



# Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013

Jiping Xie<sup>1</sup>, Laurent Bertino<sup>1</sup>, François Counillon<sup>1</sup>, Knut A. Lisæter<sup>1</sup>, and Pavel Sakov<sup>2</sup>

<sup>1</sup>Nansen Environmental and Remote Sensing Center, Bergen 5006, Norway

<sup>2</sup>Bureau of Meteorology, Melbourne VIC3001, Australia

Correspondence to: Jiping Xie (jiping.xie@nersc.no)

Received: 23 May 2016 – Discussion started: 27 June 2016

Revised: 22 December 2016 – Accepted: 9 January 2017 – Published: 13 February 2017

**Abstract.** Long dynamical atmospheric reanalyses are widely used for climate studies, but data-assimilative reanalyses of ocean and sea ice in the Arctic are less common. TOPAZ4 is a coupled ocean and sea ice data assimilation system for the North Atlantic and the Arctic that is based on the HYCOM ocean model and the ensemble Kalman filter data assimilation method using 100 dynamical members. A 23-year reanalysis has been completed for the period 1991–2013 and is the multi-year physical product in the Copernicus Marine Environment Monitoring Service (CMEMS) Arctic Marine Forecasting Center (ARC MFC). This study presents its quantitative quality assessment, compared to both assimilated and unassimilated observations available in the whole Arctic region, in order to document the strengths and weaknesses of the system for potential users. It is found that TOPAZ4 performs well with respect to near-surface ocean variables, but some limitations appear in the interior of the ocean and for ice thickness, where observations are sparse. In the course of the reanalysis, the skills of the system are improving as the observation network becomes denser, in particular during the International Polar Year. The online bias estimation successfully maintains a low bias in our system. In addition, statistics of the reduced centered random variables (RCRVs) confirm the reliability of the ensemble for most of the assimilated variables. Occasional discontinuities of these statistics are caused by the changes of the input data sets or the data assimilation settings, but the statistics remain otherwise stable throughout the reanalysis, regardless of the density of observations. Furthermore, no data type is severely less dispersed than the others, even though the lack of consistently reprocessed observation time series at the beginning of the reanalysis has proven challenging.

## 1 Introduction

The Arctic Ocean plays an important role in the global climate system, where the sea ice at the interface between atmosphere and ocean regulates the fluxes of heat, moisture and momentum. The recent warming of the Arctic and the change of its water cycle has been linked to the following manifestations: a significant reduction and thinning of the sea ice cover (Johannessen et al., 2004; Shimada et al., 2006; Rothrock et al., 2008; Kwok and Rothrock, 2009), more freshwater in the Arctic in the 2000s (Haine et al., 2015) and more mobility and faster deformations of the Arctic sea ice (Rampal et al., 2009; Spreen et al., 2011). The interpretation of such changes is severely hampered by the sparseness of the concerned observations, which should not be improved dramatically in the near future. It can be assisted by free-running model simulations, but those are usually hampered by mislocations of ice edge and certain water masses. One possibility is to study surrogate locations where similar processes are assumed to take place. Another solution is to correct the dynamical model by assimilating observations available over relevant timescales.

The latter activities thus necessitate a state-of-the-art reanalysis system able to accurately honor the observations in a physically consistent manner. Recent efforts in Arctic Ocean state estimation have delivered either long-window optimizations (Nguyen et al., 2009, 2011) or, more often, short-window estimations (Schweiger et al., 2011; Mathiot et al., 2012; Sakov et al., 2012; Chevallier et al., 2013). Long-window optimizations deliver continuous model trajectories, which are physically more consistent than those using short windows. On the other hand, slicing the opti-

mization problem into short windows makes the estimation problem more linear or better conditioned (fewer unknowns and observations) and delivers more accurate products. Besides the window length, the choice of a background error covariance matrix is also a critical aspect in a data-scarce area such as the Arctic. The background error covariance used in an ocean data assimilation system can be – by increasing order of complexity – based on fixed multivariate spatial statistics (Cummings et al., 2009), an empirical estimation by a time-invariant ensemble (Oke et al., 2008) or a seasonally variable ensemble (Brasseur et al., 2005; Xie et al., 2011). In the case of ice–ocean systems, sea ice data assimilation often relies on rudimentary ice-only nudging methods (Schweiger et al., 2011; Tietsche et al., 2013); however, the possibility to account for flow-dependent coupled ice–ocean data assimilation updates has already been demonstrated in Lisæter et al. (2003). The pilot TOPAZ4 reanalysis of Sakov et al. (2012) has shown that the forecast error covariance from a dynamical ensemble mitigates the physical inconsistencies that could be expected from a short assimilation window.

The TOPAZ4 system is a coupled ocean–sea ice data assimilation system of the physical environment in the North Atlantic and Arctic oceans (see Fig. 1), which was initially used for short-term forecasting (Bertino and Lisæter, 2008) and later on for reanalysis (Sakov et al., 2012). TOPAZ4 represents the Arctic component of the CMEMS system (marine.copernicus.eu) where it is also used with coupling to an ecosystem model (Samuelson et al., 2015; Simon et al., 2015). The present paper follows the pilot TOPAZ4 reanalysis by Sakov et al. (2012) in which the performance of the same system has been demonstrated for the period of 2003–2008. They proposed an implementation of the EnKF data assimilation method that avoids ensemble collapse, provides reliable state-dependent error estimates and improves the match to independent observations compared to a free-running simulation.

Forced by the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al., 2011), TOPAZ4 assimilates most available measurements including along-track sea level anomalies (SLAs) from satellite altimeters, sea surface temperatures (SSTs), sea ice concentrations (SICs) and sea ice drift (SID) from satellites as well as in situ temperature and salinity profiles. The proposed reanalysis is 4 times longer (1991–2013) than the pilot reanalysis, and includes data-scarce periods with poor observational coverage and more intense observing efforts, such as during the International Polar Year (IPY, 2007–2009). The focus of this study is to provide a quantitative assessment of the reanalysis performance in the pan-Arctic region (defined as north of 63° N) in order to guide the user through its skills and limitations. In particular, we investigate the stability of the ensemble reliability through changes of the Arctic observational network, the variability of the system accuracy

in different sub-areas, its seasonal cycle and its trend in the course of the reanalysis.

The outline of this paper is as follows: in Sect. 2, the reanalysis system is described including the model, the data assimilation scheme and their implementation. Section 3 evaluates the reliability of the reanalysis ensemble. In Sect. 4, we compare the ensemble mean against available observations: altimetry, SST,  $T$ – $S$  profiles, ice concentration, ice drift and ice thickness. For each of these quantities, we assess the variability of the system performance in space or in time. Section 5 summarizes and discusses the potential improvements of our system for the next version of the reanalysis.

