sites can be split into two clusters, one focusing on central West Siberia, the other stretching along the Arctic Ocean coast. The western cluster (JR-Network) is operated mainly by the Japanese National Institute for Environmental Studies (NIES) in collaboration with Russian colleagues, while the other sites have different operators, including the Russian Arctic and Antarctic Research Institute (AARI), FMI, and the German Max Planck Society. One of the Russian sites (DIK) has only been established recently in summer 2018.



Figure 6. Locations of the 29 atmospheric monitoring towers that form the pan-Arctic observation network analyzed within this study. Site selection was restricted to infrastructure >50°N (plus Fraserdale, at 49.9°N) that provided continuous records of atmospheric greenhouse gas mixing ratios within the target study period (map: GoogleEarth).

More detailed information on all of the included measurement sites is provided through an online mapping tool that MPI-BGC developed, together with international colleagues, as an INTAROS service to the Arctic research community (<https://mpi-bgc-ipas.shinyapps.io/Arctic-GHG-tool/>). Where available, this tool provides direct links to the online repositories where data can be accessed. The map also covers other observational

networks and atmospheric monitoring stations that do not fulfill the search requirements for the presented study (e.g., flask sites that provide discontinuous measurements).

Table 2: List of the 29 atmospheric monitoring towers that form the pan-Arctic observation network analyzed within this study.

Domain	Site code	Site	Latitude [°]	Longitude [°]	Elevation [m a.s.l.]	Tower height [m a.g.l.]
Alaska	CRV	CARVE Tower Fairbanks	64.99	-147.60	611	32
	BRW	Barrow	71.32	-156.61	11	16.46
Canada	FSD	Fraserdale	49.88	-81.57	210	40
	ETL	East Trout Lake	54.35	-104.99	493	105
	LLB	Lac La Biche	54.95	-112.47	540	NA
	CHL	Churchill	58.74	-93.82	29	60
	FNE	Fort Nelson	58.84	-122.57	361	15
	BCK	Behchoko	62.80	-115.92	160	60
	INK	Inuvik	68.32	-133.53	113	10
	CBY	Cambridge Bay	69.13	-105.06	35	12
	ALT	Alert	82.45	-62.51	200	10
Europe	SVB	Svartberget	64.26	19.77	267	150
	SOD	Sodankylä	67.36	26.64	179	48
	PAL	Pallas	67.97	24.12	560	20
	KJN	Kjolnes	70.85	29.24	1	4
	ZEP	Zeppelin	78.90	11.88	474	16
	SN1	Station Nord	81.60	-16.66	24	80
Russia	VGN	Vaganovo	54.50	62.32	192	85
	AZV	Azovo	54.71	73.03	110	50
	BRZ	Berezorechka	56.15	84.33	168	80
	KRS	Karasevoe	58.25	82.42	76	67
	DEM	Demyanskoe	59.79	70.87	63	63
	ZOT	Zotino	60.80	89.35	114	301
	NOY	Noyabrsk	63.43	75.78	108	43
	CHS	Pleistocene park	68.51	161.53	10	34
	AMB	Ambarchik	69.62	162.30	20	27
	TIK	Tiksi	71.60	128.89	19	10
	DIK	Dikson	73.50	80.40	15	20
	BAR	Baranov	79.28	101.62	20	10

To improve result interpretation and visualization, cumulative footprint results were aggregated to ecoregion level for some of the plots shown in Section 3.3 below. Separation of the model domain into these ecoregions followed the assignment of marine ecoregions of the world (MEOW, Spalding et al., 2007) and the largest terrestrial ecoregions of the world (TEOW, Olson, et al., 2001). A combined version of both maps, aggregated to 32 km resolution, is shown in Figure 7 below.

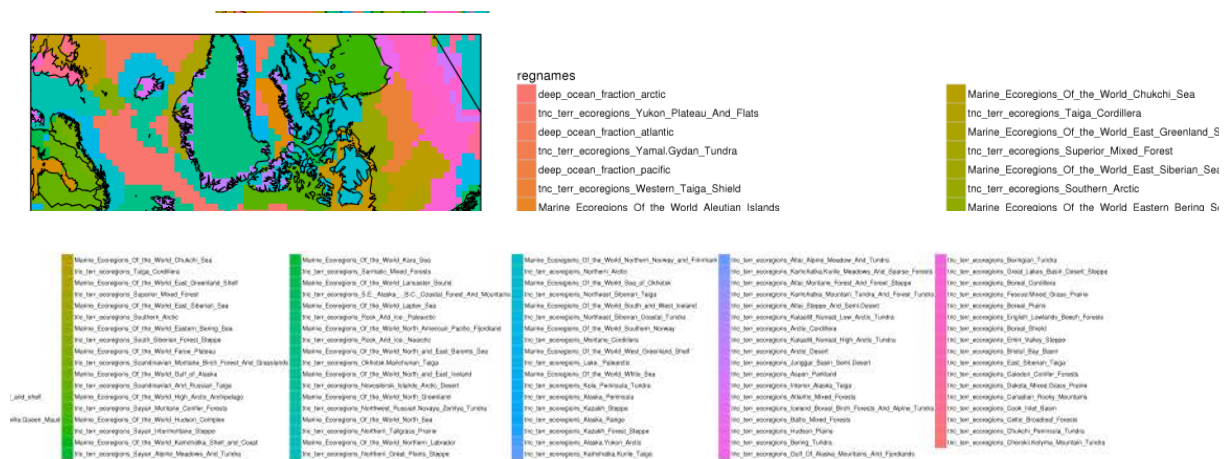


Figure 7. Combined map of marine and terrestrial ecoregions, aggregated to 32 km grids, within the core area of the pan-Arctic study domain.

3. Results

3.1. Evaluation of the ocean observing system (UHAM)

3.1.1 Statistics of Sea Surface Height and Bottom Pressure

Figure 8 shows two stations observations for the year 2005: the tide station Ny-Ålesund in Svalbard and Mooring B1 in 3824 m of water in the Beaufort Gyre. As the top panel shows, the tidal signals in Ny-Ålesund are much larger than the residual signals while it is comparable to the residual signal in Mooring B1. The de-tided signal, the IB effect, and two model simulations are shown in the middle panel. In the tide station Ny-Ålesund, the de-tided signal is dominated by the IB effect. Removing the IB effect reduced the variability significantly and compared better with the model simulations. However, the observed variability is still larger by 25% than the model simulations (Figure 8). As for Mooring B1 (Figure 8e), the IB effect shows much larger variability than the bottom pressure perturbation, and there seems to be no relation between them. It is likely that in deep water regions the IB effect is compensated by the sea-level anomaly as several studies have assumed (e.g., Calafat et al., 2013; Proshutinsky et al., 2007; Wunsch & Stammer, 1997). Therefore, we did not remove the IB effect from the three bottom pressure records. Similar to the tide gauge at Ny-Ålesund, the bottom pressure perturbations still show larger variability than the model simulation. There exists a high-frequency variability with a timescale of several days. As the bottom panels show, the model simulations and the observations show significant coherence in a broad spectral band, ranging from a couple of days to the seasonal cycle, although there are some low coherence spikes, especially in Mooring B1 observations.

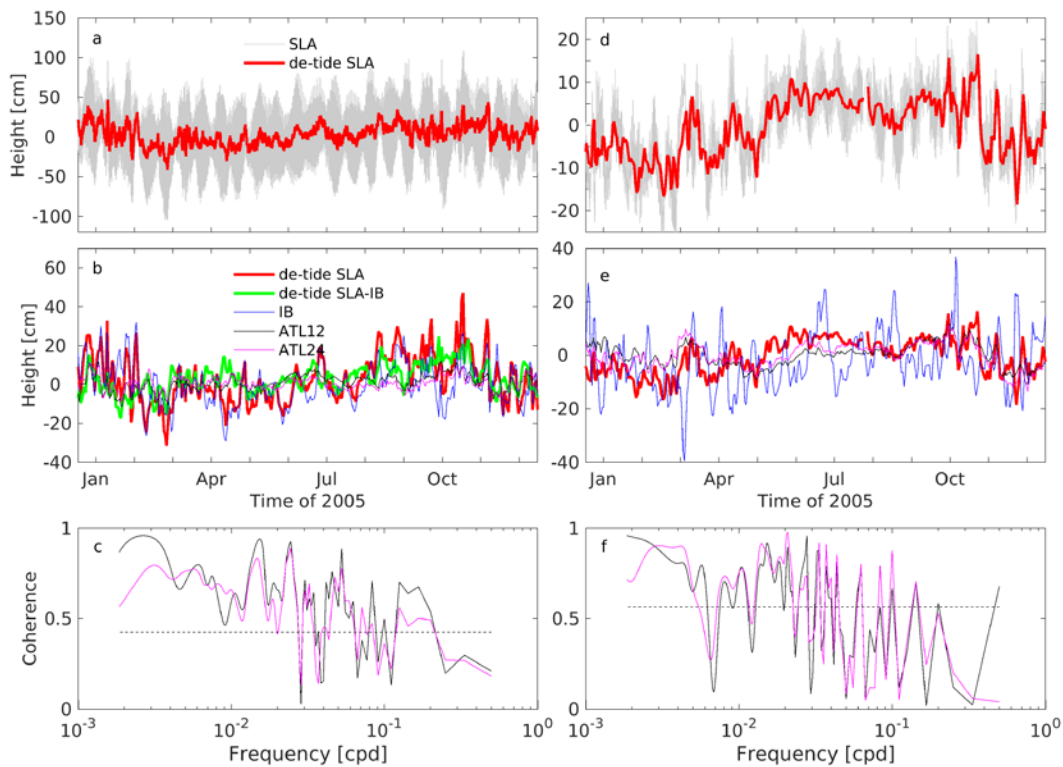


Figure 8. (a) Ny-Ålesund tide station observations and de-tided signal. (b) de-tided signal, invert barometer (IB) effect, the de-tided data with IB effect removed, and corresponding sea level signal simulated in ATL12 and ATL24. (c) The coherence of ATL12 and ATL24 simulated sea level with de-tided SLA without IB effect. The black dashed line indicates the 95% significance level. Panels (d)-(f) are the same as panels (a)-(c) except that they are based on the mooring B1 bottom pressure station.

The Taylor diagram (Taylor, 2001) is used to summarize the statistics (correlation and normalized standard deviation) of the model simulations, compared with the tide gauge observations and bottom pressure records. For high-frequency variability, as Figure 9a shows, there is a tendency that higher resolution simulation leads to both higher variability and higher correlation. However, the models can only simulate comparable high-frequency variability in Prudhoe Bay (T1) and Alert (T6). For the other stations, the model-simulated variability is smaller by 25-50%, compared to the observed variability. Fig. 9b displays statistics for the monthly data. In the Norwegian and Barents Seas, the models perform better than in the other regions, with ATL24 overwhelming the other simulations. In the East Siberian Sea, ATL24 simulates comparable variability with the observed variability, while its correlation is smaller than the other three simulations. In Svalbard, Laptev and Kara Seas, the four model simulations perform similarly, with the variability being smaller than the observed variability by 30-60%.

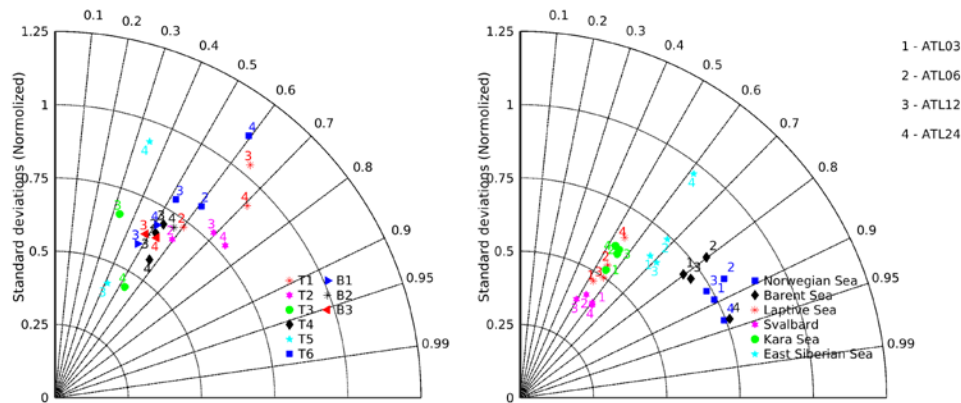


Figure 9. Taylor diagrams of the model simulations for the high frequency (left) and monthly (right) data. Comparing monthly observations with the model simulations is achieved based on six different geographical locations (see the legend in the right panel).

It is not clear what causes the different performance of model simulations in different marginal seas. However, by comparing sea surface height spectra from the model simulations and altimeter data, Biri et al. (2016) find that the higher resolution model (ATL24) performs better in reproducing the observed spectra at high frequencies and wavenumbers, especially in strong variability regions, such as in boundary currents. The presented results seem to confirm the conclusion, since there are strong coastal currents in the Norwegian, Barents and East Siberian Seas (Calafat et al., 2013). Figure 10 shows the seasonal climatology and interannual variability of the tide gauge and model simulations over the Norwegian Sea (a,b), the Barents Sea (c,d), the Kara and Laptev Seas (e,f), and the East Siberian Sea (g,h). In the Norwegian and the Barents Seas, the model simulations match well the tide gauge observations in both the seasonal and interannual variability. Large sea level anomalies occurring between the 1980s and 1990s are significant in both the observations and the model simulations. However, the observed tide gauge data in the Kara, Laptev, and East Siberian Seas show double peaks in the seasonal cycle, which the models seem unable to simulate. Only ATL24 seems to simulate double peaks in the East Siberian Sea, but with a one-month delay. As for the interannual variability, as Figure 10 (panels f and h) show, tide gauge observations have more considerable variability than all model simulations. Before the 1980s, the tide gauge data shows decadal variability and the model simulations follow the observations although the variability is small. Several factors may contribute to the discrepancy, such as river run-off and ice melting. Since we use virtual salt fluxes, the effects of these processes on the mass change of the sea water cannot be simulated in the model.

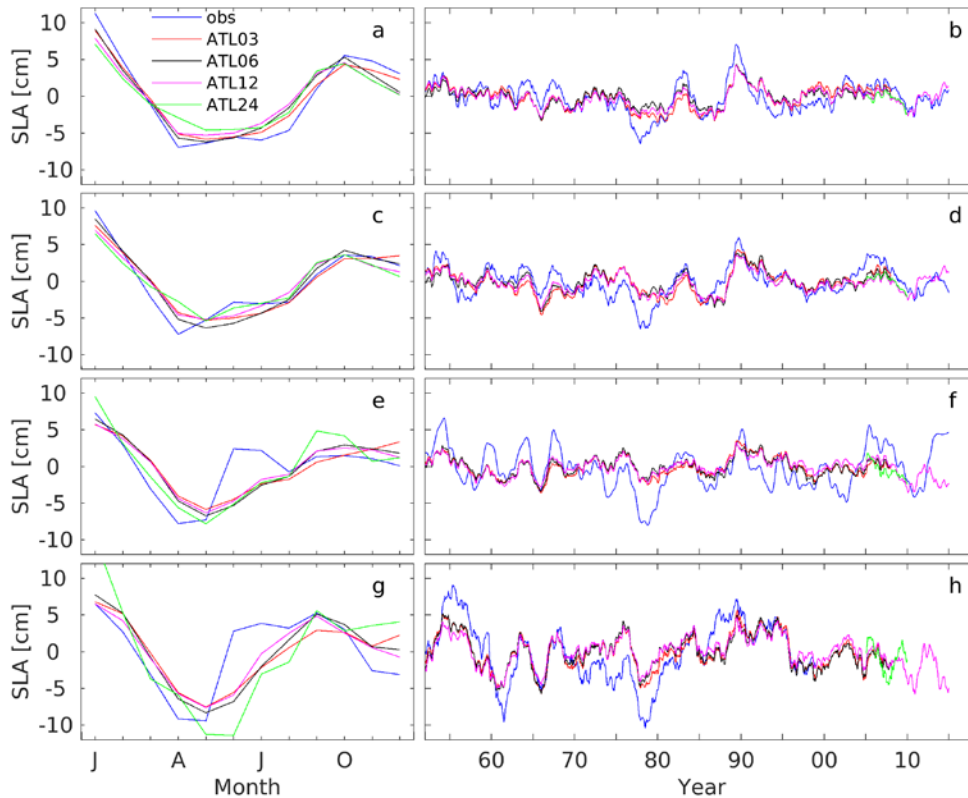


Figure 10. The mean seasonal cycle of tide gauges and model simulations averaged over the Norwegian Sea (a), the Barents Sea (c), the Kara and Laptev Seas (e), and the East Siberian Sea (g). The respective 12-month running averaged data is shown in (b), (d), (f) and (h).

3.1.2 Sea level variability

Figure 11 shows sea surface height variability for different frequency bands. In the Arctic Ocean, sea surface height shows significant variability in the marginal seas, the Alaskan coast, the Beaufort Gyre, and near the Lomonosov Ridge, where the Atlantic water spills into the Canadian basin. In the Norwegian Atlantic Current (NwAC), East Greenland Current (EGC), and West Greenland Current (WGC), sea surface height also shows significant variability. In the deep region around the North Pole, sea surface height has a root mean square variability around 3 cm.

Variability in the marginal seas is mainly at the high-frequency band (<10 months). NwAC, EGC, and WGC also show high variability for 2-10 months. The seasonal signal is also significant near in marginal seas, along with the coasts of the Norwegian Sea, the Barents Sea, and Greenland. At timescales 2-8 years (Figure 11e), considerable variability is found in the East Siberian Sea and along the NwAC. On decadal timescales, significant variability exists in the Greenland Sea, Norwegian Sea, Eurasian Basin and Canadian Basin (Figure 11f). In the deep-water region around the North Pole, the sea surface height shows both high- and low-frequency variability. Patterns of sea surface height variability from the other three model simulations are similar to Figure 11, but their amplitude is slightly different (not shown here).

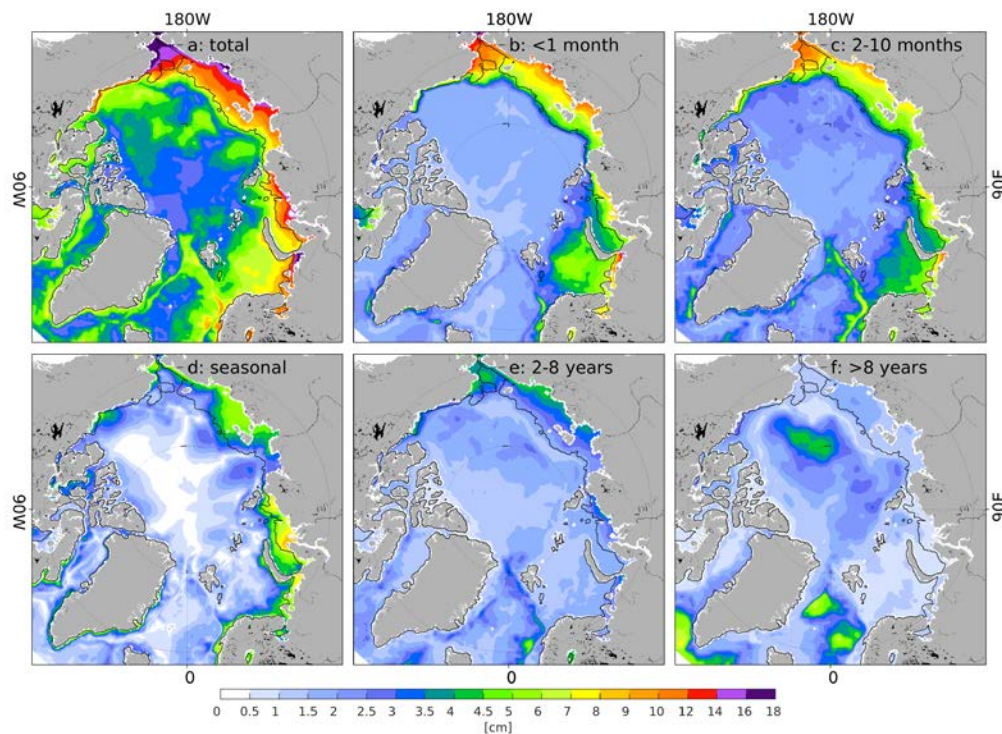


Figure 11. Root mean square variability of sea surface height in ATL12 for different frequency bands: total (a), <1 month (b), 2-10 months (c), annual (d), 2-8 years (e) and >8 years (f).

3.1.3 Relation of sea level variability to mass and steric contributions

As Equations. (1)-(3) show, sea surface height variability reflects changes in temperature, salinity, mass, and atmospheric pressure. We saw in Section 3.1.1 that the contribution of atmospheric loading could be enormous in some shallow regions. However, the model simulations do not include atmospheric loading, and therefore we mainly concentrate on the other two processes.

To identify the source of sea level variability depending on the source of sea water, we split the Arctic Ocean into three layers following Aagaard et al. (1985): the Arctic freshwater, the Arctic Atlantic layer and the bottom. We use the potential density anomaly of $\sigma_0=27.00 \text{ kg m}^{-3}$ to separate the upper layer (Arctic Water) and the intermediate layer (Atlantic Water), and the potential density anomaly of $\sigma_1=32.74 \text{ kg m}^{-3}$ to separate the intermediate layer from the bottom layer.

Figure 12a displays a slice of mean temperature across the Arctic Ocean with the two potential density isolines overlaid. They reasonably separate the Arctic Ocean into the three layers considering the temperature structure. Figure 12 (panels b and c) show the depth of those interfaces. The upper layer water exists in the entire Arctic Ocean and is thicker in the Canadian Basin than in the Eurasian Basin since freshwater is accumulated in the Canadian Basin due to Ekman pumping. The upper layer water is also transported around Greenland by the EGC and flows through the Canadian Arctic Archipelago. Throughout this study, we use the terminology “middle layer” or “intermediate layer” to represent the Atlantic water layer, although it is near the surface outside the Arctic Ocean. As Figure 12c shows, the Atlantic water sinks beneath the Arctic water layer after entering the Arctic Ocean through the Fram Strait. Some outflow west of the Fram Strait is also visible.

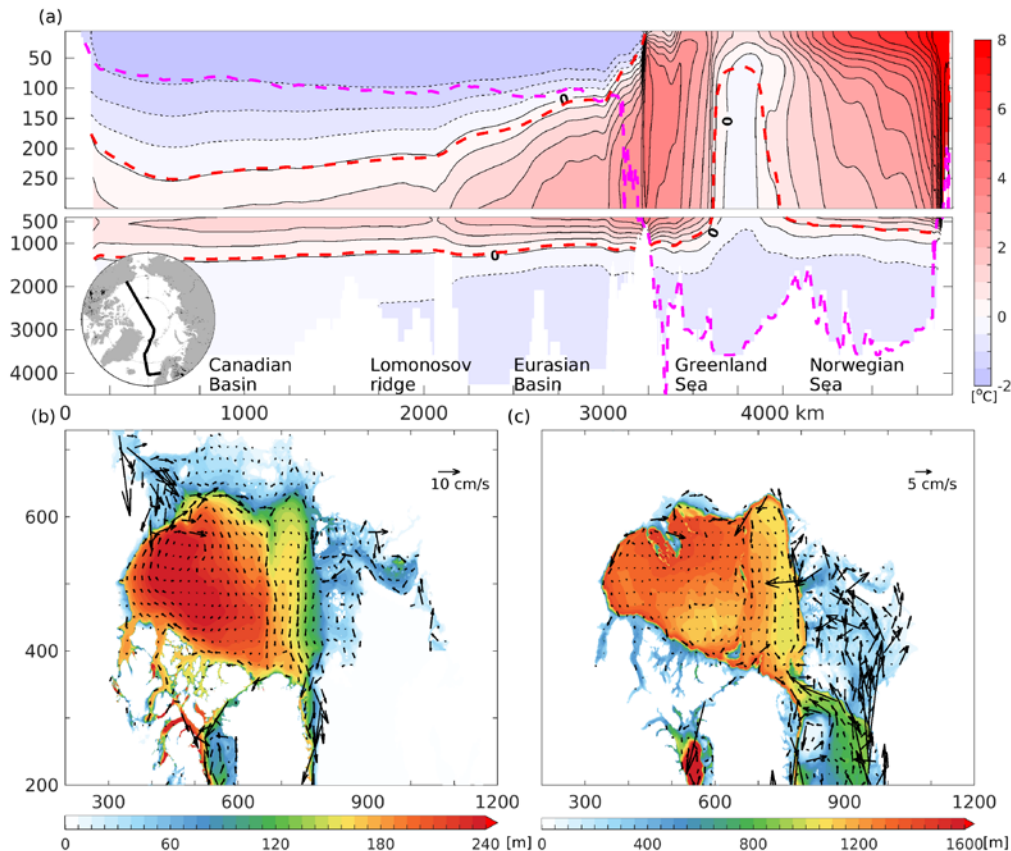


Figure 12. (a) Slice of mean temperature across the Arctic Ocean. The two solid red lines are the 0°C isotherm. The upper dashed red line corresponds to $\sigma_0=27.00 \text{ kg m}^{-3}$ and the bottom dashed red line to $\sigma_1=32.74 \text{ kg m}^{-3}$. (b) and (c) are the depth of the interfaces between the upper and intermediate, and the intermediate and bottom layers, respectively.

a. High-Frequency variability

Based on an ocean circulation model simulation, Vinogradova et al. (2007) showed that the ocean bottom pressure anomaly is barotropic at timescales shorter than seasonal. Figure 13c shows the transfer function of the sea surface height and bottom pressure for shallow-water (solid black line) and deep-water (solid blue line) regions. In the shallow water regions, bottom pressure and sea surface height are identical for timescales < 100 days. With increasing timescale, the sea surface height and bottom pressure become gradually out of phase, as the black dashed line shows. The deep-water regions show a similar pattern, but compared with the coastal areas, the transfer function is small, and the phase change starts at small timescales. Figure 13 (panels a and b) show the spatial pattern of the transfer functions for timescales < 1 month and 2-10 months. Except for the NwAC, EGC, WGC, and Alaskan Coastal Current region, sea surface height variability is identical to mass component variability for both the shallow water and deep-water regions at timescale < 1 month. At timescales of 2-10 months, mass component variability reflects sea surface height variability only in marginal seas, which indicates that barotropic processes dominate there. In the deep-water regions and strong current regions, the value of transfer function decreases, which suggests that contributions of steric height anomaly increase.

The mass component can be measured by bottom pressure observations. Considering the coherence between the mass component and sea surface height anomaly, we can use

bottom pressure observations, tide gauge data, and GRACE observations to monitor high-frequency sea surface height variability in the marginal seas and complement the altimetry system.

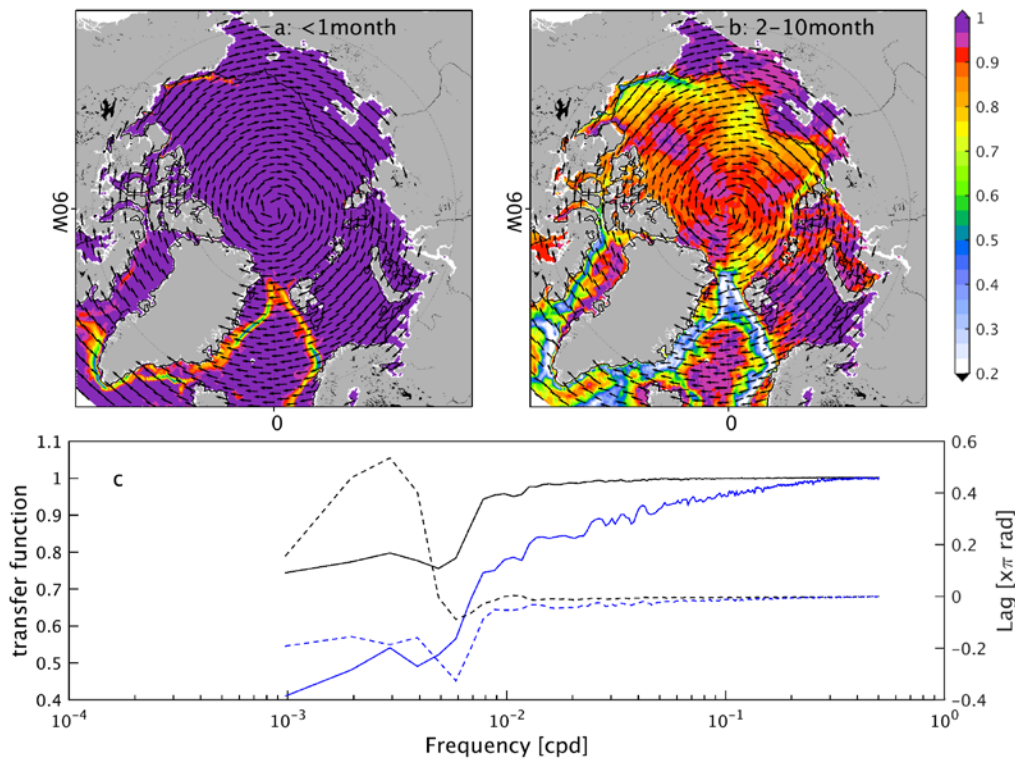


Figure 13. The transfer function of sea surface height and bottom pressure for timescales <1 month (a) and 2-10 months (b). (c) The transfer function for shallow water regions with depth ≤ 100 m (solid black line), and for deep water regions with depth > 100 m (solid blue line). The corresponding dashed lines represent the phase. The 95 % significance level for coherence is 0.38.

b. Annual cycle

Based on altimeter and GRACE data, Armitage et al. (2016) showed that sea level in the Arctic has a significant seasonal cycle, which is mainly dominated by steric effects. In this part, we will investigate the relation between the annual cycle of sea surface height and those of the bottom pressure contribution, halosteric effect, and thermosteric effect. The contribution of each layer (see Figure 12) to sea level variability is further investigated. The knowledge gained will be used to find locations for deploying observational systems and for deciding parameters to be observed.

For the annual cycle of sea surface height, significant variability exists along the coast. The high variability also crosses the 500 m isobath (Figure 14d). As Figure 14 (panels a and b) shows, both the bottom pressure and steric effect account for the sea surface height variability. Bottom pressure variability is significant along the coast, while significant steric height variability spreads to the deep-water regions, and along the NwAC, propagating into the Barents Sea. In the Arctic Ocean, halosteric effects account for the steric height variability (Figure 14h), especially from the upper layer (Figure 14i). In the Nordic Seas and south part of the Barents Sea, thermosteric effects dominate the variability, especially in the middle layer. Other layers (Figure 14, panels d, f, k) show a little contribution to the variability.

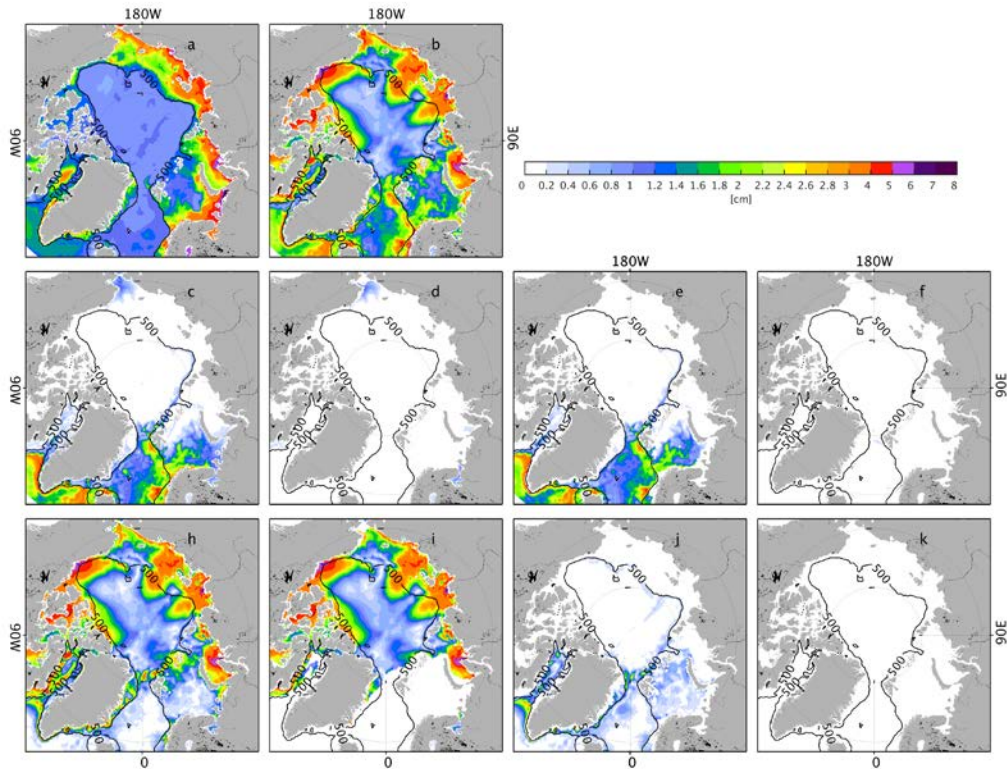


Figure 14. Root mean square variability of the annual cycle of mass component (a), steric sea surface height (b), thermosteric effect (c), and halosteric effect (h). Panels d-f and i-k are thermo/halosteric height variability components from the upper, the intermediate, and the bottom layers.

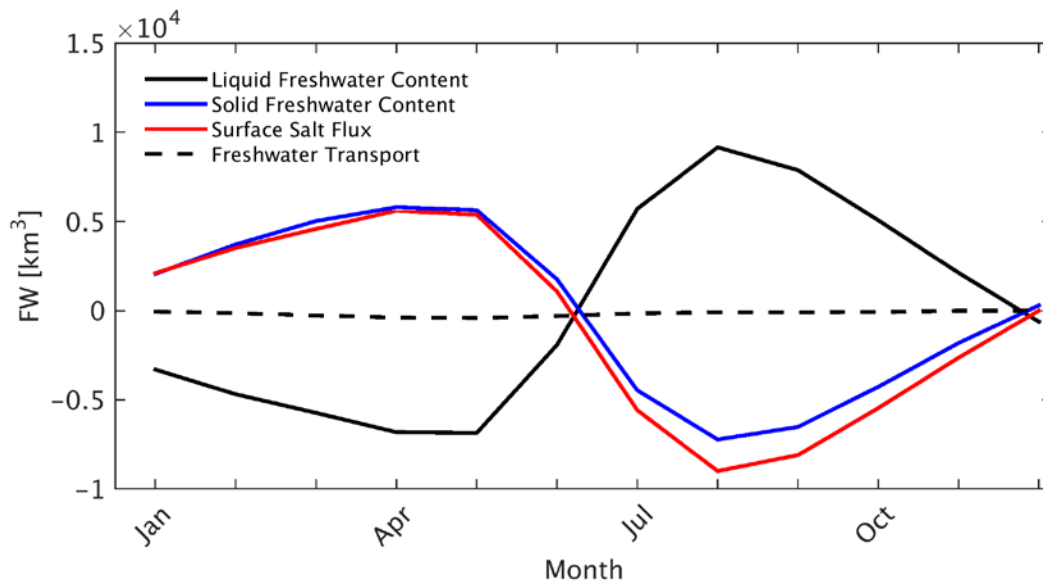


Figure 15. Annual freshwater budget: Liquid freshwater content (thick black line), solid freshwater content (blue line), accumulated salt flux over time and over the entire Arctic (red line, including ice melting, evaporation minus precipitation, runoff and nudging term), accumulated transport through all straits (dashed black line) over time.

Halosteric component accounts for a significant fraction of seasonal sea level variability in the Arctic, which related to upper layer salinity variability. In Figure 15, we show the an-

nual climatological freshwater budget. The liquid freshwater variability (black line) is mainly caused ice melting or freeze (blue line). Cumulative surface salt flux (red line) including river runoff, evaporation minus precipitation, nudging term, and ice changes are similar to the solid freshwater changes, which indicate that formation and melting of ice dominate the liquid freshwater changes. The effect of transport through the main straits is also minor. Therefore, ice observation or salinity observations, especially in the upper layer, may help to monitor the seasonal sea level variability. Bottom pressure observations may also complement altimetry system in the coastal regions. Temperature observations are essential along the NwAC and near the Barents Sea Opening.

c. Interannual Variability

On interannual timescales, the East Siberian Sea, NwAC, and EGC regions show variability ranging from 3 cm to 5 cm. The Kara Sea shows variability around 2 cm and seems to connect with a 2 cm variability in the Eurasian and Canadian Basins, which may reflect pathways of river run-off from the Eurasian Basin to the Canadian Basin (Morison et al., 2012).