## 2 The reanalysis system

### 2.1 The HYCOM ice–ocean model

The TOPAZ4 system uses version 2.2 of the Hybrid Coordinate Ocean Model (HYCOM) developed at the University of Miami (Bleck, 2002; Chassignet et al., 2003). It uses 28 hybrid  $z$ -isopycnal layers, and the top layer has a minimum thickness of 3 m. The model grid has a horizontal resolution of 12–16 km, which is eddy permitting from the Equator to the Nordic Seas but is still far from being eddy resolving in the Arctic. The lateral boundaries of temperature and salinity are relaxed to a combination of the World Atlas of 2005 (WOA05; Locarnini et al., 2006) and version 3.0 of the Polar Science Center Hydrographic Climatology (PHC; Steele et al., 2001). HYCOM is coupled to a sea ice model in which the ice thermodynamics are described in Drange and Simonsen (1996) and the elastic–viscous–plastic rheology in Hunke and Dukowicz (1997). The surface momentum fluxes use a bulk formula parameterization (Kara et al., 2000), and the related thermodynamic fluxes are computed as described in Drange and Simonsen (1996).

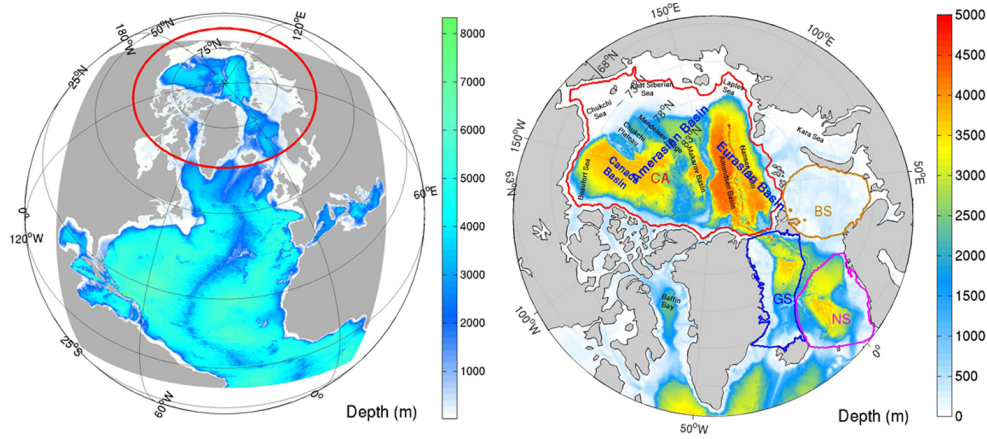
The model has been initialized from the same climatology data as used at the boundaries. The Pacific water inflow is imposed by a barotropic inflow through the Bering Strait at the model boundary and balanced by an outflow at the southern boundary of the domain. Unlike in Sakov et al. (2012), the inflow varies seasonally as found in observations (Woodgate et al., 2005): with a maximum in June (1.3 Sv), a minimum in January (0.4 Sv), and the mean transport is 0.8 Sv.

### 2.2 Data assimilation with the EnKF

Given observations, a model forecast and assumptions on their respective uncertainties and at time  $t_i$ , the analyzed model states can be estimated by data assimilation using the least squares minimization (Evensen, 1994, 2003):

$$\mathbf{X}_i^a = \mathbf{X}_i^f + \mathbf{K}_i(\mathbf{Y}_i - \mathbf{H}\mathbf{X}_i^f), \quad (1)$$

where  $\mathbf{Y}_i$  is the matrix of perturbed observations,  $\mathbf{X}_i$  is the ensemble of model state vectors and  $\mathbf{H}$  is the observation



**Figure 1.** Left: bottom topography in the whole TOPAZ4 domain. The red line delimits the pan-Arctic region north of 63° N. Right: definition of sub-basins and marginal seas. The domain is divided into the four subregions delimited by the colored lines: the central Arctic in red (CA), the Greenland Sea in blue (GS), the Barents Sea in orange (BS) and the Norwegian Sea in magenta (NS).

operator denoting the projection from the model state variables to the measurements. The superscripts “a” and “f” refer to the analyzed and the forecast states, respectively. We use the deterministic form of the EnKF (DEnKF; Sakov and Oke, 2008), which solves the analysis without the requisite to perturb the observations. The term in the parentheses in Eq. (1) is the departure from the model simulations to the observations (named innovations). As opposed to Sakov et al. (2012), the 1% multiplicative inflation, which becomes problematic when used with spatially varying observational network (Anderson et al., 2001), has been removed near to the end of the reanalysis (January 2010). Multiplicative inflation leads to an exponential increase of the spread in absence of observation (such as in the interior of the Arctic Ocean). When combined with a multivariate update, it will amplify the biases of the observed variables. For instance, the passive microwave satellite images of sea ice confuse melt ponds (not considered in TOPAZ4) with open water (Ivanova et al., 2015). This results in a bias that in turn leads to a degradation of the stratification in the Arctic due to the multiplicative inflation. The bias estimation procedure has also been modified as explained below (see Sect. 2.4).

### 2.3 Assimilated observations

The observations assimilated into the reanalysis are same types as used in Sakov et al. (2012) except for some updates in the data sources. They are the satellite SST, SLA, in situ temperature and salinity profiles, SIC and low-resolution SID data from satellites. An overview of the observations used in the reanalysis is given in Table 1. The preprocessing, temporal averaging and observation errors are mostly following the procedure described in Sakov et al. (2012).

At the beginning of the time period, the assimilated SST data are the 1° resolution Reynolds SST from NOAA (Reynolds and Smith, 1994). In June 1998, they are re-

placed by the high-resolution OSTIA data (Stark et al., 2007) from the UK Met Office. The assimilated SLA data are the delayed-time product (version vxxc) from Collecte Localisation Satellites (CLS) which is validated, unfiltered and not subsampled by CLS. The SIC from the Ocean & Sea Ice Satellite Application Facility (OSISAF) are assimilated into the TOPAZ4 system. Before 19 June 2002, this assimilated product is derived from the Special Sensor Microwave/Imager (SSM/I) at 25 km resolution, and later is derived from the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) 89 GHz brightness temperature at 12.5 km resolution. In the last 3 years, this product has been upgraded to a 10 km resolution. The temperature and salinity profiles include Argo floats, ice-tethered profiles (ITPs) from the Damocles project and a large collection of hydrographic cruise data. With the exception of the Reynolds SST, all assimilated data are available through the CMEMS portal.

### 2.4 Bias estimation in the TOPAZ4 reanalysis

Two bias fields (for SST and mean sea surface height (MSSH)) are estimated online by model state augmentation, thus the analysis state of Eq. (1) is modified as

$$\begin{pmatrix} \bar{\mathbf{x}}_i^a \\ \mathbf{c}_i^a \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}}_i^f \\ \mathbf{c}_i^f \end{pmatrix} + \mathbf{K}_i (\mathbf{y}_i - \mathbf{H}\bar{\mathbf{x}}_i^f + \mathbf{H}\mathbf{c}_i^f), \quad (2)$$

where  $\bar{\mathbf{x}}_i$  is the ensemble mean of the model state vector at the analysis time  $i$ ,  $\mathbf{y}_i$  is the vector of observations and  $\mathbf{c}_i^f$  represents the estimated bias correction inherited from the analyzed bias correction at time  $i - 1$ . In order to avoid inconsistencies between assimilation of SST and temperature profile, the SST bias is propagated downwards into the model mixed layer and decays exponentially (into the  $\mathbf{H}$  operator).