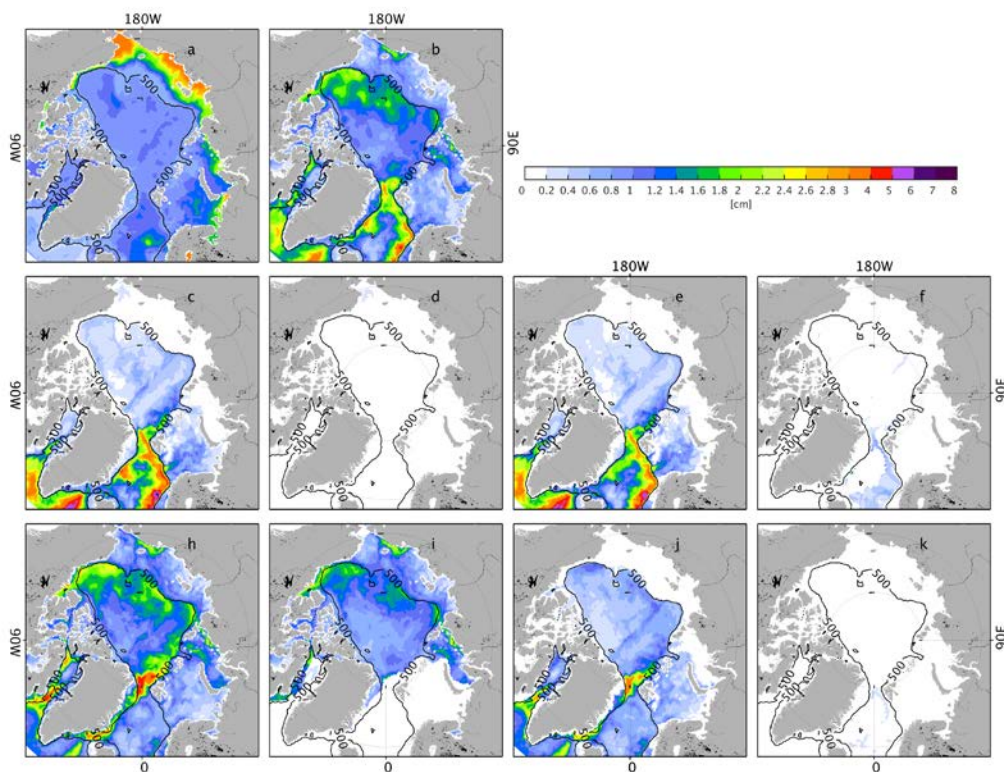


Figure 16. Same as Figure 14 but for interannual timescales (2-8 years).

In the coastal regions, especially of the East Siberian Sea, bottom pressure reflects the sea surface height variability (Figure 16a). Steric height is vital in the Eurasian and Canadian Basins and along the NwAC, WGC, and EGC regions. The Chukchi Sea also shows a steric height variability of 2 cm, which is related to Pacific water inflow through the Bering Strait. Contributions from the thermosteric and halosteric effects at each layer are displayed in Figure 16 (panels c and k). Thermosteric effect dominates along the NwAC, EGC, and WGC (Figure 16c), mainly in the middle layer (Figure 16e). The footprint of Atlantic water circulating the Eurasian and Canadian Basins is also visible (Figure 16e), although its contribution to

steric height variability is small (<0.8 cm). Halosteric effects dominate sea level variability in the deep-water locations of the Arctic Ocean (Figure 16h). In the upper layer, the halosteric effect is essential in the deep-water area near the East Siberian Sea and the Alaskan coast (Figure 16i). The halosteric effect in the middle layer shows a variability of 2.5 cm (Figure 16j) in the region near the Fram Strait, which likely reflects the shifting of the Atlantic Water and Arctic Water front. The Atlantic water footprint in the Arctic Ocean is more pronounced than that in Figure 16e, especially near the Alaskan Coast, where eddy-induced mixing transports freshwater downward to the middle layer (Spall, 2013).

On interannual timescales, sea surface height variability reflects mass component changes, which can be monitored by bottom pressure observations. Equipments that observing bottom pressure in the coastal region, especially in the East Siberian Sea, may help to improve the altimetry system. Salinity observations near the East Siberian section and Alaskan section are required. Halosteric effects are also visible in the Chukchi Sea and near Fram Strait. Thermosteric effects still dominate along the NwAC, EGC, and WGC. The footprint of Atlantic Water in the Arctic Ocean is shown both on halosteric and thermosteric effects in the middle layer. However, it is relatively small compared with halosteric effect in the upper layer.

d. Decadal Variability

The Arctic Ocean shows considerable decadal variability on freshwater content driven by large-scale wind (Proshutinsky & Johnson, 1997). Altimeter data is used to estimate freshwater content (Armitage et al., 2016; Giles et al., 2012), which points out the importance of salinity on monitoring decadal variability. On decadal timescales, the Arctic sea level shows a variability of 2~5 cm in the Beaufort Gyre, the East Siberian Sea and Eurasian Basin. The Greenland and Norwegian Seas also have a variability of 5 cm.

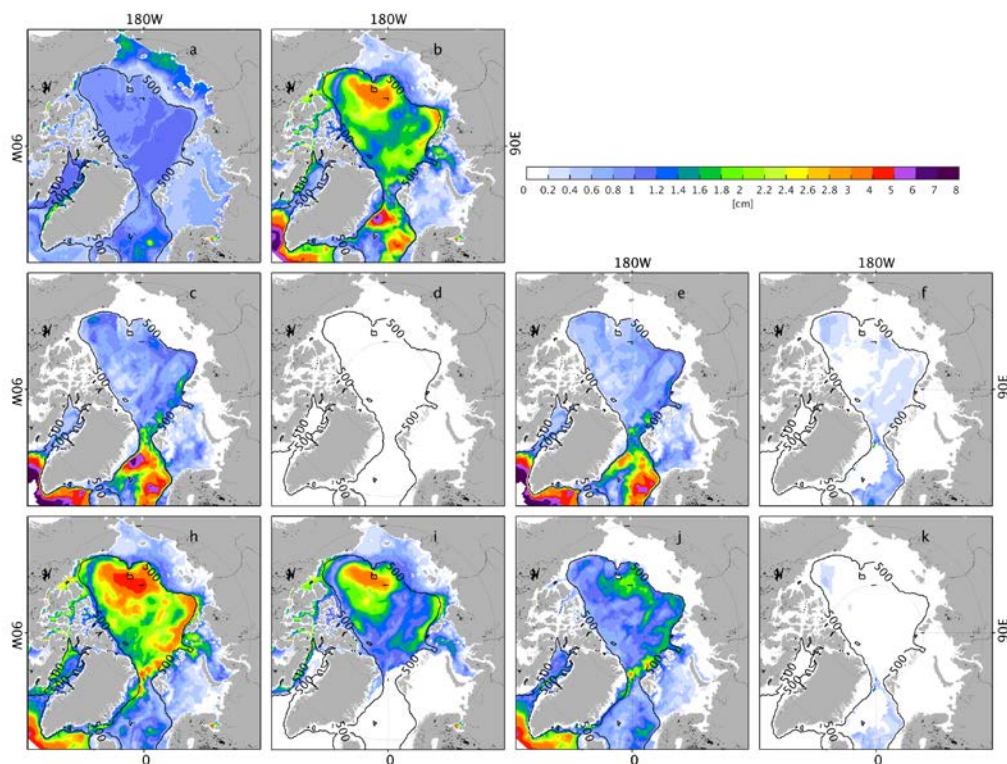


Figure 17. Same as Figure 14 but for decadal timescales (>8 years).

For the bottom pressure variability, as Figure 17a shows, the East Siberian Sea shows a variability of 1.5 cm. The Nansen and Makarov Basins show considerably more variability than the Canadian Basin. Steric height variability explains the high sea surface height variability in the Beaufort Gyre, the Eurasian Basin, the Norwegian Sea, and the Greenland Sea (Figure 17b). Decadal sea surface variability in the Norwegian and Greenland Seas is dominated by the thermosteric effect (Figure 17c), especially in the middle layer (Figure 17e). Halosteric variability dominates the Beaufort Gyre and Eurasian Basin (Figure 17h). The halosteric effect in the upper layer reflects freshwater changes in the Beaufort Gyre (Figure 17i). North of the Kara Sea, the upper layer halosteric height has a variability of 2 cm along the 500 m isobath. In the middle layer, halosteric effects on decadal timescales are more significant than those on other timescales (Figure 14j, and Figure 16j). The Arctic water interacts with Atlantic water near Fram Strait, which reduces the salinity of the Atlantic Water. Then it sinks beneath the upper layer and propagates along the 500 m isobaths in the Nansen Basin. Between the boundary of the Makarov and Canadian Basins, the halosteric height shows a considerable variability of 2 cm. It is not clear what causes this variability, and further studies are needed.

On decadal timescales, freshwater content variability dominates the Arctic Ocean, which can be monitored through sea surface height variability. In the Arctic Ocean, salinity is observed based on shipborne instruments (e.g., CTD) or moorings with limited spatial resolution. The results indicate that only observing the upper layer is not enough. Figure 17j shows that the Atlantic Water layer should also be observed. The relation between the decadal variability of sea surface height and the halosteric height also suggests one way of estimating freshwater content through assimilating altimeter data. To what extent freshwater content can be reconstructed by assimilating altimeter data will be investigated in the next stage with data assimilation experiment.

3.1.4 Existing observational system and gaps revealed by adjoint sensitivity

a. Altimeter observations and gaps

At timescales smaller than 30 days, sea level shows significant variability (Figure 11) in the coastal regions, which is related to bottom pressure (Figure 13). The significant variability may be related to local wind stress curl, and may also be propagated from upstream regions. Here, we use two adjoint simulation to demonstrate the causes.

Figure 18a shows the mean vertically-averaged velocity in the Barents Sea over 15 days. Two inflows from the North Atlantic are visible: one is along the Norwegian coast through the Norwegian Coastal Current, and one is separated from the Norwegian Atlantic Current. Part of the Norwegian Atlantic Current branch goes to the St. Anna trough following the topography, and the other branch seems to join the Norwegian Coastal Current and goes to St. Anna Trough along Novaya Zenlya island.

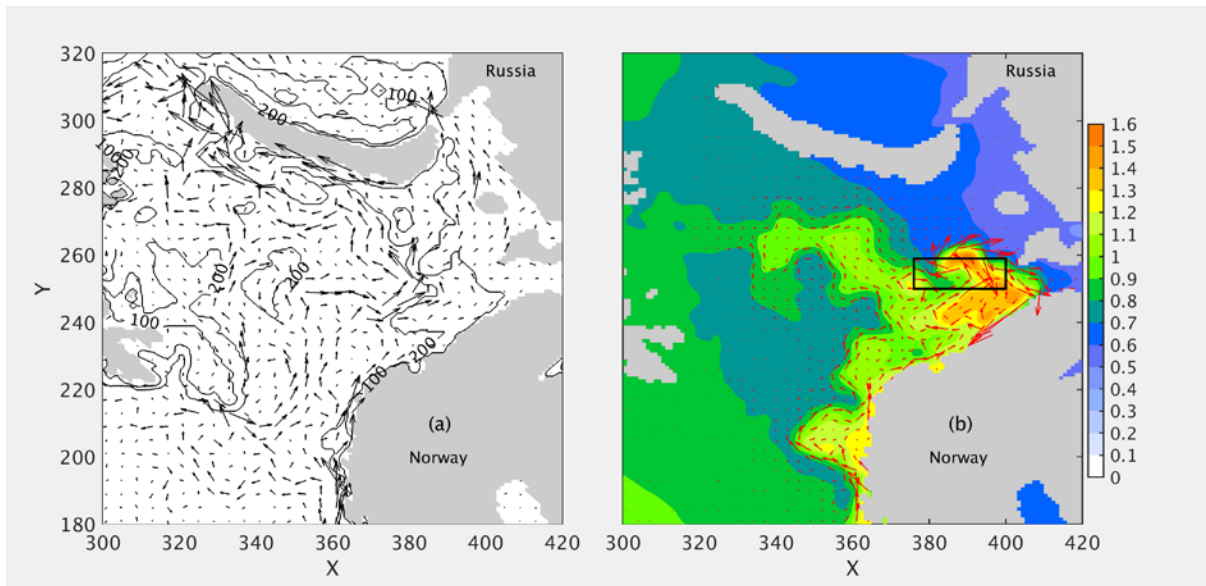


Figure 18. (a) Vertically-averaged velocity over the integration period. (b) The sensitivity of target J with respect to sea level (shade) and wind stress (vector) averaged over 15 days. The black box in panel b indicates the average region for J. The sensitivity is scaled by 1×10^5 .

Figure 18b shows time-averaged sensitivities for sea level and wind stress with respect to the mean sea level in the black box region. The sensitivities give the optimal perturbation pattern that influences the target. For instance, if we put sea level perturbation with an amplitude of 1 and with the pattern as shown in Figure 18b, the mean sea level over the red box region will increase by a value of 1 after the 15th day. The value of the sensitivity represents the contribution of each point to the target function. The averaged sensitivity in Figure 18b indicates that mean sea level in the target region is sensitive to sea level perturbations along the Norwegian Coast and the center of the Barents Sea. The sensitivity pattern matches the mean flow, indicating that sea level perturbation can be propagated to the target region through circulation. The wind stress sensitivity pattern and the sea level sensitivity pattern seems to follow the Ekman dynamic: within the most region of Figure 18b, an anticyclonic wind-stress pattern leads to convergence of water, which increases sea levels, and the sea level variability is then propagated downstream following the mean circulation.

Figure 19 shows vertically-averaged velocity (a) and mean sensitivities (b) over 15 days for the experiment in the East Siberian Sea. The velocity in the East Siberian Sea is influenced by the Bering Strait inflow and along coast current from the Laptev Sea. The sensitivity map in Figure 19b reveals that the wind stress curl lead to the high-frequency sea level variability in the target region. Changes of wind stress curl in the East Siberian Sea, the Laptev Sea, and the Kara Sea lead to sea level anomaly through Ekman dynamics, and then the sea level anomaly propagates to the target region.

In the two experiments above, we show that the high-frequency variability is related to Ekman dynamics, and the sea level anomaly can propagate following the current. The sensitivity analysis shown above is based on a 15-day simulation starting January 01, 2001. For a different season with a different circulation pattern, the sensitivities may indicate a different optimal pattern. For instance, considerable sensitivity may occur in the Bering Strait region in summer time in Figure 19. However, the Ekman dynamic may still dominate the sea level variability.

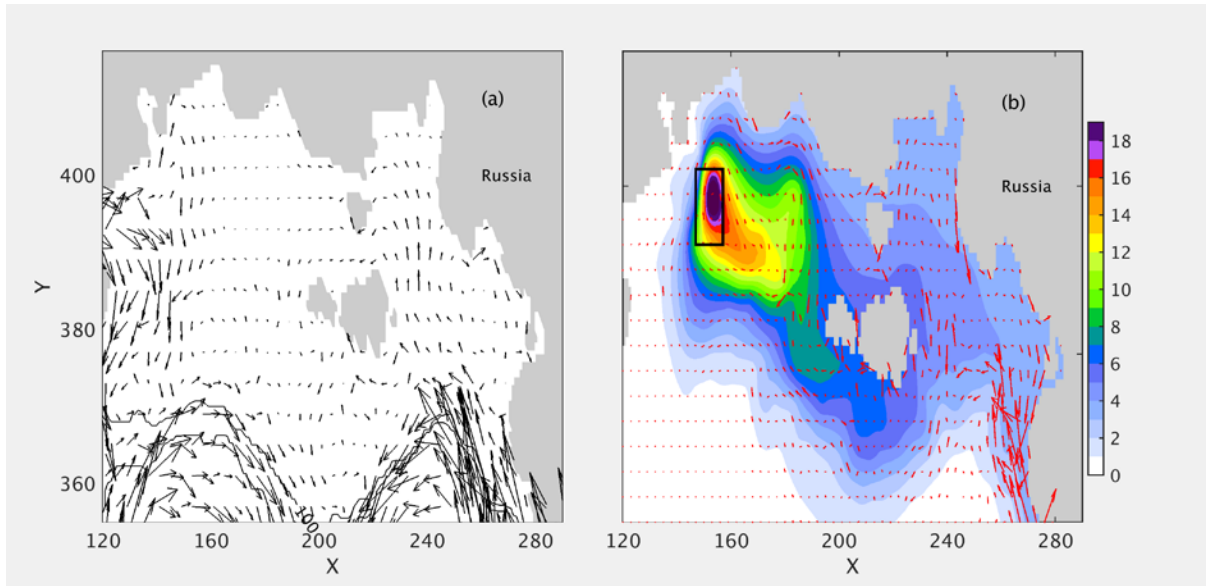


Figure 19. (a) Vertically-averaged velocity over the integration period. (b) The sensitivity of target J with respect to sea level (shade) and wind stress (vector) averaged over the integration period. The black box in panel b indicates the average region. The sensitivity is scaled by 1×10^5 .

In the Beaufort Gyre, sea level changes are related to salinity change on the decadal time scale (Figure 17). Figure 20 shows $\partial J / \partial S$ [10^{-6} $\text{m} \cdot \text{psu}^{-1}$] averaged over the first two years for the upper layer and the middle layer. The mean sea level is sensitive to salinity within the Beaufort Gyre in the upper layer (Figure 20a). The sensitivity indicates that reducing the salinity increases the sea surface height two years later. The vectors in Figure 15a shows the sensitivity of mean sea levels to wind stress curl. Increase of the anticyclonic wind stress pattern leads to convergence of the freshwater, which reduces density and increases the sea level. In the middle layer (Figure 20b), $\partial J / \partial S$ is smaller than in the upper layer. The salinity sensitivity pattern expands to the Makarov Basin, and is connected with the Laptev Sea, Kara Sea, and the Barents the Sea, which indicates that the freshwater may be transported to the Canada Basin from the marginal seas through circulation.

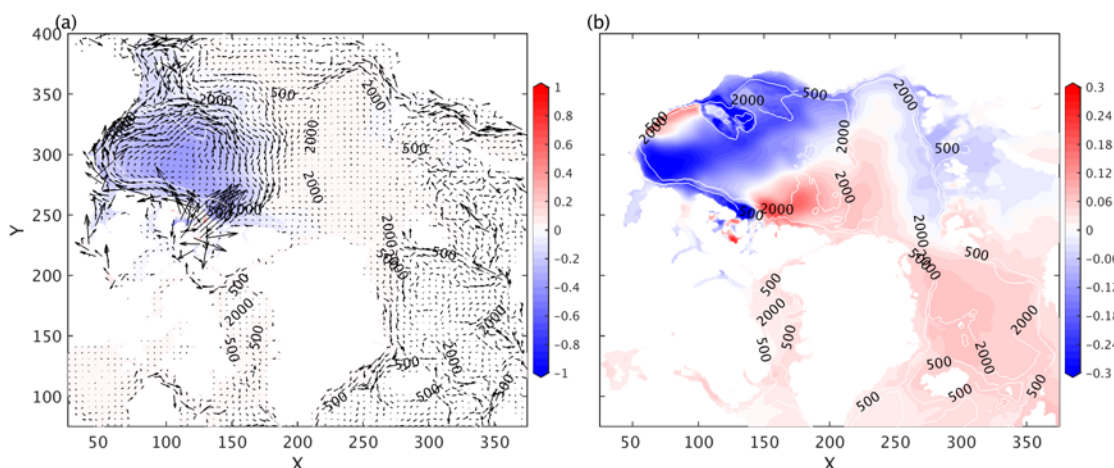


Figure 20. Sensitivity (shade) of mean sea level over the Beaufort Gyre with respect to salinity [10^{-6} $\text{m} \cdot \text{psu}^{-1}$] in the upper layer (a) and the middle layer (b). The vectors are sensitivity with respect to wind stress. The sensitivities are averaged over the first two years.

To monitoring the high frequency variability, the along-track altimeter observations may provide valuable information. However, sea level observations in the large area of the Arctic are not retrieved from the satellite altimeter because processing the satellite radar altimeter breaks down in the presence of sea ice. As more altimeter satellites are launched, coverage of altimeter data can be improved, especially in the marginal seas. However, the observing frequency will not change (Figure 3). Since sea level and bottom pressure are coherent at timescale smaller than 30 days, bottom pressure observation systems and tide gauges may be used to complement the existing altimeter system. At seasonal timescale, sea surface height variability reflects mass both mass transports changes and salinity changes related to ice melting and freezing. Ice observations and bottom pressure observation may complement the altimetry system for monitoring the seasonal variability. In the Beaufort Gyre, sea surface height variability is related to salinity changes caused by wind stress anomaly. Extracting sea surface height from the satellite by using specialized processing algorithms (Armitage et al., 2016) may improve our observing capability in the Beaufort Gyre.

b. The mooring system

Temperature and salinity anomalies affect the heat and freshwater transport mainly through two aspects: 1) the anomalies are advected to downstream regions by the mean circulation, and 2) the anomalies change the density and therefore the circulation through the thermal-wind relationship. Hence, directly analyzing the sensitivity of the heat transport and freshwater transport with respect to temperature and salinity can be complicated. Following Marotzke et al. (1999), we decompose the sensitivities of J_h and J_f to dynamic sensitivity and kinematic sensitivity. The kinematic sensitivities are given by:

$$\left(\frac{\partial J_f}{\partial S}\right)_\rho = \left(\frac{\partial J_f}{\partial S}\right)_T + \frac{\beta}{\alpha} \left(\frac{\partial J_f}{\partial T}\right)_S \quad (10)$$

$$\left(\frac{\partial J_h}{\partial T}\right)_\rho = \left(\frac{\partial J_h}{\partial T}\right)_S + \frac{\alpha}{\beta} \left(\frac{\partial J_h}{\partial S}\right)_T \quad (11)$$

where the dynamic parts are the last term on the right-hand side of the equations with negative signs, α and β are the thermal and haline expansion coefficients. The sensitivity of heat/freshwater transport to temperature/salinity is separated to two parts: the kinematic part (left hand in equations 10 and 11) reflects the sensitivity to temperature/salinity that is advected by the circulation, and the dynamic part represents the sensitivity to temperature/salinity that changes density. Via thermal wind relation, the dynamic sensitivity indicates circulation changes caused by temperature/salinity changes.

Figure 21a shows the sensitivity of freshwater transport with respect to salinity in the upper layer. Sensitivity to salinity is positive north of the Fram Strait, which indicates that a positive salinity anomaly there will increase northward freshwater transport or reduce southward freshwater transport. Since the mean velocity is southward (Figure 12b), increasing salinity there will reduce the southward freshwater transport, assuming salinity anomalies are advected to the Fram Strait. However, the kinematic sensitivity in Figure 21c indicates that advection of negative salinity anomalies to the Fram Strait can increase northward freshwater transport or reduce southward freshwater transport. The explanation for this contradiction is that: positive salinity perturbation (Figure 21a) leads to positive density anomaly (dynamic part, Figure 21b) which flatten and deepen the isopycnal; the southward current is reduced (Figure 21b) which reduces southward freshwater transport. Due to the

deepen isopycnal, fresher water near the surface is transported to this region through along isopycnal motions (Figure 21c). Therefore, freshwater transport observed by moorings in the Fram Strait mainly reflects changes of circulation. Sensitivities of freshwater transport with respect to salinity (Figure 21d), dynamic part (Figure 21e), and kinematic part (Figure 21f) in the middle layer are much smaller than those in the upper layer.

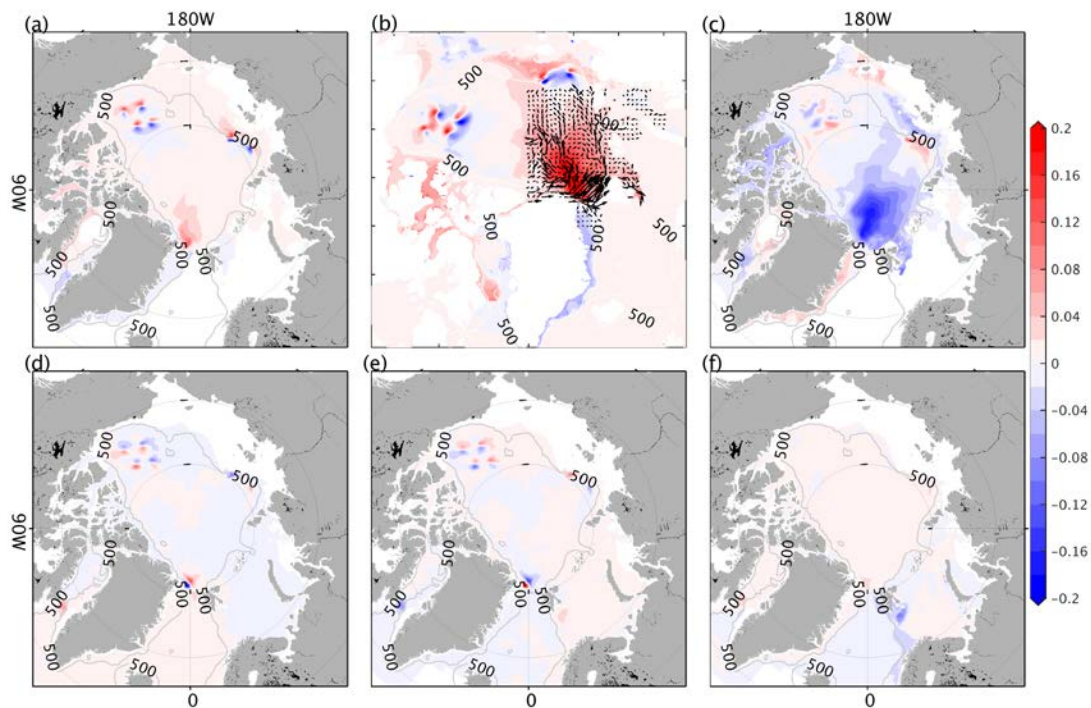


Figure 21. The sensitivity of freshwater transport observed by the moorings system [$\text{m}^3/(\text{s}\times\text{psu}\times\text{m})$] with respect to (a) salinity, (b) dynamic part, and (c) kinematic part in the upper layer. (d)-(f) is the same as (a)-(c) but in the middle layer. The sensitivities are divided by layer thickness.

Overall, the freshwater transport observed by the mooring observing system mainly monitors the dynamic changes in the Arctic Ocean. The Fram Strait system monitors changes in large part of the Nansen Basin. Moorings M1 and M2 can monitor dynamic changes following the circulation pattern. The four moorings in the Beaufort Gyre monitor local changes. Impact of the Davis Strait system is also limited in the Baffin Bay. Since the Bering Strait is near the model boundary, and the mean current is from the Pacific Ocean to the Arctic Ocean, the adjoint sensitivity that provides upstream information probably fails to recover the sensitivity. The effect of the Bering Strait on the model state will be evaluated through data assimilation in the next stage.

Sensitivities of the heat transport with respect to temperature, dynamic part, and kinematic part are shown in Figure 22. In the upper layer, large $\partial J/\partial T$ (Figure 22a) exist north of the Fram Strait, near Spitsbergen, north of the Laptev Sea, and in the Baffin Bay. This sensitivity is mainly caused by the kinematic part (Figure 22c). In the middle layer, large $\partial J/\partial T$ exist along the Norwegian Atlantic Current, the Norwegian Coastal current, and following the main pathway of the Arctic Atlantic water. The sensitivities in the Baffin Bay and the Beaufort Gyre system are also considerable, although they are located near the mooring location. Along the Norwegian Atlantic Current, the Atlantic inflow through the Fram strait and trough of Franz Josef land, and Baffin Bay, kinematic part dominates the sensitivity while dynamic part tends to compensate the kinematic effects. In the Beaufort Gyre, the

dynamic sensitivity dominates and kinematic part has no effect. In the Norwegian Coast Current, both the dynamic and kinematic parts show positive sensitivities, indicating that heat transport through the Barents Sea Opening can be increased by both advection of positive temperature anomaly upstream and velocity changes caused by positive density perturbations.

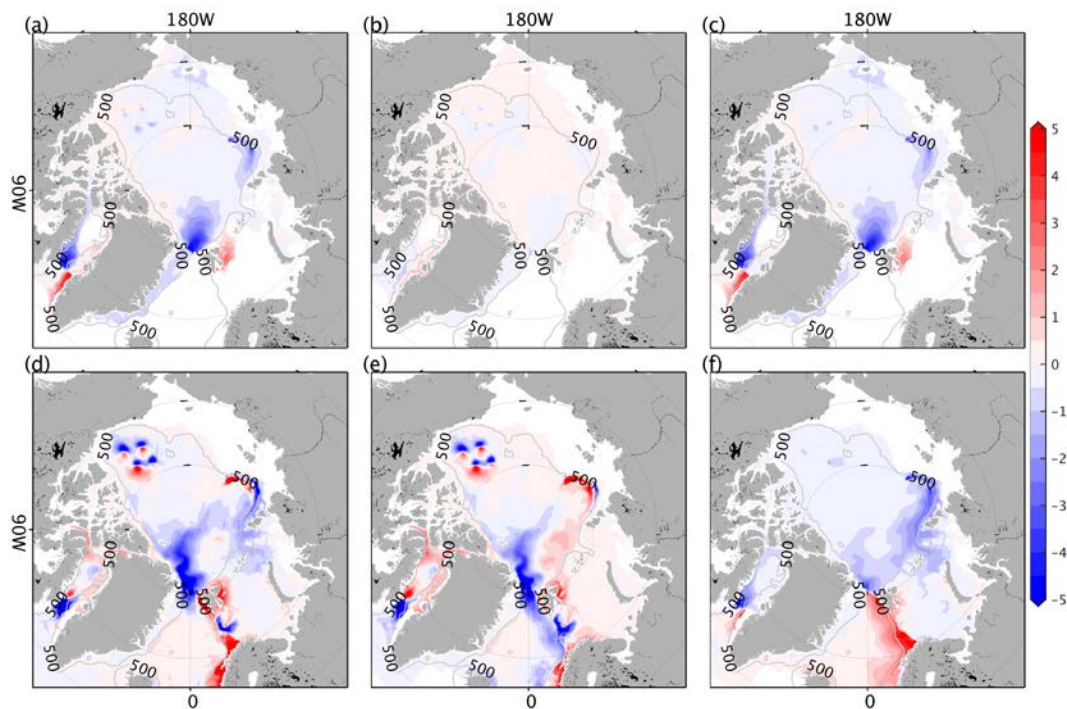


Figure 22. The sensitivity of heat transport observed by the moorings system [$\text{m}^3/(\text{s}\times\text{psu}\times\text{m})$] with respect to (a) temperature, (b) dynamic part, and (c) kinematic part in the upper layer. (d)-(f) is the same as (a)-(c) but in the middle layer. The sensitivities are divided by layer thickness, and therefore unit [m] appears in the denominator of the sensitivity unit.

The heat transport observed by the Fram Strait system, the Barents system, M1, and M2 depicts the pathways of the Atlantic inflow in the Norwegian Sea, in the Nansen Basin before its separate into two branches: one goes to the Makarov Basin, and the other one recirculates in the Nansen Basin. Sensitivity patterns in the Baffin Bay indicate the important role of the west Greenland Current and the Baffin Island Current. The Beaufort Gyre sensitivity is dominated by dynamic component and located near the moorings which indicate that dynamic process (changes of circulation related to density changes) dominates in this region. Observing changes of temperature and salinity are essential in the Beaufort Gyre.

3.2 Evaluation of the atmosphere observing system (FMI)

3.2.1 Differences between analyses and forecasts

Effects of radiosonde observations on operational analyses were identified by comparing analyses, and 12 h forecasts valid at the analysis time. The differences between forecasts and analyses show the impact of observations in the analysis fields. However, the impact of radiosonde observations only cannot be distinguished in this kind of analysis as all the assimilated observations influence the analysis. Locations, where observations contribute to a large difference between the forecast and analysis, are those where the existing observations are particularly valuable for numerical weather prediction models, whereas locations

with a large difference between the forecast and analysis but no in-situ observations nearby suggest locations where new in-situ observations may be valuable.

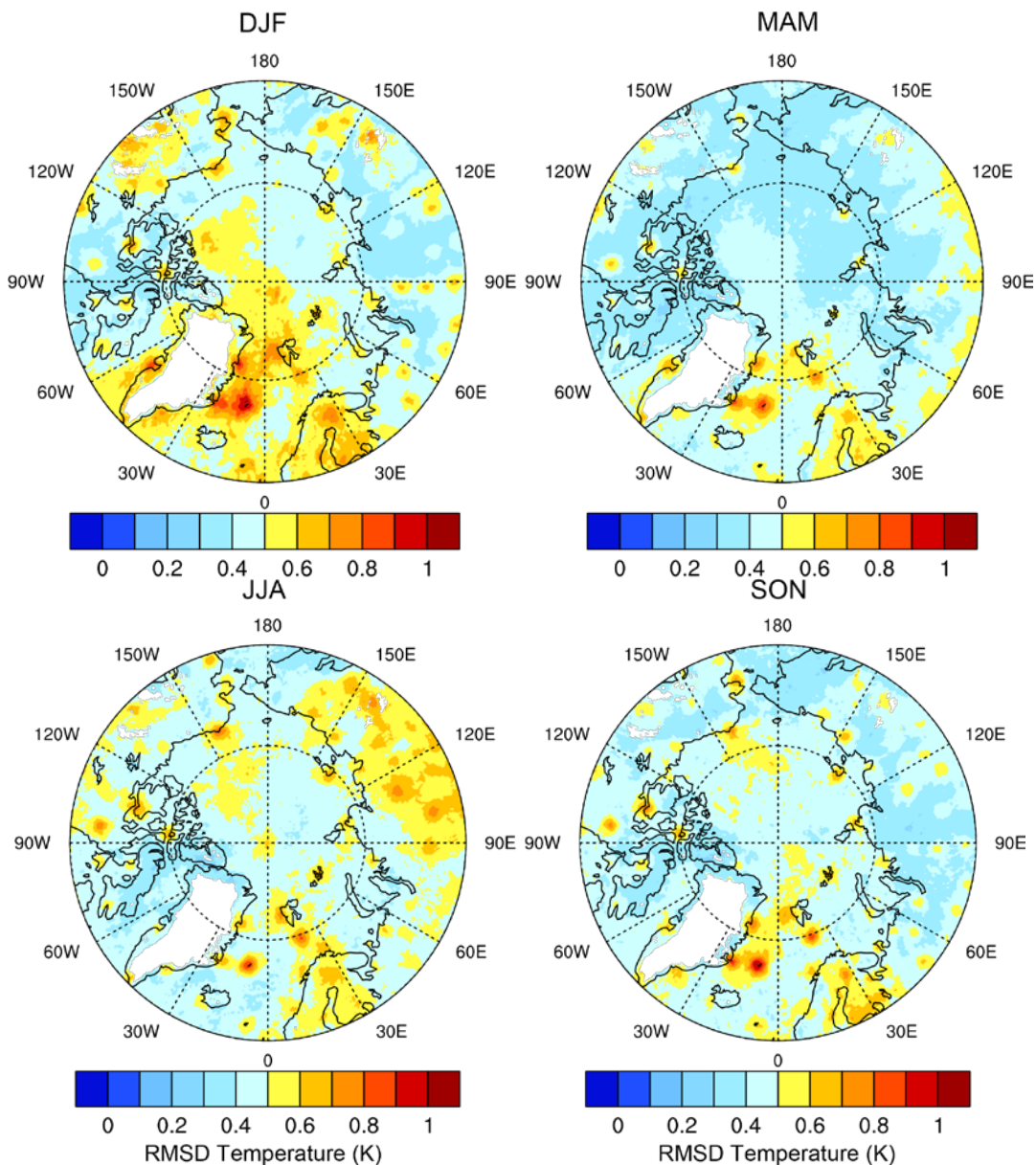


Figure 23. Root-mean-square-difference between analyses and forecasts (RMSD-FA) in 850 hPa-level temperature in January 2016–September 2018 (seasons: winter (DJF), spring (MAM), summer (JJA) and autumn (SON)).

The difference between 12 h forecasts and analyses shows that the largest differences occur nearby the sounding stations. Figure 23 shows the Root-Mean-Square difference between 12 h forecasts and analyses (RMSD-FA) in 850 hPa temperature. The signal of many of the sounding stations is visible in Figure 23 as a larger RMSD-FA than in the surrounding areas (compare to Figure 4). The same kind of pattern also occurred on other pressure levels (1000hPa, 925hPa, 700hPa, 500hPa). The small geographical radius of the RMSD-FA maxima suggests that many soundings had a significant contribution to the analysis, but their direct impact was limited to a relatively small area. The results suggest that the density of the sounding station network may have impacts on the accuracy of resulting analyses. In the

Arctic Ocean in winter, a large area with a relatively large value of RMSD-FA is seen, which is presumably related to the assimilation of satellite data.