The initial biases for each ensemble member are random values, homogeneous in space and uniformly distributed. The

**Table 1.** Overview of assimilated observations per cycle, with average numbers for the cycles during which the observations are present.

Type	Number	After SO	Spacing	Resolution	Period	Provider
SLA	$9 \times 10^4$	$5 \times 10^4$	Track	7 km	1992–2013	CLS
SST	$6 \times 10^3$	$6 \times 10^3$	Gridded	100 km	1990–1998	Reynolds SST from NCDC ( <a href="http://www.nhc.noaa.gov/aboutsst.shtml">http://www.nhc.noaa.gov/aboutsst.shtml</a> )
SST	$2 \times 10^6$	$2.4 \times 10^5$	Gridded	5 km	1998–2013	OSTIA from UK Met Office
In situ $T/S$	$3 \times 10^4$	$5 \times 10^3$	Point	–	1990–2013	Ifremer + other
SIC (SSM/I)	$9 \times 10^4$	$5 \times 10^4$	Gridded	25 km	1990–2002	OSISAF
SIC (AMSR-E)	$1.6 \times 10^5$	$5 \times 10^4$	Gridded	12.5 km*	2002–2013	OSISAF
SIC (AMSR-E)	$1.6 \times 10^5$	$5 \times 10^4$	Gridded	12.5 km	2008–2009	AMSR-E ( <a href="http://nsidc.org/data/amsre/">http://nsidc.org/data/amsre/</a> )
Ice drift (CERSAT)	$6 \times 10^3$	$10^3$	Gridded	35 km	2002–2010	Ifremer
Ice drift (OSISAF)	$4 \times 10^3$	$10^3$	Gridded	62.5 km	2011–2013	OSISAF
Total	$2.3 \times 10^6$	$4 \times 10^5$				

\* The resolution of ice concentration product increased to 10 km. Unless specified, all observations are from <http://marine.copernicus.eu>. NCDC is the National Climatic Data Center.

initial SST biases are sampled in the interval  $[-4, 4]$  °C, and within  $[-0.6, 0.6]$  m for the MSSH.

The bias fields are updated according to the sample covariance from the forecast ensemble, but are not integrated forward. To avoid a collapse of the bias ensembles, a multiplicative inflation is used (2 % for SLA and 6 % for SST). The multiplicative inflation of bias did not handle well the changes of observation coverage: it has been re-initialized and capped at 5 °C for SST bias in April 2001 (hereafter called event E1). Later on, in May 2006, it was re-initialized again and replaced by an additive inflation of identical amplitude (event E2), using an auto-regressive temporal process with one order, which definitively prevented further divergence. After several assimilation steps, the bias fields converge to temporally stable and spatially variable fields. Figure 2 shows the bias estimates at the end of the reanalysis for the SSH and the SST. The bias patterns compare well with those obtained in Sakov et al. (2012)<sup>1</sup>. There are small discrepancies because the bias is estimated at a different time – December 2009 in Sakov et al. (2012) instead of December 2013 here – and the bias estimation is the result of a longer estimation period for which the signal-to-noise ratio is reduced. The misfits using the online-bias-corrected values are slightly lower than the bias estimate of the last analysis step (not shown). Although the static part of the bias would theoretically be better estimated on the last assimilation of the reanalysis, the online bias approach can follow decadal trends in the errors, as well as seasonal biases and changes of the observational network. The online bias estimate is provided together with the model output. In the following validation sections, the online bias estimates  $c_i^a$  are used to offset the reanalysis state.

<sup>1</sup> Sakov et al. (2012) present the mean SSH bias of the opposite sign.

### 3 Probabilistic reliability analysis

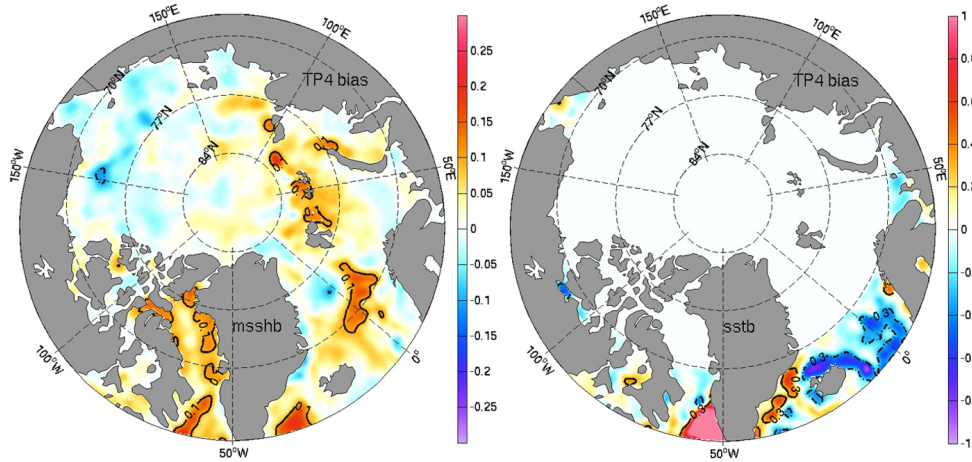
The main selling point of an ensemble data assimilation system is the probabilistic evaluation of the uncertainties, which follows the model dynamics and thus varies both in time and space. This ability comes at a risk of divergence of the Kalman filter: if the ensemble collapses, the Kalman gain tends to zero and the assimilation system behaves as one – expensive – free run. The EnKF is designed to support a very heterogeneous observational network: when observations become denser, the ensemble spread is supposed to shrink, but the forecast accuracy should be improved accordingly. However, in practice, maintaining the reliability through the course of the reanalysis requires careful analysis and handling of ill-specified model or observation error terms, and verifies that one observational data set is not “over-assimilated” at the expense of the others. Here, a simple method is used to assess the system reliability and whether the uncertainty predicted by the EnKF is commensurate with actual deviations from observations. The ensemble resolution, as well as more oceanographic interpretation of the bias, will be presented in Sect. 4.

The ensemble statistics of the assimilated variables have been stored at each assimilation time (every week) and in observational space. This allows the evaluation using the modified reduced centered random variable (RCRV; Talagrand et al., 1999; Candille et al., 2007) to measure the reliability of the TOPAZ4 system. Considering one observation  $y$  and the ensemble mean of model state  $\bar{x}^f$ , the scalar variable  $q$  can be defined as the innovation normalized by the observation and model uncertainties:

$$q = \frac{y - \mathbf{H}\bar{x}^f}{\sqrt{\sigma_o^2 + \sigma_{en}^2}}, \quad (3)$$

where  $\sigma_o$  is the observation error and  $\sigma_{en}$  is the standard deviation of the corresponding forecast ensemble, including the





**Figure 2.** Estimates of the mean SSH bias (left) and the SST bias (right) obtained at the last analyzed date by online parameter estimation. In the left panel, the solid (dashed) line indicates the 10 (–10) cm isolines. In the right panel, the solid (dashed) line indicates the 0.3 °C (–0.3 °C) isolines. There is no bias estimation for SST in the white area north of 70° N.

uncertainty of bias estimation for SLA and SST. In the framework of the Kalman filter,  $q$  is assumed to be a reduced centered Gaussian variable.

In the following, we will assess the time evolution of the averaged bias:

$$b = E[q] = \frac{1}{M} \sum_j^M \frac{y_j - \mathbf{H}\mathbf{x}^f}{\sqrt{\sigma_{oj}^2 + \sigma_{enj}^2}}, \quad (4)$$

where  $M$  is the total number of observations at the assimilation time. Furthermore, the standard deviation of  $q$ ,

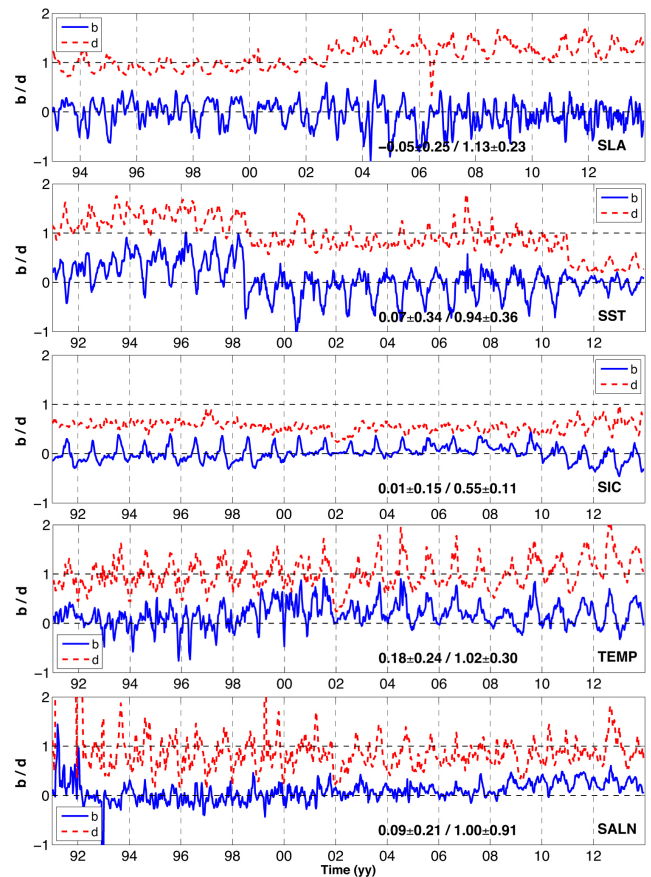
$$d = \sqrt{\frac{M}{M-1} E[(q-b)^2]}, \quad (5)$$

measures the ensemble dispersion with respect to the normalized misfits.

The first two moments of the RCRV,  $b$  and  $d$ , provide simple diagnostics of whether the forecast ensemble obtained from TOPAZ4 provides a reliable estimate of the uncertainty of the ensemble mean, which is trusted in view of the observations with the assumed uncertainties. Assuming that we can neglect all cross-covariances between innovations, a perfectly reliable system would have no bias (i.e.,  $b = 0$ ) and a dispersion equal to 1 (Candille et al., 2007). A  $d$  smaller than 1 is a sign that the assimilation system could be too optimistic about its uncertainties and vice versa. Both cases indicate that the EnKF system is not well calibrated, which in turn leads to suboptimal performance of the reanalysis system.

The two first moments of the reanalysis RCRV are presented for the different observational types. The time series of  $b$  and  $d$  in the 23 years are shown in Figs. 3 and 4.

The dispersion and seasonal bias of SLA increased after the launch of ENVISAT in 2002, when previously unobserved areas at high latitude got to be included in the calculation of the statistics. We can notice that the bias stabilizes



**Figure 3.** Time series of  $b$  (blue line) and  $d$  (dashed red line) of SLA, SST, SIC, in situ temperature and salinity observations, respectively, in the Arctic region. They are filtered by a smoothing average within 28 days. The average (standard deviation) of  $b$  and  $d$  is shown in the panels.

later on when the multiplicative inflation is replaced by the auto-regressive bias correction (event E2 in 2006).

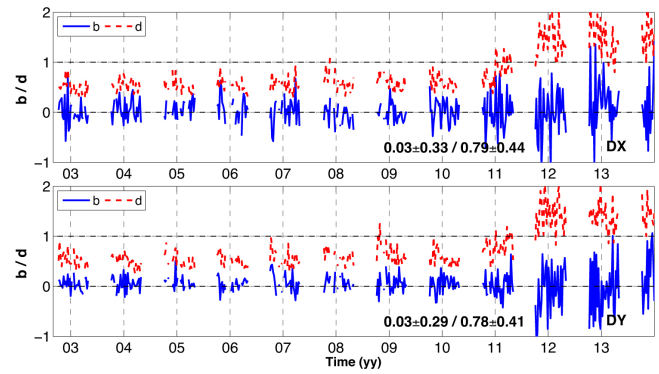
The SST panel of Fig. 3 exhibits a cold winter bias and a slight overdispersion during the time when Reynolds SST is assimilated (until 1998). The transition to OSTIA initially improves the reliability statistics with a dispersion close to 1 and a reduced bias fluctuating around 0, which relate to the changes of observation errors and land mask. The warm bias is dominant in summer. During the last 3 years of the reanalysis, the summer warm bias  $b$  is reduced but the dispersion shrinks dramatically. This coincided with the time when the observation error was increased and the quality control of the observations (based on observation uncertainty) was softened, which resulted in assimilating more observations in the Gulf Stream and near the ice edge. Although it is somewhat counterintuitive that increasing the observation error leads to a degradation of the reliability, this can happen if the misfits to the observations increase more than the model uncertainty. Furthermore, the new observation coverage includes regions close to the ice edge where the spatiotemporal interpolation of SST may have degraded the reliability (this will be further discussed in Sect. 4.2).

In the SIC panel of Fig. 3, the dispersion is underestimated throughout the reanalysis, with  $d$  on average at 0.55. The bias fluctuates around 0 with a standard deviation of 0.15 mostly related to a summer bias (Lis ater et al., 2003). A bias degradation and a dispersion improvement are jointed with clear seasonality during the last 3 years, which relates to the aforementioned change of SST assimilation settings.

The RCRVs for in situ temperatures reveal a cold bias in the reanalysis, especially salient after 1998 following developments of the observational network. A seasonal cycle in both  $b$  and  $d$  is detected during the IPY period, which may have been present before but insufficiently observed. The RCRVs for in situ salinities are initially noisy due to lack of observations. The IPY data also reveal a fresh bias as they sample regions of the central Arctic that were previously unobserved. The ensemble dispersion of salinity is good, with a tendency to be on the low side, and especially after 2002 the observation samples increase remarkably due to Argo floats.