RMSD-FA averaged over the whole Arctic (Table 3) was only 0.1 K smaller at 850 hPa level than RMSD-FA averaged for the locations of the sounding stations (Table 4). Some individual stations, especially in the North Atlantic, had large RMSD-FA and proved to be very valuable for numerical weather prediction. Nearby Jan Mayen (station 01001) the annual mean RMSD-FA was 0.9 K, which was almost twice as large as the average for the whole Arctic.

Seasonal variability in RMSD-FA was seen in many regions in the Arctic (Figure 23). For example, in Siberia, the RMSD-FA was generally largest in summer, and relatively small in autumn and spring. On the other hand, RMSD-FA was large during all seasons in the North Atlantic. The average RMSD-FA was largest near the surface and decreased upwards (Table 3 and Table 4).

Table 3. Average RMSD in temperature [K] between analyses and forecasts (RMSD-FA) in the area north of 60°N.

Level (hPa)	DJF	MAM	JJA	SON
1000	0.58	0.50	0.50	0.47
925	0.57	0.51	0.58	0.50
850	0.49	0.42	0.48	0.45
700	0.35	0.31	0.35	0.33
500	0.31	0.27	0.30	0.29

Table 4. Average RMSD in temperature [K] between analyses and forecasts (RMSD-FA) at the sounding stations

Level (hPa)	DJF	MAM	JJA	SON
1000	0.74	0.65	0.71	0.56
925	0.68	0.61	0.71	0.58
850	0.61	0.53	0.60	0.57
700	0.43	0.40	0.45	0.42
500	0.38	0.35	0.37	0.36

3.2.2 Differences between soundings and forecasts

We compared 12 h forecasts with radiosonde observations to identify regions where the forecasts have the largest deviation from the observations. Although significant differences typically suggest lower accuracy of forecasts, it is important to note that they may also be, at least partly, due to notable errors in the observations. The root-mean-square-difference between 12 h forecasts and radiosonde observations (RMSD-FO) was approximately twice as large as RMSD-FA. Furthermore, the geographical distributions of RMSD-FA and RMSD-FO were not similar (Figure 23 and Figure 24). Accordingly, differences between forecasts and analyses did not always correspond to differences between forecasts and radiosonde observations. Figure 24 shows that the largest RMSD-FO occurred in Siberia and the smallest in the Northern Europe and Alaska. This suggests that forecasts and/or radiosonde observations were less accurate in Siberia than in the other areas. It is noteworthy that, due to the high surface elevation, the 850 hPa level in Siberia is closer to the surface than over the oceans or in the Northern Europe or Western Russia. As the RMSD-FO was

typically largest near the surface (Table 5), the high surface elevation may contribute to the large RMSD-FO at the 850 hPa level in Siberia. However, regional distribution of RMSD-FO at the 700 hPa and 500hPa levels was similar to distribution at the 850 hPa level, suggesting that the regional distribution of RMSD-FO at the 850 hPa level is related to real geographical variation of accuracy of soundings and/or forecasts.

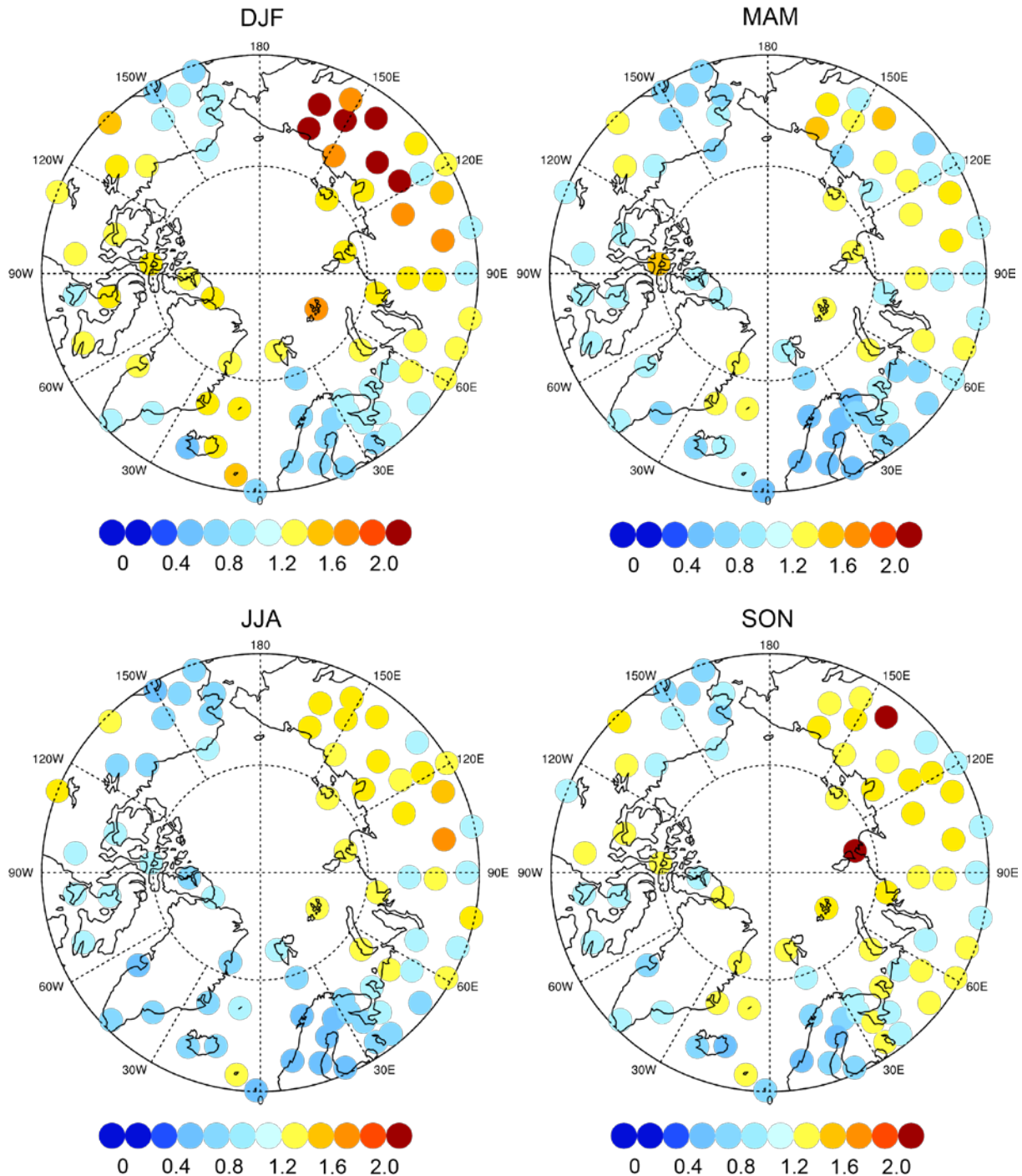


Figure 24. Seasonal root-mean-square-difference between soundings and forecasts (RMSD-FO) in 850 hPa-level temperature [K] in January 2016–September 2018 (seasons: winter (DJF), spring (MAM), summer (JJA) and autumn (SON)).

Table 5. Average RMSD in temperature [K] between soundings and forecasts (RMSD-FO) at the sounding stations

Level (hPa)	DJF	MAM	JJA	SON
1000	3.32	2.71	1.90	2.00
925	2.15	1.48	1.35	1.41
850	1.41	1.10	1.11	1.20
700	0.98	0.88	0.91	0.97
500	0.97	0.85	0.87	0.90

3.2.3 Weight of soundings

The weight of the soundings in the analyses was investigated by comparing RMSD between 1) analyses and forecasts (RMSD-FA), 2) radiosonde observations and forecasts (RMSD-FO), and 3) radiosonde observations and analyses (RMSD-AO) (Table 6). Where RMSD-FA was large but RMSD-AO was relatively small, the radiosonde observations can be seen as very important for the quality of analysis. On the other hand, in regions where RMSD-AO was large, the weight of observations was small in the assimilation and analysis. RMSD-AO might also be large due to a large uncertainty of these particular observations.

RMSD-AO (0.94 K) averaged over all the stations in the Arctic was 77 percent of RMSD-FO (1.22 K). However, average RMSD-FA was only 0.58 K. This means that, on the average, the assimilation has brought the analyses closer to the observations but, on average, analyses were still closer to forecasts than observations. Only at the stations in Northern Europe and at the Barrow station in Alaska (which is a station included in the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN)), the RMSD-AO is smaller than RMSD-FA. This indicates that, at these stations, the analyses were in good agreement with soundings. As RMSD-FA in the Northern Europe was even larger than the pan-Arctic average, the weight of soundings in the analyses was large in Northern Europe.

The weight given to soundings in the analysis was typically inversely related to RMSD-FO. This was seen as RMSD-FO and RMSD-AO being related to each other so that RMSD-FO was large where RMSD-AO was large. However, at Jan Mayen (station 01001), Barrow (station 70026), and Bear Island (station 01028), the RMSD-FO was large, but soundings also had a large impact on analyses, i.e., RMSD-AO was small. These findings suggest that these stations were critical for the quality of analysis.

In Figure 25, the scatter of instantaneous values of radiosonde observations minus forecasts (inst_O-F) is shown as a function of instantaneous values of analyses minus forecasts (inst_A-F) for four observation stations. The distance from the diagonal line is proportional to the difference between analyses and radiosonde observations. If the points are close to the diagonal line, the analysis is close to radiosonde observations, and it suggests a large weight of these soundings in the analysis. If the dots are, in turn, largely deviated from the diagonal line, the weight given to those radiosonde observations in the assimilation is small. The scatter had remarkable variations between the observation stations, and the four stations shown in Figure 25 represent different types of scatter seen at the stations.

Figure 25a (Jan Mayen, station 01001) represents a station in which the RMSD-FA was large, but RMSD-AO was relatively small. The smallest RMSD-AO was, however, found in Luleå Kallax (station 02185) (Figure 25b). Stations for which RMSD-AO was small were typically located in the North Atlantic, Northern Europe or Alaska. Figure 25c (Cambridge Bay, station 71925) represents a scatter that is typical for most of the stations in Canada and

Western Russia. Figure 25d (station Zyryanka, 25400) shows a large scatter typical for East Siberian stations.

Table 6. RMSD in 850 hPa Temperature [K] between analyses and forecasts (RMSD-FA), between soundings and forecasts (RMSD-FO) and between soundings and analyses (RMSD-AO)

Region	Station	RMSD Analysis – Forecast (RMSD-FA)	RMSD Soundings – Forecast (RMSD-FO)	RMSD Soundings – Analysis (RMSD-AO)
Alaska	70200	0.63	1.03	0.67
	70133	0.55	1.08	0.81
	70219	0.56	0.91	0.61
	70026	0.70	1.03	0.55
	70231	0.48	0.94	0.76
	70273	0.49	0.81	0.61
	70261	0.52	1.00	0.79
Canada	71964	0.59	1.44	1.14
	71957	0.55	1.03	0.77
	71043	0.54	1.27	1.08
	71934	0.63	1.24	0.97
	71925	0.70	1.23	0.79
	71926	0.66	1.15	0.74
	71924	0.63	1.38	1.06
	71917	0.43	1.10	0.91
	71915	0.53	1.08	0.75
	71081	0.56	1.19	0.86
	71909	0.52	1.19	0.93
	71082	0.51	1.26	1.05
Greenland	04220	0.67	1.01	0.69
	04270	0.61	1.11	0.93
	04417			
	04360	0.57	0.96	0.75
	04339	0.78	1.27	0.93
	04320	0.73	1.18	0.79
North Atlantic	04018	0.55	0.85	0.60
	04089	0.46	1.16	1.01
	01001	0.90	1.27	0.74
	06011	0.60	1.30	1.11
	03005	0.58	0.81	0.55
	01004	0.66	1.18	0.76
	01028	0.76	0.94	0.55
Northern Europe	01241	0.55	0.77	0.57
	01010	0.59	0.82	0.54
	02365	0.64	0.86	0.53
	02185	0.58	0.75	0.44

	02963	0.64	0.78	0.46
	02836	0.70	0.83	0.47
Western Russia	22217	0.50	0.97	0.79
	22113	0.53	0.97	0.75
	22820	0.57	1.00	0.78
	22522	0.55	1.03	0.76
	22845	0.53	1.02	0.77
	22543	0.54	1.15	0.91
	22271	0.52	1.10	0.85
	23802	0.58	1.06	0.76
	20744	0.51	1.25	1.00
	23205	0.55	1.11	0.82
	23415	0.56	1.14	0.84
	20046	0.63	1.51	1.23
Central Siberia	23921	0.54	1.25	0.97
	23330	0.53	1.21	0.93
	23933	0.55	1.25	1.10
	23955	0.58	1.25	1.00
	20674	0.56	1.30	1.03
	23472	0.58	1.25	0.93
	23078	0.52	1.33	1.07
	23884	0.58	1.11	0.79
	24507	0.56	1.69	1.40
	24908	0.56	1.07	0.76
	20292	0.60	1.54	1.28
	24125	0.59	1.53	1.24
	24726	0.60	1.59	1.33
	Eastern Siberia	24944	0.58	1.19
24641		0.54	1.36	1.12
24343		0.51	1.54	1.35
21824		0.53	1.42	1.21
24959		0.53	1.16	0.90
24266		0.50	1.85	1.65
21432		0.63	1.33	1.03
24688		0.65	3.57	3.29
21946		0.58	1.39	1.05
25400		0.53	1.64	1.42
25703		0.59	1.47	1.15
25428		0.53	1.78	1.53
25123		0.55	1.79	1.56

In the areas where radiosonde observations were close to forecasts, the radiosonde observations had a presumably larger weight in the analysis, but also satellite data assimilation may have contributed to this. Alternatively, in these areas the short forecasts were very accurate, which kept the difference between the forecasts and radiosonde observations small.

In some cases, the analysis was actually further away from a radiosonde observation than the forecast valid at the analysis time, indicating the influence of other observational data (e.g., satellites) in the assimilation process and analysis.

It can be concluded that the spatial variability in the weights of radiosonde observations in these analyses was large. It is, however, difficult to point out whether this variability arises from differences in the quality of radiosonde observations or different levels of utilization of radiosonde observations in the assimilation process. A key finding is that the variability in weights of radiosonde observations in the analysis seems to have a much larger impact on the quality of forecasts than the density of the sounding network.

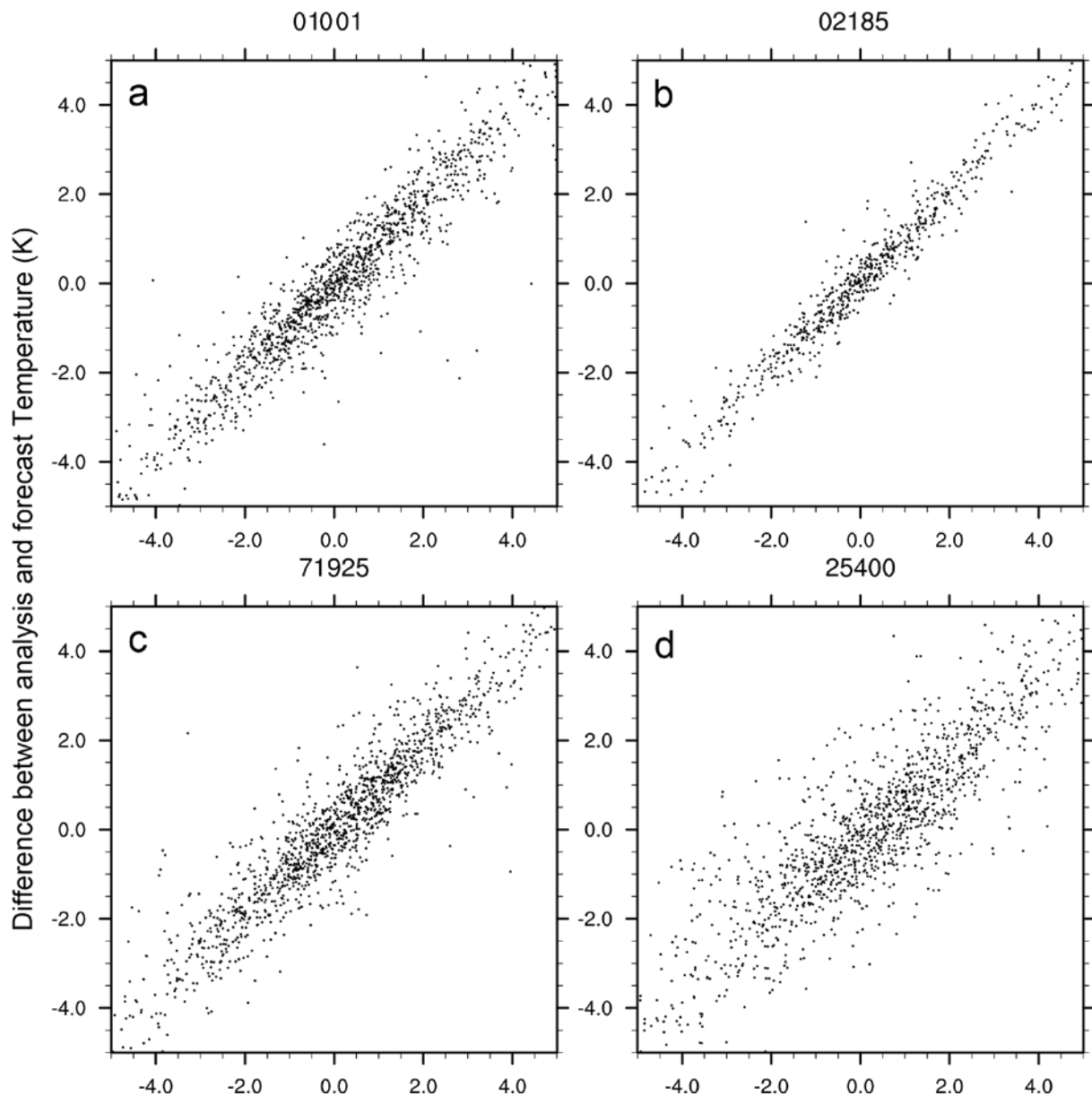


Figure 25. Scatter of instantaneous values of radiosonde observations minus forecasts (inst_O-F) as a function of instantaneous values of analyses minus forecasts (inst_A-F) in 850 hPa-level temperature for four observation stations: (a) Jan Mayen (station 01001), (b) Luleå-Kallax (station 02185), (c) Cambridge Bay (station 71925), and (d) Zyryanka (station 25400).