The RCRVs for SID show initially too little dispersion ( $d = 0.56$ ) from 2002 to 2010, shown in Fig. 4 (consistent with Sakov et al., 2012). Afterward, the dispersions increase when the drag coefficient is reduced in 2011, leaving more freedom for the ice to drift following the ocean currents, but the system becomes overdispersive ( $\sim d = 1.36$ ) when the SID data source is switched from 3-day drifts on 35 km resolution to 2-day drifts on a 62.5 km resolution grid. The system shows no clear bias but the bias variability increases with the new observation product; its features will be discussed in Sect. 4.

Overall, the statistics presented are relatively stable throughout the reanalysis. There is a good balance between the different data types assimilated: none of the data types are severely less dispersed than the others. For most of the



**Figure 4.** Time series of  $b$  (blue line) and  $d$  (dashed red line) about the zonal (DX) and meridional (DY) drifts of sea ice in the Arctic. The average (standard deviation) of  $b$  and  $d$  is shown in the panels.

assimilated observation data sets, the biases fluctuate around 0 with amplitudes no larger than 0.1 (except for the in situ temperatures); the dispersions mostly fluctuate around 1 and the departures from 1 are smaller than 0.15 (except for the assimilated SIC and SID) without any sign of general ensemble collapse. However, there are some clear discontinuities caused by the introduction of new data sets with different spatial coverage (polar orbit, land mask, sea ice mask) or the related error variance adjustments. Providing a consistent reanalysis is thus challenging in the absence of continuous reprocessed observations marked with the time period.

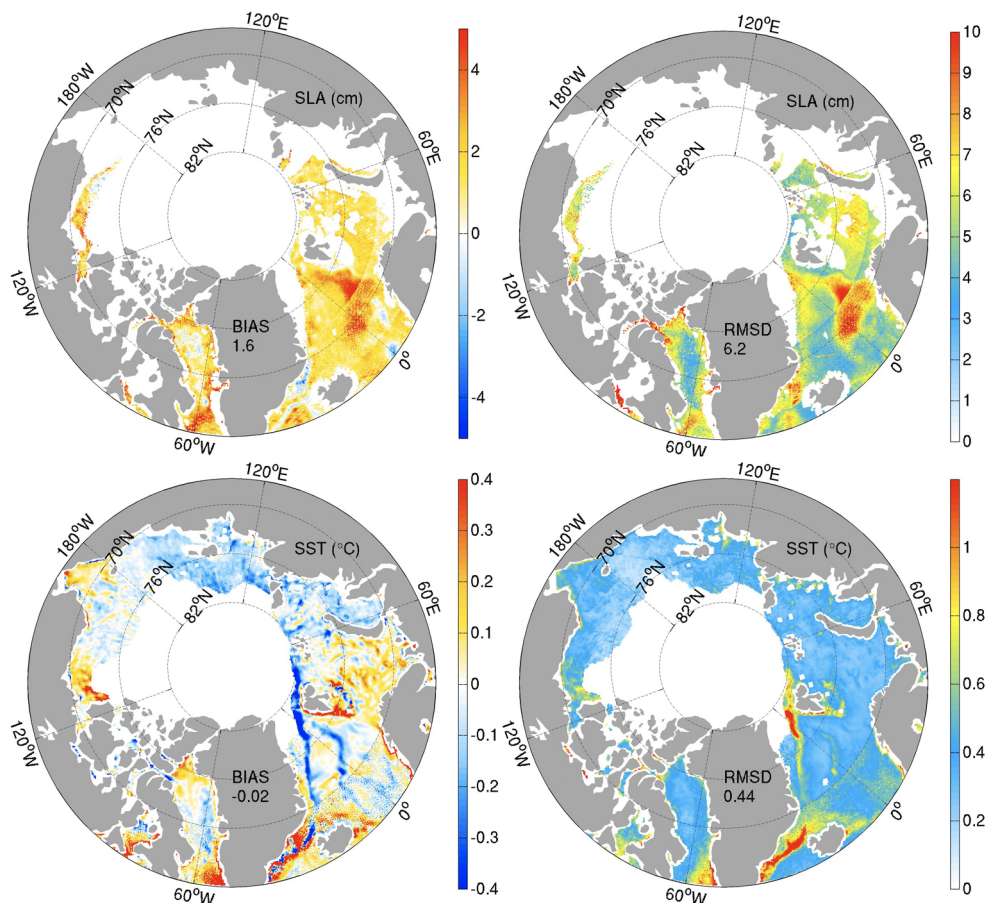
#### 4 Quantitative deterministic accuracy

In this section, we investigate whether the accuracy of the reanalysis ensemble mean (also called resolution in Candille et al., 2007) varies spatially, seasonally or interannually. Such information is necessary for potential users of the reanalysis product. It also pinpoints the model limitations that motivate further developments of modeling and assimilation approach. The misfits of the reanalysis are calculated by the daily averages of the ensemble mean and the observations. The bias and the root mean square differences (RMSDs) of the misfits are calculated as described in Eqs. (6) and (7):

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (\mathbf{H}_i \bar{\mathbf{x}}_i^f - \mathbf{y}_i - \mathbf{H} \mathbf{c}_i^f) \quad (6)$$

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{H}_i \bar{\mathbf{x}}_i^f - \mathbf{y}_i - \mathbf{H} \mathbf{c}_i^f)^2}, \quad (7)$$

where  $\bar{\mathbf{x}}_i^f$  is the forecasted daily average from the ensemble mean, which is compared to the observations  $\mathbf{y}_i$  on the same day.  $N$  is the number of times sampling was conducted over the diagnostic period (either 365 or 366 yearly). For SST and SLA, the bias term of  $\mathbf{c}_i^f$  is the online estimated correction ( $\mathbf{c}_i^f = \mathbf{c}_{i-1}^a$ , as in Eq. 2). Error bars are used to represent the standard deviations of these quantities – i.e., the variability



**Figure 5.** Top: residual bias (left) and RMSD (right) between the daily average SLA from the reanalysis and the assimilated along-track SLA data averaged over the period 1993–2013 (unit: cm). Bottom: the corresponding residual bias (left) and RMSD (right) between the daily average SST from the reanalysis and the assimilated observations averaged over the period 1999–2013 (unit: °C). Areas with less than 30 observations have been masked in white.

of the RMSD or bias estimate through the calculation period. For assimilated observations, the bias is the same as the  $b$  term in the RCRV.