3.2.4 Effects of air-mass origin

The influence of air mass origin was analysed by calculating the difference between the forecast and sounding and associating each value with the starting point of the 12 h backward trajectory that ended up to the sounding station. This is hereafter called the difference associated with the trajectory starting point (DSP). Typically, at a single grid box there were starting points of trajectories ending up at several sounding stations. Figure 26a shows the spatial distribution of grid box RMS values of DSP between soundings and forecasts. Results were also very similar for 24 h backward trajectories (not shown). Spatial distribution of grid box RMS values of DSP between soundings and forecasts reflected the spatial distribution of RMSD between soundings and forecast at sounding station (Figure 26b). This was an expected result, as we noticed that the starting points of 12 h trajectories were typically not very far from sounding station, generally less than 800 km from the sounding stations.

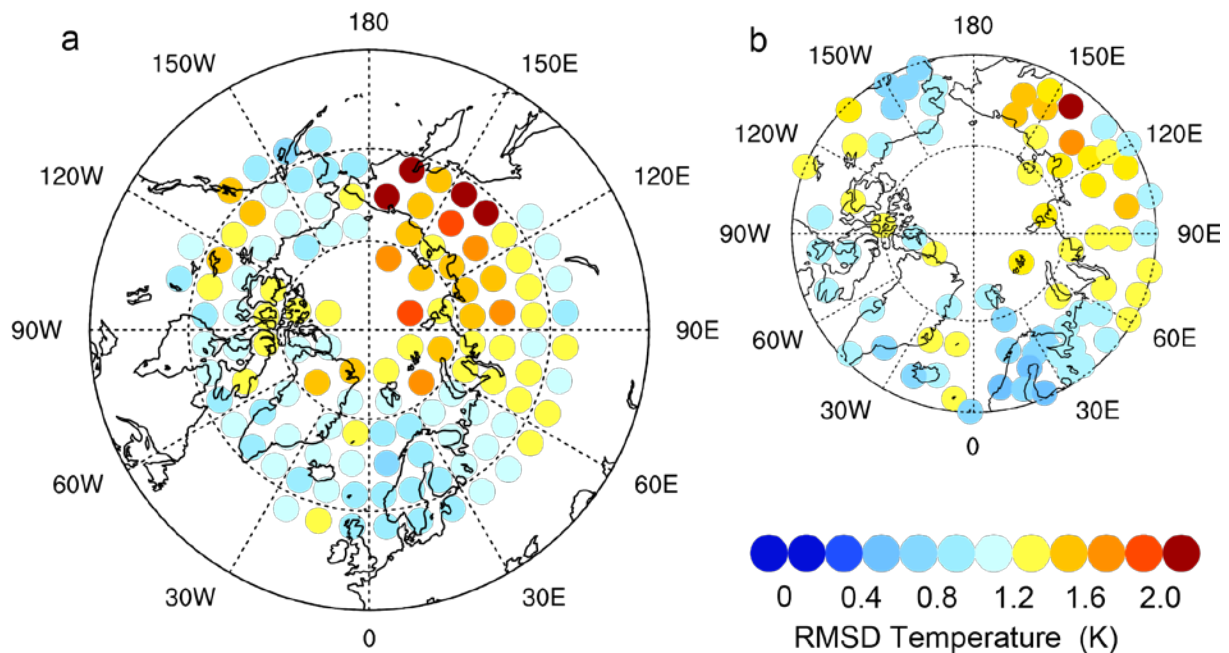


Figure 26. (a) The RMS value of differences in 850 hPa temperature between soundings and forecasts. The RMSD value is plotted for the starting points of 12 h backward trajectories ending at some of the 76 Arctic radiosonde stations. Only grid boxes which contain more than 100 trajectory starting points were plotted. (b) RMSD in 850 hPa level temperature between soundings and forecast at the sounding stations.

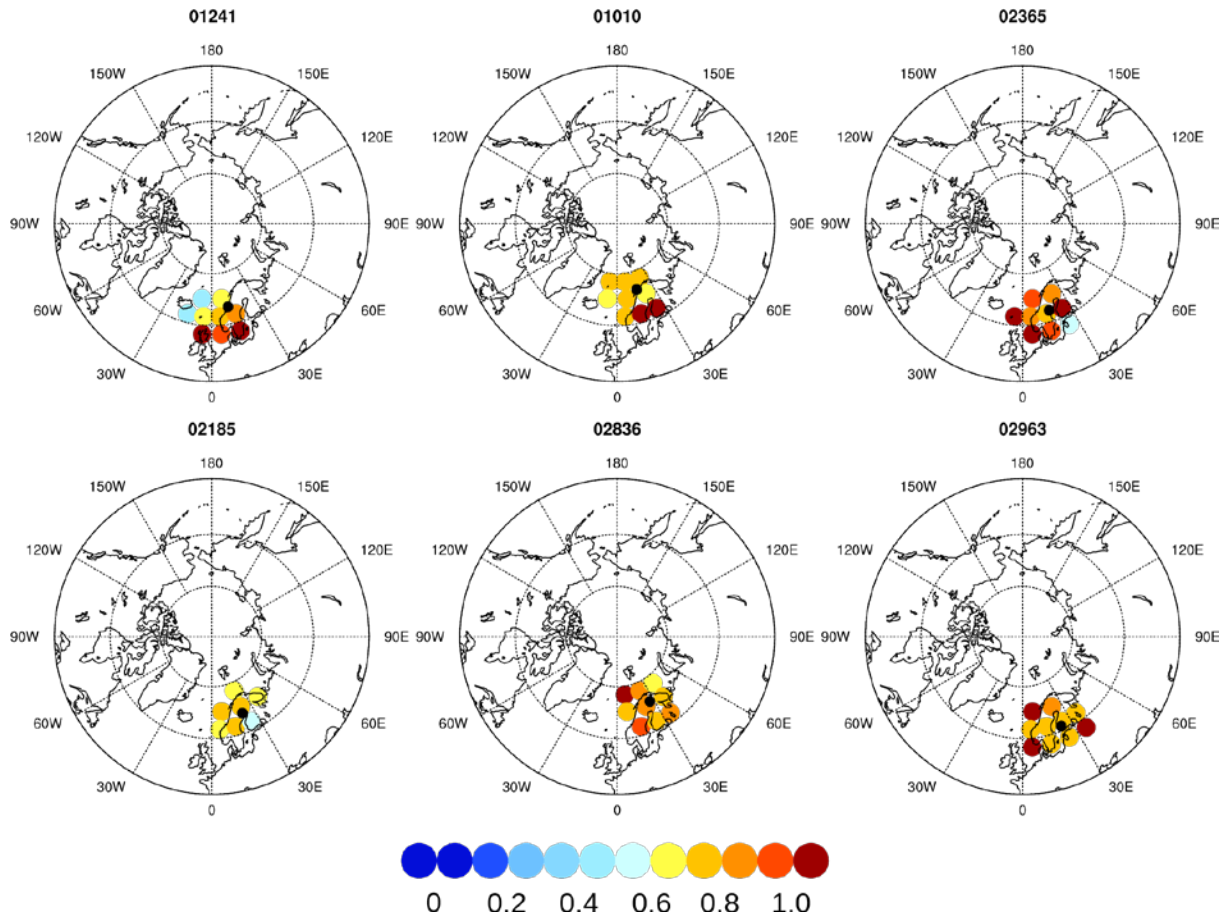


Figure 27. The RMS value of differences in 850 hPa-level temperature between soundings and forecasts. The RMSD value is plotted for starting points of 12 h backward trajectories ending at Orland (01241), Andoya (01010), Sundsvall Harnosand (02365), Luleå Kallax (02185), Sodankylä (02836), and Jokiainen (02963) sounding stations. The location of each sounding station is marked by a black dot. Only grid boxes which contain more than ten trajectory starting points are plotted.

Figure 27 presents grid box RMS values of DSP. When sounding stations were studied separately, the deviation of grid box RMS values of DSP between soundings and forecasts were large in the same grid box, when the air mass ended up in a different sounding station. This suggests that circulation patterns which affected the trajectories also affected the differences between soundings and forecasts. The exception to this is the central Arctic Ocean where the grid box RMS values of DSP between soundings and forecast indicate larger differences when trajectories starting points were in the central Arctic Ocean than on continents or near them. This seemed to be the case in five of the six sounding stations surrounding the Arctic Ocean (Figure 28).

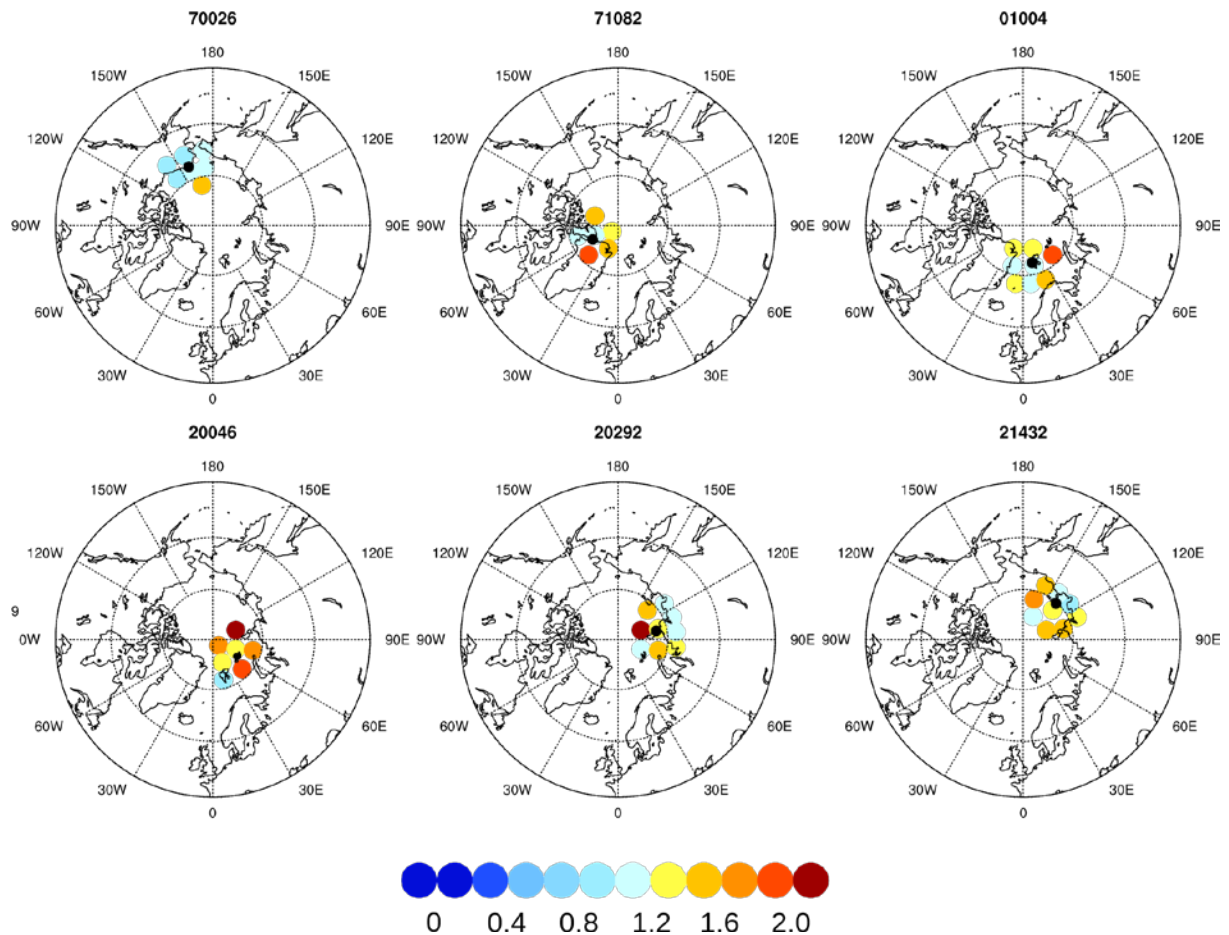


Figure 28. The RMS value of differences in 850 hPa-level temperature between soundings and forecasts. The RMSD value is plotted for starting points of 12 h backward trajectories ending Barrow (70026), Alert (71082), Ny Ålesund (01004), Polargmo Im. E.T. Krenkelja (20046), Gmo Im. E.K. Fedorova (20292), Ostrov Kotel’nyj (21432) sounding stations. The location of each sounding station is marked by a black dot. Only grid boxes which contain more than ten trajectory starting points are plotted.

Analyses of grid box RMS values of DSP between soundings and forecasts did not show that air mass origin had a large systematic effect on differences between soundings and forecasts. Instead, the station-to-station variations in the results suggest that it is not possible to robustly identify spatial gaps of sounding network without analyses of flow patterns.

3.2.5 Effects of synoptic-scale circulation patterns

Effects of synoptic-scale circulation patterns on the differences between forecasts and soundings were examined by categorizing atmospheric sea-level circulation patterns using the SOM method. A 4×5 arrays (20 nodes) of characteristic atmospheric circulation patterns was obtained. The results demonstrated seasonal differences: nodes 1–12, characterized by large pressure gradients, mostly represented circulation regimes occurring in winter (October–April) whereas nodes 13–20, characterized by a relatively weak pressure gradients, represented circulation regimes occurring in summer (May–September).

In winter, the variation in RMSD-FO between circulation regimes (i.e., the SOM nodes) was noteworthy, and relationships between the geographical distribution of RMSD-FO and the circulation regimes could be identified. However, among the nodes common in summer the geographical variations in RMSD-FO were less systematic, and RMSD-FO was generally smaller in summer than in winter.

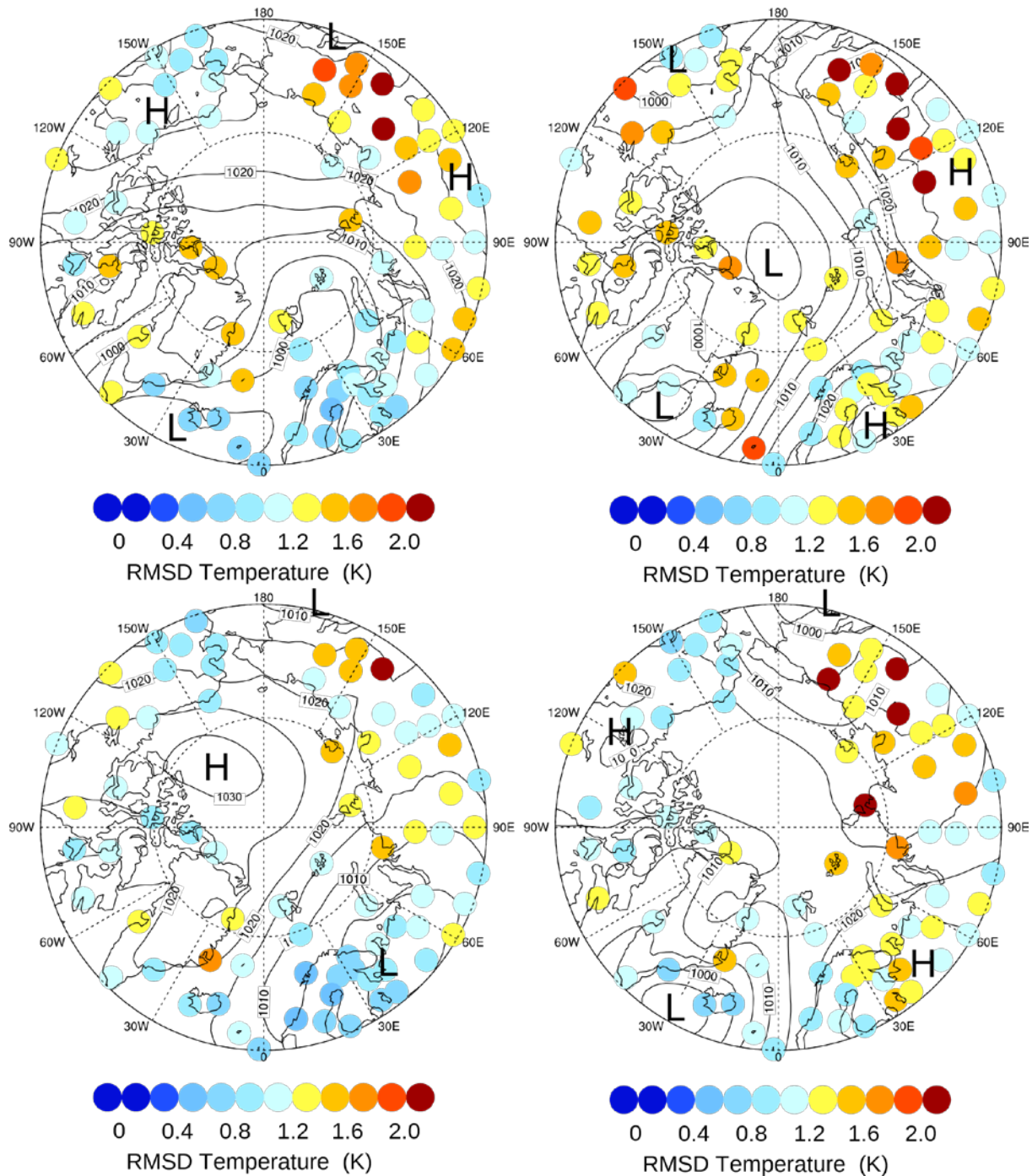


Figure 29. RMSD-FO in 850 hPa-level temperature and average mean sea-level pressure in four characteristic circulation regimes (i.e., SOM nodes) in winter.

Figure 29 shows four distinctive and characteristic circulation regimes (SOM nodes) for winter and geographical distribution of RMSD-FO in them. For example, in Northern Europe in winter, the 850 hPa air temperature (T_{850}) differences between forecasts and soundings were larger than average when and where the mean sea-level pressure was high (1020-1030 hPa in Figure 29 b and d, compared to 990-1010 hPa in Figure 29 a and c). However, the relationships between mean sea-level pressure and T_{850} varied between geographical locations. Thus, a certain kind of synoptic conditions did not affect RMSD-FO in a similar manner everywhere. Hence, the impacts of synoptic-scale conditions cannot be much generalized.

Differences in RMSD-FO between the SOM nodes were in the same order of magnitude with the differences between air-mass origins. This supports the conclusion that the effects of spatial gaps in the sounding network remarkably varied related to synoptic-scale circulation patterns.