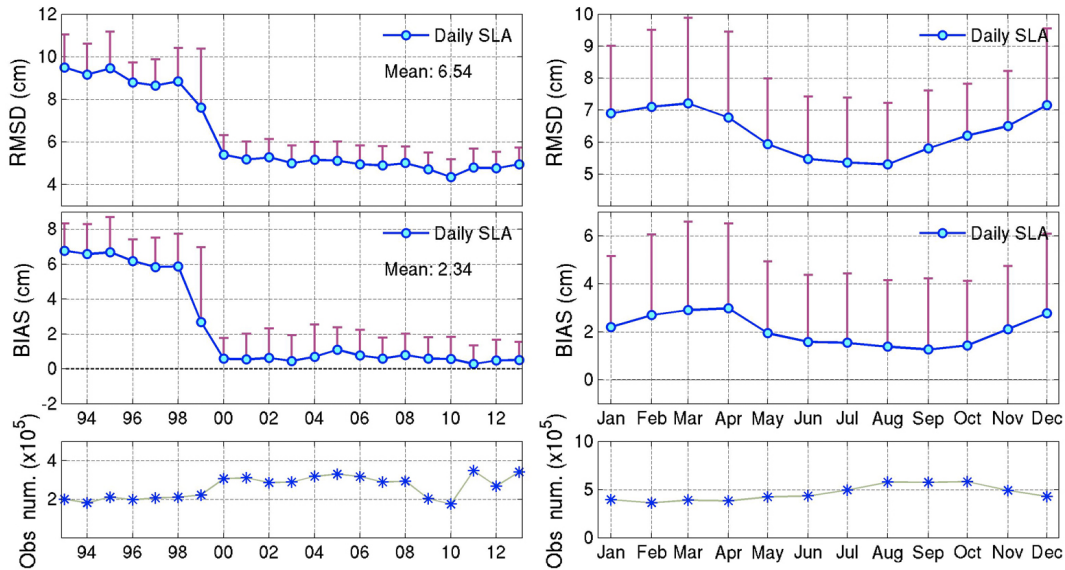
#### 4.1 Sea level anomalies

The SLA accuracy in the reanalysis is evaluated in the pan-Arctic region (defined to the north of 63° N; see Fig. 1). The spatial variability of the bias and RMSD, calculated over the whole reanalysis period (1993–2013), is shown at the top of Fig. 5. The residual bias is mainly positive, with much smaller amplitude than the estimated bias (see Fig. 2). Some positive biases reach over 4 cm around the Lofoten Basin and south of the Baffin Bay. Except for the sea ice edge in the Greenland Sea, the high RMSDs (over 9 cm) match the areas of large bias shown in Fig. 5. The spatially averaged bias is 1.6 cm, and the RMSD is about 6.2 cm.

The yearly time series of the SLA misfits and the observation number are shown on the left side of Fig. 6. The number of assimilated observations evolves with the launch or com-

pletion of satellite missions. The number of observation increases in 2000 with the launch of the GEOSAT Follow On (GFO) mission. The missions of Topex, Jason 1 and Jason 2 do not contribute directly to the pan-Arctic region as their inclination is 66°, unlike 70° for GFO. A low observation period is in 2009–2010 with the end of GFO mission (Le Traon et al., 2015), followed by an increase in 2011 with Cryosat-2, a decrease in 2012 with the end of Envisat and a last increase with the Saral/AltiKa mission in 2013. From 1993 to 2013, the RMSD decreases gradually from over 9 cm to less than 6 cm. After 2000, the residual bias stabilizes around 1 cm but remains positive. The RMSD gradually reduces with the introduction of new and more accurate observations. The reduced altimeter constellation in 2009–2010 does not cause an increase of the misfits. This demonstrates the advantage of assimilating multiple types of observations, as improved SSH may also be the result of improved SST or temperature and salinity profiles. Meanwhile, the temporal standard deviation of the RMSD during the year (shown as the half-error





**Figure 6.** Left: yearly averaged estimates of daily SLA RMSD (upper) and the residual bias (middle) of the TOPAZ reanalysis calculated against the along-track SLA available in the pan-Arctic region (unit: cm). The error bars denote the standard deviations of the daily statistics within each year. The bottom panel is the number of available observations in each year. Right: similar plot for monthly averaged estimate of daily SLA RMSD (upper), and the residual bias (middle). The error bars denote the standard deviations of the daily statistic within each month. The bottom panel shows the number of observations available for each month in the pan-Arctic during 1993–2013.

bar) also reduces from 1–2 cm to less than 1 cm, indicating the system is getting more stable with time.

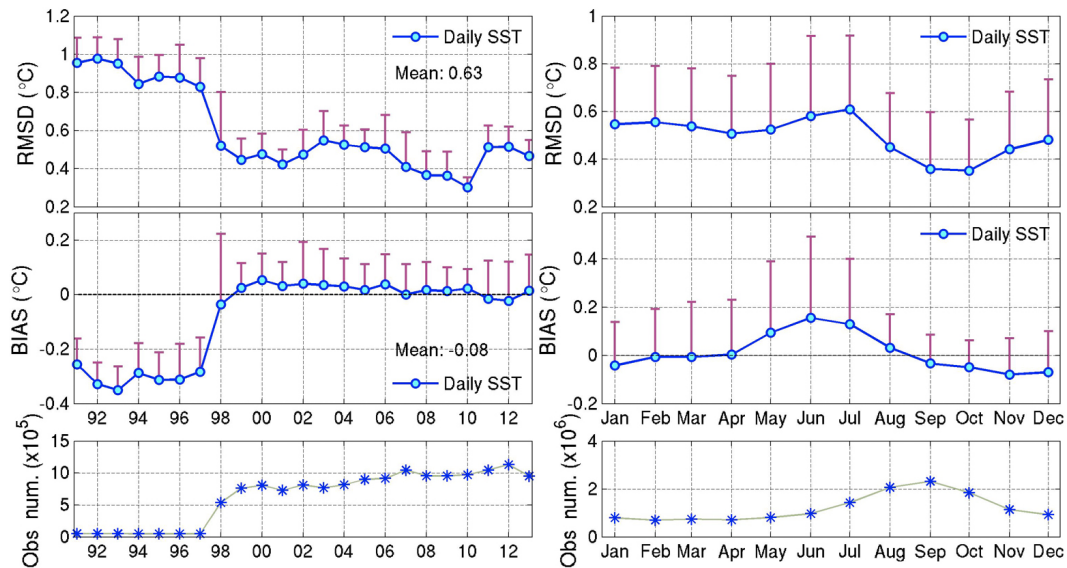
The seasonal cycle of the accuracy is shown on the right side of Fig. 6. The SLA being masked by sea ice, the number of observations varies seasonally in opposition to the sea ice cover. The RMSD ranges from 5 to 7 cm as a consequence of the seasonal spatial coverage. The residual bias is positive throughout 1 year but reaches a maximum in April. This may be explained as well by the seasonal sea ice coverage, but also by a possible underestimation of the thermal expansion. The standard deviations of the residual bias and RMSD have no visible seasonality.