3.3 Evaluation of the pan-Arctic atmospheric greenhouse gas monitoring network (MPG)

3.3.1. Single site footprint coverage: an example of the Ambarchik monitoring site

To demonstrate the spatial dimension of the field of view, or footprint, covered by a single atmospheric mixing ratio monitoring site, in this chapter we present the results for Ambarchik as an example for large-scale data coverage in the Arctic domain. The shading in Figure 30 highlights which level of information on greenhouse gas exchange processes can be retrieved within each area, based on the time series of mixing ratio observations, aggregated at eco-region scale (top panels) and 32 km grid resolution (bottom panels). In all panels, darker bluish colors indicate areas with high information content. These areas stretch over thousands of km², covering larger areas of the Arctic Ocean as well as Eastern and Central Siberia.

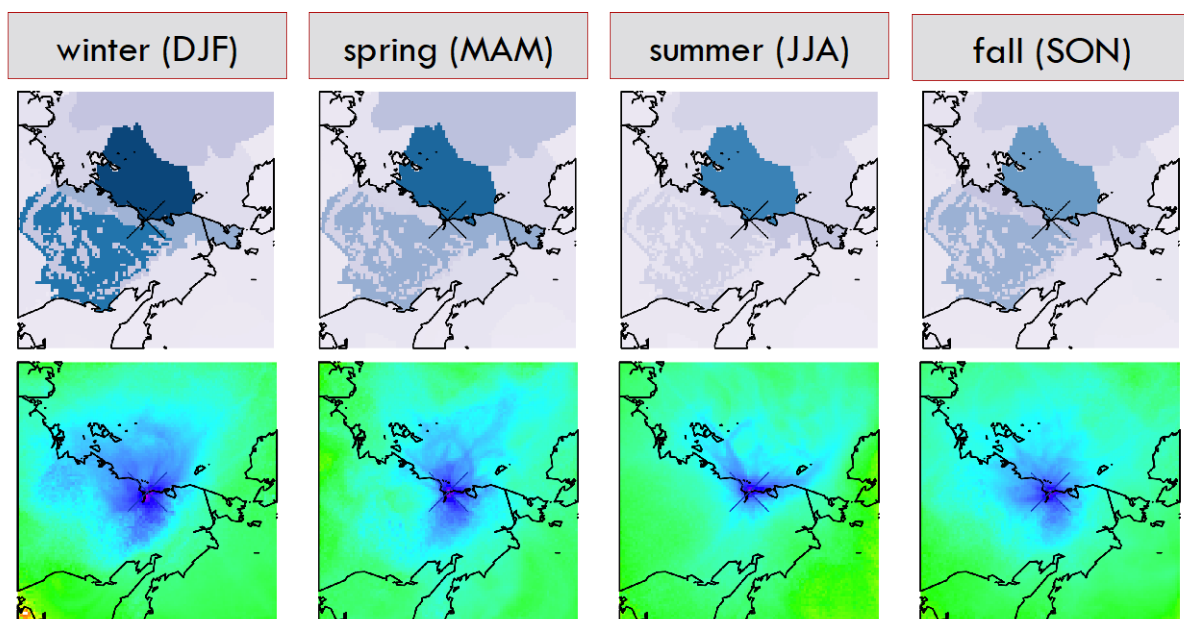


Figure 30. Example of a cumulative annual footprint function for a single observation site, Ambarchik (black cross). (top) Footprint function integrated at ecoregion scale (dark blue: high influence); (bottom) Footprint functions at pixel resolution (logarithmic scale, with purple and blue showing high influence).

Particularly the bottom panels of Figure 30 demonstrate the high temporal variability of the domains covered by the footprint, which is influenced by the shifts in atmospheric transport patterns that can yield very different fields of view both at short (daily) and longer (seasonal) timescales. The oceanic domain of the East Siberian Shelf that is situated just north of the observation site receives continuous footprint coverage throughout the year, but even here the extent of the dark blue core footprint areas varies strongly across seasons. For the terrestrial regions south of Ambarchik, this seasonal shift in wind climatology indicates that the prevalence of northern winds during spring and summer yields only marginal

information for the northwestern (spring) and central (summer) Siberian domains during these seasons.

The patterns outlined above are also reflected in the ecoregion-scale footprints, which are shown in the top panels of Figure 30. Here, all grid pixels that fall within a specific ecoregion are aggregated to a single sum, based on the assumption that environmental conditions within domains are fairly homogeneous so that data covering one part of them is also representative for the other sections. The ecoregion situated just north of Ambarchik, the East Siberian Sea, is covered by a high integrated footprint function, summed up across multiple towers throughout the year. In contrast, all terrestrial ecoregions show a strong seasonality in signal strength, and particularly during summer, the information yield is comparatively low even when summed up over larger areas. At the same time, the continuous coverage of the deep Arctic Ocean, even though at a low absolute level, demonstrates that also frequent coverage in the far field of individual footprints can lead to a decent information gain for larger domains.

3.3.2. Network footprint coverage: Seasonality of pan-Arctic footprints

As demonstrated for the single site evaluation shown above, also the spatial and temporal distribution of the information retrieved by the pan-Arctic network is dependent on site locations and prevailing atmospheric transport patterns. In this section, we present several aspects of the network coverage by plotting the footprint in different formats. One issue that complicates the evaluation is that absolute metrics of information content are rather abstract in this context, and the preferred (since more intuitive) format is to plot relative changes in data coverage within the target domain. The major objective here is, therefore, to follow a qualitative approach to highlight regions with limited data coverage within the pan-Arctic study domain.

Figure 31 shows the spatial patterns in relative footprint coverage at pixel scale across the Arctic, and how these patterns change over the seasons. As to be expected, the cumulative footprint functions focus strongly on the tower locations and their immediate surroundings (a logarithmic scale is used). Not surprisingly, it becomes obvious that high data coverage can mostly be achieved for regions that contain site clusters (e.g., Western Siberia, Scandinavia). At the same time, however, also the central Arctic Ocean as a rather remote domain shows an elevated relative influence since the far-field footprints of multiple towers frequently overlap here. Regions that stick out with rather limited data coverage throughout the year are the Russian Far East, most parts of Alaska, and the Eastern Canadian Provinces.

The highly resolved patterns shown in Figure 31 are well suited to demonstrate the seasonal variability in data coverage within the pan-Arctic network. The most obvious shifts between seasons, e.g., seen over Central Siberia or Western Russia with high information content in winter and limited coverage in summer, can be linked to seasonal shifts in prevailing wind directions. However, also the general mixing patterns within the atmosphere plays a role in this context. As noted in Section 2.3, the footprint is calculated as the residence time of simulated particles within the lower half of the atmospheric boundary layer. With more efficient convective mixing taking place in the warmer seasons, a higher fraction of particles gets vented into higher atmospheric levels, and therefore does not contribute to the footprints anymore. This effect can, e.g., be seen within the Canadian monitoring network, where the relative footprint contributions are high in most of the Western Provinces in the cold seasons, but a strong focus on tower locations exists during summer.

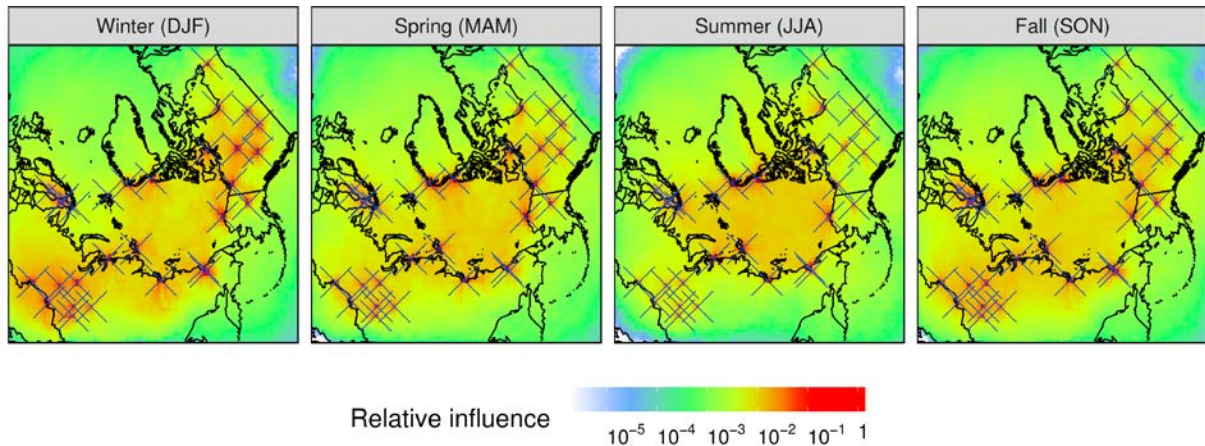


Figure 31. Extension of the gridded footprint functions (resolution: 100 km), accumulated for the entire pan-Arctic observation network of 29 sites. All values have been plotted on a logarithmic scale, with warm colors indicating areas where high information content could be retrieved by the atmospheric observations.

To better differentiate between regions with high and low footprint information content, respectively, Figure 32 shows the cumulative influence of the network coverage, binned into categories of 25 percent. For example, the first category (25 %) encompasses the pixels with the highest relative influence, which when summed up contribute one quarter to the total signal retrieved at the network of towers. The next shading level in Figure 32 indicates the area that is responsible for another 25 percent of the total signal, and so on. This version of plotting the network footprint further emphasizes the major findings indicated in Figure 31:

- The area that is responsible for the largest portion (75 %) of the retrieved signal concentrates on regions that contain tower clusters, i.e., Western Siberia, Scandinavia, Western Canada, and the Arctic coastlines.
- Due to overlapping footprints from multiple tower locations, the Arctic Ocean contributes substantially to the total footprint function.
- There is a seasonal shift in relative weights between the Arctic Ocean and terrestrial domains, with the latter receiving a higher cumulative influence during winter.
- Limited data coverage is indicated for Eastern Canada, Far Eastern Siberia, and Western Alaska. Western Russia and Central Siberia display pronounced seasonal shifts in coverage, with lower values in summer.

Signal aggregation to ecoregion scale can be used to evaluate the information content within the network footprint for combinations of different biome types and climate zones. As mentioned above, this strategy builds on the concept that conditions within each assigned ecoregion are largely homogeneous, so that footprints that only cover parts of a region will also provide information on all other parts outside the tower field of view. Figure 33 shows the relative influence (linear scale) of marine and terrestrial ecoregions (see also Figure 7) within the pan-Arctic study domain. These patterns clearly deviate from those shown in Figure 31 and Figure 32 above, since here also the size of the ecoregions plays a role. As a consequence, the deep Arctic Ocean clearly gets the highest cumulative footprint coverage, and the information content for most other regions appears relatively low, since all values were normalized with the maximum value across panels. Due to the size of the Atlantic Ocean deep water region, also this domain sticks out with a high relative information content, even though it does not contain any of the high-information areas when plotting footprints at pixel scale.

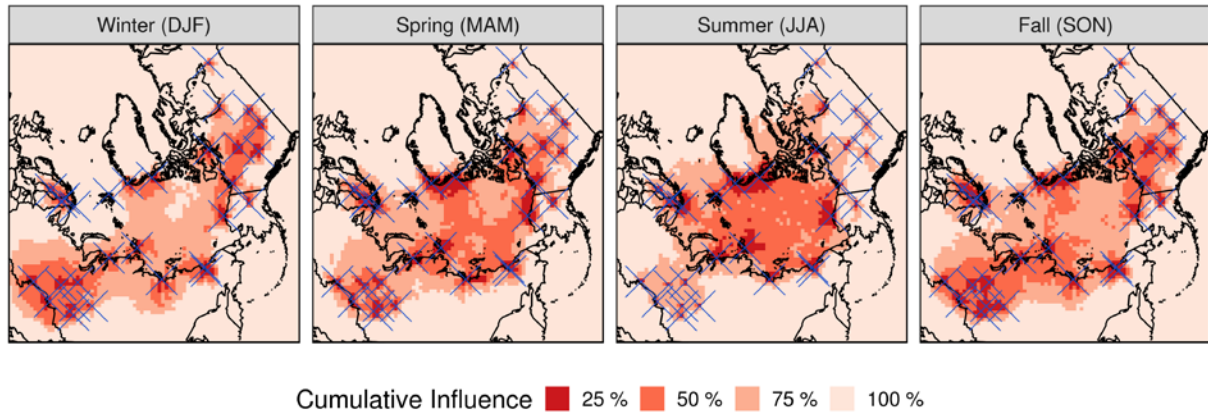


Figure 32. Gridded network footprint functions, aggregated into bins of equal information content. For example, 25% of the information retrieved by the towers comes from the areas shaded in the darkest red, which mark the high-information regions usually clustered around the tower locations.

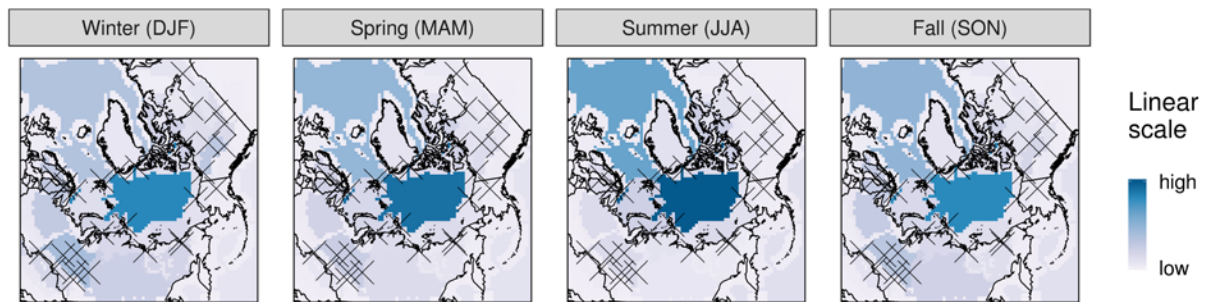


Figure 33. Relative information content of the aggregated network footprint at ecoregion scale. Based on a linear scale, dark colors indicate regions with relative high information content, while lighter colors show diminishing impact for the signals retrieved at the tower locations.

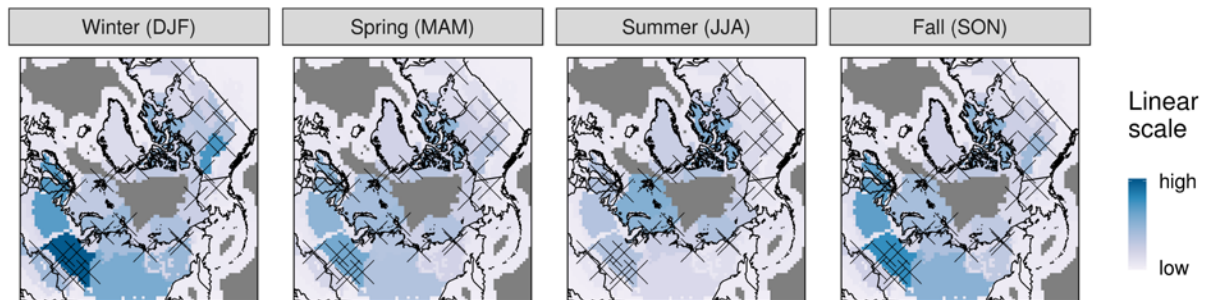


Figure 34. Same as Figure 33, only with remote ocean regions (grey shading) removed from the linear color scale that indicates relative importance of a region to the total information retrieved by the towers.

Since the dominance of the large deep ocean ecoregions in Figure 33 does not allow for a meaningful interpretation of pan-Arctic footprint information content, in a modified version of this ecoregion scale map (Figure 34) we masked out the deep ocean (grey areas). Normalized against a new reference, these patterns now largely confirm the general findings derived based on Figure 31 and Figure 32 above. Differences mostly concern the Russian domain, where the dense coverage by the Western Siberian cluster of sites (situated in the West Siberian Taiga) also projects into the neighboring East Siberian Taiga and Scandinavian and Russian Taiga, therefore seemingly extending the areas with high information content.

4. Summary of the identified gaps and recommendations

4.1. The ocean observing system (UHAM)

In the Arctic Ocean, the high-frequency (<30days) sea level variability is significant (order of 10 cm) in the marginal seas. Ekman dynamics dominate the high-frequency variability: wind stress anomaly leads to convergence (divergence) of sea water and therefore increase (decrease) of sea levels. The adjoint sensitivity in Figure 18 and Figure 19 reveals that upstream sea level anomaly can propagate to downstream. Depending on space and time scales of the relevant variability in the Arctic Ocean, several observing components are essential.

- 1) The satellite altimeter system is a critical system to monitor the high-frequency variability. With current satellite altimeter systems (Sentinel-3B, CFOSat, and Swot, Figure 3 c,d), a large area in the marginal Seas can be observed every 1-2 days in the summer time, which can be useful for monitoring the high-frequency variability. However, in the winter time, most of the area can be observed only every 5-10 days, leading to large observing gaps. As Figure 3 shows, the coverage of the satellite altimeter improves with an increased number of satellites, but the observing frequency does not improve, especially in winter time. The observing gaps are mainly caused by the presence of sea ice, where retrieving algorithms break down. Since sea level variability is coherent with bottom pressure variability over the entire Arctic Ocean, equipment that measures bottom pressure such as tide gauges or moorings in the ocean bottom may complement the satellite altimeter system.

High-frequency oscillations also matter in the in- and outflow of the Arctic, e.g., the Fram or Davis Straights. Transport monitoring systems in those dynamical boundary areas capturing the transport without eddy aliasing will be an additional essential ingredient for the high-frequency monitoring (see below).

- 2) The seasonal sea level is related to changes in freshwater and mass related components (bottom pressure). Changes in liquid freshwater content are mainly related to ice melting and formation. Therefore, sea ice observing systems may help to monitor the seasonal variability. Bottom pressure observations from tide gauges and GRACE can be used to monitor the mass related variability.
- 3) Decadal sea level variability in the Arctic Ocean is most significant in the Beaufort Gyre. It is due to freshwater content changes. Based on a sensitivity analysis (Figure 20), we see that the sea level is sensitive to local salt concentrations within the upper layer. Salinity anomalies may also come from the marginal seas around the Eurasian Basin which indicates the potential effect of observing salinity there on monitoring the Beaufort Gyre variability.

On decadal timescales, it is important to have a sufficient hydrographic observing component capable of capturing temperature and salinity changes over the entire Arctic Ocean from the surface to the bottom. Also, new algorithms (Armitage et al., 2016) that can recover sea level from sea ice covered areas may help to improve current satellite altimeter systems and to improve the ability of monitoring the Beaufort Gyre.