## 4.2 Sea surface temperatures

The spatial variability of the SST misfits during 1999–2013 is shown at the bottom of Fig. 5. Note that SST is masked under sea ice, as done during assimilation. There are stripes of cold residual bias and high RMSD along the ice edge from north of the Svalbard Island until south of the Greenland Sea. These are contradictory to the sea ice concentration biases in the same areas in Sect. 4.4, where a cold bias corresponds with too little ice. The accuracy of SST observations near ice edge is poor and relies on strong ad hoc assumptions. Another salient feature is the warm bias ( $>0.3^{\circ}\text{C}$ ) north of the Denmark Strait where the recirculation of Atlantic water inflow is excessive in TOPAZ4 (Lien et al., 2016). This pattern was also visible in the estimated bias shown in Fig. 2, suggesting that the estimated bias accounts for most of the bias but that it still underestimates the true bias. An addi-

tional stripe of the cold residual bias and higher RMSD is clear along Mohns Ridge, also pointing to topographic steering issues. In the Barents Sea, a relative weak bias is noticeable. Besides these areas, most of the SST RMSD is lower than  $0.6^{\circ}\text{C}$ . On average, in the whole Arctic region, the SST RMSD is about  $0.44^{\circ}\text{C}$  during the period 1999–2013.

The evolution of SST accuracy of the TOPAZ4 reanalysis is shown on the left side of Fig. 7, together with the number of observations. In June 1998, the coarse-resolution Reynolds SST is swapped to the higher-resolution OSTIA SST and the number of observations increases drastically. On average, over the period 1991–2013, the SST RMSD is about  $0.63^{\circ}\text{C}$ , and the bias  $-0.08^{\circ}\text{C}$ . In the first years, the SST RMSDs are initially about  $1^{\circ}\text{C}$  but decrease gradually down to  $0.8^{\circ}\text{C}$  before 1998. During this period, the model has a cold SST bias around  $-0.3^{\circ}\text{C}$  with  $0.1^{\circ}\text{C}$  standard deviation. After the introduction of OSTIA, the SST bias settles down closer to zero, but a slight positive in summer is still noticeable before 2011. Meanwhile, the RMSD decreases rapidly below  $0.6^{\circ}\text{C}$  as a direct consequence of the bias reduction and the more abundant observations. In 2010, the RMSD reaches the minimum below  $0.4^{\circ}\text{C}$ . At that time, the ensemble spread was getting too small, and the system performance was too constrained by SST, as can be seen in the standard deviation of RMSD. It was thus decided to artificially increase the SST observation errors, which resulted in a small increase of the misfit up to  $0.5^{\circ}\text{C}$ . It is clear from the above that the transition to high-resolution SST in our system has led to a higher SST accuracy.



**Figure 7.** Same as the previous figure but for SST over the period 1991–2013 (unit: °C).

Furthermore, the seasonal performance of SST is shown in Fig. 7. As for SLA, the number of observations varies seasonally with the sea ice mask and causes the changes of the bias and RMSD. The RMSD is minimum in September and October with less than 0.4 °C owing to more observations, and is maximum at 0.6 °C in June and July when the bias is maximum as well. The reason for the larger bias in summer months is indeterminate but should relate to the inaccuracies of the mixed layer depths and the atmospheric radiative forcing.

### 4.3 In situ temperature and salinity profiles

There are  $1.1 \times 10^5$  temperature and salinity profiles assimilated in the pan-Arctic region during the period 1991–2013, but their distributions and the respective uncertainties are very uneven both in time and space, with more observations in ice-free areas and during the IPY. In order to limit variability of the uncertainty, the bias normalized by the uncertainties of the observation and model error (i.e.,  $b$  as defined in Eq. 4), is shown in Fig. 8. For temperature, there is a cold (warm) bias along the west (east) coast of the Svalbard archipelago, which indicates a northward Atlantic water flow that is too weak across the Fram Strait and a southward flow of Arctic water east of Svalbard that is too weak. There are biases that are too saline on both coasts of the Svalbard archipelago and along the Norwegian coast. They likely result from an underestimation of river discharges.

To investigate the vertical structures of the biases, the averaged temperature and salinity profiles from the reanalysis and the climatology WOA13 (Locarnini et al., 2013), and their misfits are shown in Fig. 9. The analysis is separated

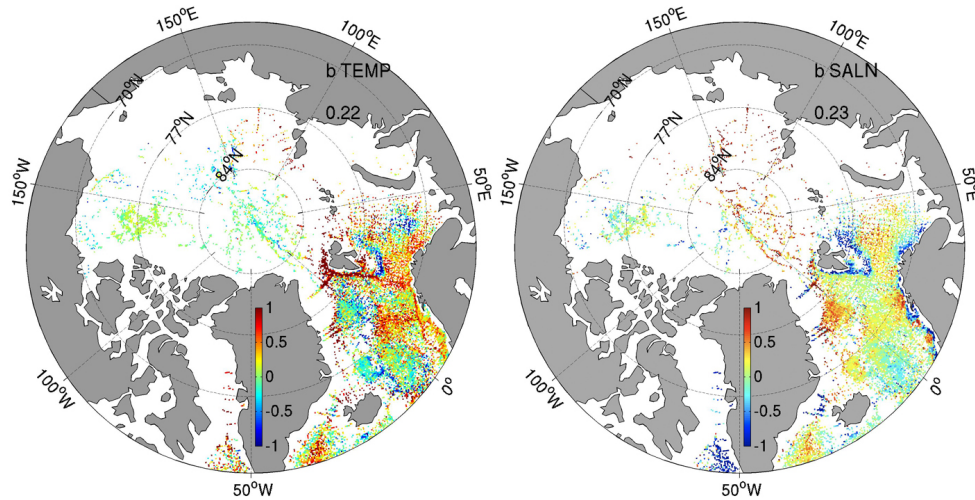
into four subregions: the central Arctic, the Barents Sea, the Greenland Sea and the Norwegian Sea (see Fig. 1).

In the central Arctic, the average profiles depict well the cold halocline water near the surface and warm saline water around 400 m associated with Atlantic water (AW). Near the surface (deeper than 200 m), the salinity misfits of TOPAZ4 are slightly smaller than the climatology. The core Atlantic water is clearly too diffuse in TOPAZ4 (not pronounced enough and vertically too broad), leading to a cold bias (−0.3 °C) and 0.5 °C RMSD around that depth. Another large RMSD is noticeable around 1000 m (0.6 °C and 0.3 psu). Since the bias at that depth is low and since the climatology has lower RMSD, it suggests that TOPAZ4 has too much variability at depths. That variability is likely due to the data assimilation setup with the combined effect of multiplicative inflation and spurious correlations (see Sect. 2.2).

In the Greenland Sea, the temperature RMSDs and biases are again slightly smaller than the climatology near the surface (upper 200 m), but degrade very near below, reaching the maxima of RMSD (> 1 °C and 0.1 psu) and bias around 800 m.

In the Norwegian Sea, the features are similar: the model has some skills near the surface but deteriorates at depths where the AW is present but is too diffuse. It is too broad and does not capture the maximum at the same depth as in the observation. It is a well-known limitation of ocean models nowadays (Ilıcak et al., 2016).

In the Barents Sea, the RMSD for temperature and salinity can be reduced near the surface, even compared to that of the climatology. But the AW (temperature > 3 °C and salinity > 35 psu, Blindheim and Østerhus, 2003) of the TOPAZ4 is too warm and saline, which suggests there is too much AW inflow or too weak a vertical mixing.



**Figure 8.** Spatial distribution of  $b$  for in situ temperature (left) and salinity (right) during the period from 1991 to 2013. The observation number in a grid is required to be more than 30. Note that profiles may end at different depths and cause spottiness.

Furthermore, we investigate the time evolution of the misfits throughout the reanalysis. Figure 10 shows the time series of the root mean square innovations (RMSIs) of temperatures and salinities in the whole Arctic at depths of 300–800 m, indicative of the Atlantic water layers. As in Sakov et al. (2012), the total uncertainty is added to assess the time reliability of the system. However, in this study, we use the formulation of  $\sigma_{\text{tot}}$  from Rodwell et al. (2016), which assumes that for a perfect reliable system RMSI is equal to  $\sigma_{\text{tot}}$ , with bias included:

$$\sigma_{\text{tot}}^2 = \text{BIAS}^2 + \sigma_{\text{en}}^2 + \sigma_0^2. \quad (8)$$

Here, the term “BIAS” refers to the innovation mean equivalent to the misfit at assimilation time.

For temperature profiles, the BIAS is negative, especially during the period of 1994–2005, indicating a warm bias at 300–800 m depths. This bias is persistent in the whole period, but reduces during the international Polar Year (IPY) period. Concurrently, the RMSI (red line in Fig. 10) also decreases after 2006. Since the reliability remains constant during the IPY (see Sect. 3), the enhanced accuracy can be considered a performance improvement, directly caused by the intensive observation efforts. The diagnosed uncertainty  $\sigma_{\text{tot}}$  (blue dashed line) and the RMSI are evolving in phase, which indicates a good potential for probabilistic forecasting. After the E2 event, the diagnosed  $\sigma_{\text{tot}}$  slightly underestimates the RMSI, which may result from the removal of the multiplicative inflation.

For salinity, the model seems too saline until the start of the IPY. The bias does not reemerge post-IPY when the number of salinity observations is very much reduced but still covers the same regions. The RMSI is also reduced during the IPY. Although there is some similarity in the evolution of the two curves, the diagnosed  $\sigma_{\text{tot}}$  is overestimating the

RMSI. This result seems to contradict the underdispersion in Fig. 3, but the difference relates to the depths at which the metrics are calculated (300–800 m here against full observation depth in Fig. 3). The cause of the overestimation stems from too large an observation error (not shown) and suggests a revision of the observation error settings for salinity profiles.

#### 4.4 Sea ice concentration

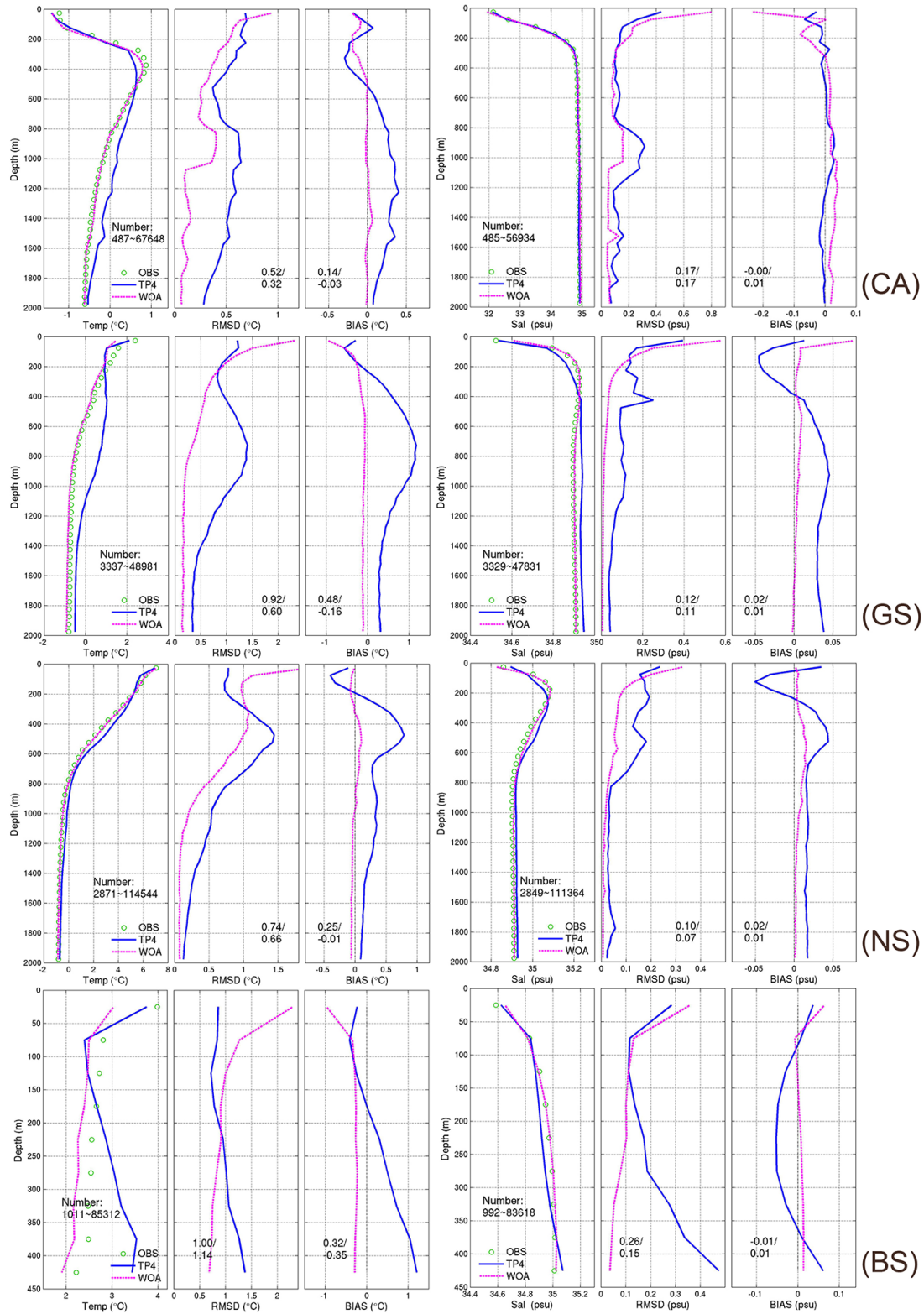
Relative to the daily sea ice concentration product from OSISAF (CMEMS OSI TAC product), the spatial variability of the SIC misfits is shown in Fig. 11. As a large seasonal variability in the sea ice extent, this is carried out at two characteristic times of one year: the maximum (March) and minimum ice extent (September).

In March, there is a dipole anomaly on either side of the ice edge in the Greenland Sea. The ice edge in TOPAZ4 is transiting too sharply from pack ice to open water because the heat capacity of the ice is neglected. This leads to a dipole bias (positive inside the ice and negative outside) during the melting season. There is also a weak bias over regions that are usually ice-free. Indeed, OSISAF does not employ weather filtering and places a thick band of low concentration (< 10 %) in ice-free regions (Ivanova et al., 2015).