- 4) At last, we also analyzed the sensitivity of freshwater and heat transport observed by mooring systems with respect to model parameters. The Fram Strait mooring system is essential for monitoring the inflow of the Norwegian Atlantic Current, and outflow from the Arctic to the Atlantic. The Barents Sea Opening is essential for monitoring

inflow of the Norwegian Atlantic Current and the Norwegian Coastal Current to the Barents Sea. The two moorings north of the Laptev Sea are also crucial for monitoring the variability of the Atlantic inflow through the trough of Franz Josef Land, and circulation changes nearby. The Davis Strait mooring system mainly observes changes in the Baffin Bay. The four moorings in the Beaufort Gyre only provide sensitivities near the moorings. Effects of the Bering Strait mooring are not revealed since the adjoint sensitivity provide upstream information which is near the model boundary for the Bering Strait mooring system.

4.2. The atmosphere observing system (FMI)

The results of this investigation suggest that the density of radiosonde observation network is not the most critical factor for the quality of T850 forecast. The largest differences between forecasts and radiosonde observations were not directly found in data-sparse regions. Instead, the largest differences between forecasts and soundings typically occurred in Siberia, where the density of the measurements is not particularly low (Figure 4), but the observation quality seems to be a challenge. Furthermore, in the Atlantic Ocean, the differences between radiosonde observations and forecasts were relatively small even though the radiosonde observation network is sparse.

The results pointed out that stations on small islands in the middle of the Atlantic Ocean (Jan Mayen and Bear Island) have a remarkable impact on the analysis, and that they are critical for the quality of analysis. The impact of radiosonde observations varied between geographical locations. Only at the stations in Northern Europe and at the Barrow station, the difference between analyses and radiosonde observations was smaller than the difference between forecasts and analyses. The spatial distribution of RMSD-FO (Root-mean-square difference between forecasts and radiosonde observations) was closely related to the impact of soundings in the analyses, which cause the largest geographical variations in RMSD-FO. The regional differences are probably related to the quality of radiosonde observations, as different radiosonde types are used in different regions of the Arctic.

In addition, synoptic-scale circulation patterns affected RMSD-FO, and made identification of spatial gaps in radiosonde sounding network challenging. Large RMSD-FO was not systematically found in conditions when the air mass originated from regions with no radiosonde observations available. Therefore, we could not directly specify locations where the establishment of new sounding stations would inevitably improve quality of analysis and forecasts. However, the central Arctic Ocean and the Northern North-Atlantic would probably benefit most from new sounding stations.

Efforts to verify the quality of radiosonde observations, especially in Russia, and homogenizing the quality of radiosonde observations would presumably be very beneficial for the quality of T850 forecasts in the Arctic and sub-Arctic. In addition, current data assimilation systems are probably not adequate to optimally exploit the information from the existing observational network, as has already been pointed out by Jung et al. (2016).

The final interpretation of our results should be coordinated with the interpretation of results of data-denial experiments carried out above all in the context of the Year of Polar Prediction (YOPP). These approaches complement each other. Via data denial experiments we can better attribute the impact of individual observations, groups of observations and various types of observations. However, data-denial experiments require extensive computational resources, and can typically only address shorter periods, e.g., in YOPP they focus on two months in winter 2018 and three months in summer 2018, which may not be long

enough to be entirely representative. The approach we have applied here allows much longer study periods (here 33 months) making it more representative.

4.3. The GHG fluxes observing system (MPG)

We identified an existing network of 29 pan-Arctic atmospheric monitoring sites that provides continuous, well-calibrated observations on atmospheric greenhouse gas mixing ratios. Since each tower has a field of view that covers several thousands of km², and this 'footprint' shifts over time with atmospheric transport and mixing conditions, we employed atmospheric transport modeling to determine the spatio-temporal variability of the integrated network footprint function. Results were shown separated by major seasons, and at two different spatial aggregation levels (pixel and ecoregion scale).

Our analyses revealed that basic information for conducting atmospheric inversion studies to quantify surface-atmosphere greenhouse gas exchange processes is available for most regions in Canada, Europe, and Western Russia. This is particularly the case when assuming that carbon exchange processes are homogeneous on the ecoregion scale, which mostly in Russia extends the network coverage substantially into formerly poorly sampled areas. Also, the Arctic Ocean receives good overall footprint coverage, even though this region is rather remote, and contains no observational infrastructure itself; however, due to overlapping footprints from the large number of sites situated at or close to the Arctic coastline, the summed-up far-field contributions from individual sites yield a high relative information content. Major regions showing persistently limited coverage include the Russian Far East, Western Alaska, and the Eastern Canadian Provinces. Areas where footprint coverage gaps additionally exist seasonally include parts of Western Russia and Central Siberia. Accordingly, investments in additional observational infrastructure in any of these areas would be the most efficient approach to increase the overall coverage of the pan-Arctic atmospheric network for greenhouse gases.

List of references

- Aagaard, K., Swift, J., & Carmack, E. (1985). Thermohaline circulation in the Arctic Mediterranean seas. *Journal of Geophysical Research: Oceans*, *90*(C3), 4833-4846.
- Armitage, T. W., Bacon, S., Ridout, A. L., Thomas, S. F., Aksenov, Y., & Wingham, D. J. (2016). Arctic sea surface height variability and change from satellite radar altimetry and GRACE, 2003–2014. *Journal of Geophysical Research: Oceans*, *121*(6), 4303-4322.
- Biri, S., Serra, N., Scharffenberg, M. G., & Stammer, D. (2016). Atlantic sea surface height and velocity spectra inferred from satellite altimetry and a hierarchy of numerical simulations. *Journal of Geophysical Research: Oceans*, *121*(6), 4157-4177.
- Boyer, T., Levitus, S., Garcia, H., Locarnini, R. A., Stephens, C., & Antonov, J. (2005). Objective analyses of annual, seasonal, and monthly temperature and salinity for the World Ocean on a 0.25 grid. *International Journal of Climatology*, *25*(7), 931-945.
- Calafat, F., Chambers, D., & Tsimplis, M. (2013). Inter-annual to decadal sea-level variability in the coastal zones of the Norwegian and Siberian Seas: The role of atmospheric forcing. *Journal of Geophysical Research: Oceans*, *118*(3), 1287-1301.
- Curry, B., Lee, C., & Petrie, B. (2011). Volume, freshwater, and heat fluxes through Davis Strait, 2004–05. *Journal of Physical Oceanography*, *41*(3), 429-436.
- de Steur, L., Hansen, E., Gerdes, R., Karcher, M., Fahrbach, E., & Holfort, J. (2009). Freshwater fluxes in the East Greenland Current: A decade of observations. *Geophysical Research Letters*, *36*(23).
- Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., . . . Bauer, d. P. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553-597.
- Desai, A. R., Xu, K., Tian, H., Weishampel, P., Thom, J., Baumann, D., . . . Kolka, R. (2015). Landscape-level terrestrial methane flux observed from a very tall tower. *Agricultural and Forest Meteorology*, *201*(0), 61-75. doi: <http://dx.doi.org/10.1016/j.agrformet.2014.10.017>
- Dmitrenko, I. A., Polyakov, I. V., Kirillov, S. A., Timokhov, L. A., Frolov, I. E., Sokolov, V. T., . . . Walsh, D. (2008). Toward a warmer Arctic Ocean: Spreading of the early 21st century Atlantic Water warm anomaly along the Eurasian Basin margins. *Journal of Geophysical Research: Oceans*, *113*(C5).
- Fieg, K., Gerdes, R., Fahrbach, E., Beszczynska-Möller, A., & Schauer, U. (2010). Simulation of oceanic volume transports through Fram Strait 1995–2005. *Ocean Dynamics*, *60*(3), 491-502.
- Göckede, M., Turner, D. P., Michalak, A. M., Vickers, D., & Law, B. E. (2010). Sensitivity of a subregional scale atmospheric inverse CO₂ modeling framework to boundary conditions. *Journal of Geophysical Research-Atmospheres*, *115*, D24112. doi: 10.1029/2010jd014443
- Giles, K. A., Laxon, S. W., Ridout, A. L., Wingham, D. J., & Bacon, S. (2012). Western Arctic Ocean freshwater storage increased by wind-driven spin-up of the Beaufort Gyre. *Nature Geoscience*, *5*, 194. doi: 10.1038/ngeo1379
- Goodrich, J. P., Oechel, W. C., Gioli, B., Moreaux, V., Murphy, P. C., Burba, G., & Zona, D. (2016). Impact of different eddy covariance sensors, site set-up, and maintenance on the annual balance of CO₂ and CH₄ in the harsh Arctic environment. *Agricultural and Forest Meteorology*, *228–229*, 239-251. doi: <http://dx.doi.org/10.1016/j.agrformet.2016.07.008>

- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., . . . Yuen, C. W. (2002). Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models. *Nature*, *415*(6872), 626-630. doi: 10.1038/415626a
- Häkkinen, S., & Proshutinsky, A. (2004). Freshwater content variability in the Arctic Ocean. *Journal of Geophysical Research: Oceans*, *109*(C3).
- Henderson, J. M., Eluszkiewicz, J., Mountain, M. E., Nehr Korn, T., Chang, R. Y. W., Karion, A., . . . Miller, C. E. (2015). Atmospheric transport simulations in support of the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE). *Atmospheric Chemistry and Physics*, *15*(8), 4093-4116. doi: 10.5194/acp-15-4093-2015
- Ingleby, B., Rodwell, M., & L., a. I. (2016). Global radiosonde network under pressure. *ECMWF Newsletter No. 149*.
- Inoue, J., Yamazaki, A., Ono, J., Dethloff, K., Maturilli, M., Neuber, R., . . . Yamaguchi, H. (2015). Additional Arctic observations improve weather and sea-ice forecasts for the Northern Sea Route. *Scientific reports*, *5*, 16868.
- Jung, T., Gordon, N. D., Bauer, P., Bromwich, D. H., Chevallier, M., Day, J. J., . . . Goessling, H. F. (2016). Advancing polar prediction capabilities on daily to seasonal time scales. *Bulletin of the American Meteorological Society*, *97*(9), 1631-1647.
- Köhl, A. (2015). Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic. *Quarterly Journal of the Royal Meteorological Society*, *141*(686), 166-181.
- Köhl, A., & Serra, N. (2014). Causes of Decadal Changes of the Freshwater Content in the Arctic Ocean. *Journal of Climate*, *27*(9), 3461-3475. doi: 10.1175/jcli-d-13-00389.1
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., . . . Woollen, J. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, *77*(3), 437-471.
- Kittler, F., Eugster, W., Foken, T., Heimann, M., Kolle, O., & Göckede, M. (2017). High-quality eddy-covariance CO₂ budgets under cold climate conditions. *Journal of Geophysical Research: Biogeosciences*, *122*(8), 2064-2084. doi: 10.1002/2017JG003830
- Koldunov, N. V., Serra, N., Köhl, A., Stammer, D., Henry, O., Cazenave, A., . . . Gao, Y. (2014). Multimodel simulations of Arctic Ocean sea surface height variability in the period 1970–2009. *Journal of Geophysical Research: Oceans*, *119*(12), 8936-8954.
- Macdonald, R., Carmack, E., McLaughlin, F., Falkner, K., & Swift, J. (1999). Connections among ice, runoff and atmospheric forcing in the Beaufort Gyre. *Geophysical Research Letters*, *26*(15), 2223-2226.
- Marotzke, J., Giering, R., Zhang, K. Q., Stammer, D., Hill, C., & Lee, T. (1999). Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *Journal of Geophysical Research: Oceans*, *104*(C12), 29529-29547.
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans*, *102*(C3), 5753-5766.
- McPhee, M. G., Stanton, T. P., Morison, J. H., & Martinson, D. G. (1998). Freshening of the upper ocean in the Arctic: Is perennial sea ice disappearing? *Geophysical Research Letters*, *25*(10), 1729-1732.
- Michalak, A. M., Bruhwiler, L., & Tans, P. P. (2004). A geostatistical approach to surface flux estimation of atmospheric trace gases. *Journal of Geophysical Research-Atmospheres*, *109*(D14), D14109. doi: 10.1029/2003jd004422

- Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., . . . Zhang, B. (2014). Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions. *Global Biogeochemical Cycles*, *28*(2), 146-160. doi: 10.1002/2013gb0004580
- Morison, J., Kwok, R., Peralta-Ferriz, C., Alkire, M., Rigor, I., Andersen, R., & Steele, M. (2012). Changing Arctic Ocean freshwater pathways. *Nature*, *481*, 66. doi: 10.1038/nature10705
- Morison, J., Wahr, J., Kwok, R., & Peralta-Ferriz, C. (2007). Recent trends in Arctic Ocean mass distribution revealed by GRACE. *Geophysical Research Letters*, *34*(7).
- Notz, D., & Stroeve, J. (2016). Observed Arctic sea-ice loss directly follows anthropogenic CO₂ emission. *Science*, *354*(6313), 747-750.
- Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., . . . Kassem, K. R. (2001). Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience*, *51*(11), 933-938. doi: 10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2
- Pawlowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers & Geosciences*, *28*(8), 929-937.
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., . . . Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(48), 18925-18930. doi: 10.1073/pnas.0708986104
- Ponte, R. M. (1999). A preliminary model study of the large-scale seasonal cycle in bottom pressure over the global ocean. *Journal of Geophysical Research: Oceans*, *104*(C1), 1289-1300.
- Proshutinsky, A., Ashik, I., Häkkinen, S., Hunke, E., Krishfield, R., Maltrud, M., . . . Zhang, J. (2007). Sea level variability in the Arctic Ocean from AOMIP models. *Journal of Geophysical Research: Oceans*, *112*(C4).
- Proshutinsky, A., Bourke, R., & McLaughlin, F. (2002). The role of the Beaufort Gyre in Arctic climate variability: Seasonal to decadal climate scales. *Geophysical Research Letters*, *29*(23).
- Proshutinsky, A. Y., & Johnson, M. A. (1997). Two circulation regimes of the wind-driven Arctic Ocean. *Journal of Geophysical Research: Oceans*, *102*(C6), 12493-12514.
- Quadfasel, D., SY, A., WELLS, D., & TUNIK, A. (1991). Warming in the Arctic. *Nature*, *350*(6317), 385.
- Rödenbeck, C., Gerbig, C., Trusilova, K., & Heimann, M. (2009). A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models. *Atmospheric Chemistry and Physics*, *9*(14), 5331-5342.
- Rödenbeck, C., Zaehle, S., Keeling, R., & Heimann, M. (2018). How does the terrestrial carbon exchange respond to inter-annual climatic variations? A quantification based on atmospheric CO₂ data. *Biogeosciences*, *15*(8), 2481-2498. doi: 10.5194/bg-15-2481-2018
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Julio Bacmeister, Liu, E., . . . Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, *24*(14), 3624-3648. doi: 10.1175/jcli-d-11-00015.1
- Sato, K., Inoue, J., Yamazaki, A., Kim, J. H., Maturilli, M., Dethloff, K., . . . Granskog, M. A. (2017). Improved forecasts of winter weather extremes over midlatitudes with extra Arctic observations. *Journal of Geophysical Research: Oceans*, *122*(2), 775-787.

- Schauer, U., Loeng, H., Rudels, B., Ozhigin, V. K., & Dieck, W. (2002). Atlantic water flow through the Barents and Kara Seas. *Deep Sea Research Part I: Oceanographic Research Papers*, 49(12), 2281-2298.
- Sena Martins, M., Serra, N., & Stammer, D. (2015). Spatial and temporal scales of sea surface salinity variability in the Atlantic Ocean. *Journal of Geophysical Research: Oceans*, 120(6), 4306-4323.
- Serra, N., Kaese, R. H., Koehl, A., Stammer, D., & Quadfasel, D. (2010). On the low-frequency phase relation between the Denmark Strait and the Faroe-Bank Channel overflows. *Tellus A*, 62(4), 530-550.
- Skagseth, Ø., Furevik, T., Ingvaldsen, R., Loeng, H., Mork, K. A., Orvik, K. A., & Ozhigin, V. (2008). Volume and heat transports to the Arctic Ocean via the Norwegian and Barents Seas *Arctic-Subarctic Ocean Fluxes* (pp. 45-64): Springer.
- Skamarock, W. C., Klemp, J. B., Didhi, J., Gill, D. O., Barker, D. M., Duda, M. G., . . . Powers, J. G. (2008). A description of the Advanced Research WRF Version 3: NCAR Technical Note.
- Spalding, M. D., Fox, H. E., Halpern, B. S., McManus, M. A., Molnar, J., Allen, G. R., . . . Robertson, J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience*, 57(7), 573-583. doi: 10.1641/b570707
- Spall, M. A. (2013). On the circulation of Atlantic Water in the Arctic Ocean. *Journal of Physical Oceanography*, 43(11), 2352-2371.
- Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J., Cohen, M. D., & Ngan, F. (2015). NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society*, 96(12), 2059-2077.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183-7192.
- Thompson, R. L., Sasakawa, M., Machida, T., Aalto, T., Worthy, D., Lavric, J. V., . . . Stohl, A. (2017). Methane fluxes in the high northern latitudes for 2005–2013 estimated using a Bayesian atmospheric inversion. *Atmos. Chem. Phys.*, 17(5), 3553-3572. doi: 10.5194/acp-17-3553-2017
- Vinogradova, N. T., Ponte, R. M., & Stammer, D. (2007). Relation between sea level and bottom pressure and the vertical dependence of oceanic variability. *Geophysical research letters*, 34(3).
- Woodgate, R. A., Weingartner, T. J., & Lindsay, R. (2012). Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column. *Geophysical Research Letters*, 39(24).
- Woodworth, P., & Player, R. (2003). The permanent service for mean sea level: an update to the 21st Century. *Journal of Coastal Research*, 287-295.
- Wunsch, C., & Stammer, D. (1997). Atmospheric loading and the oceanic “inverted barometer” effect. *Reviews of Geophysics*, 35(1), 79-107.
- Yadav, V., Mueller, K. L., Dragoni, D., & Michalak, A. M. (2010). A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America. *Biogeosciences*, 7(9), 2655-2671. doi: 10.5194/bg-7-2655-2010
- Zhang, J., & Rothrock, D. (2000). Modeling Arctic sea ice with an efficient plastic solution. *Journal of Geophysical Research: Oceans*, 105(C2), 3325-3338.

----- END of DOCUMENT-----



INTAROS

This report is made under the project
Integrated Arctic Observation System (INTAROS)
funded by the European Commission Horizon 2020 program
Grant Agreement no. 727890.



Project partners:

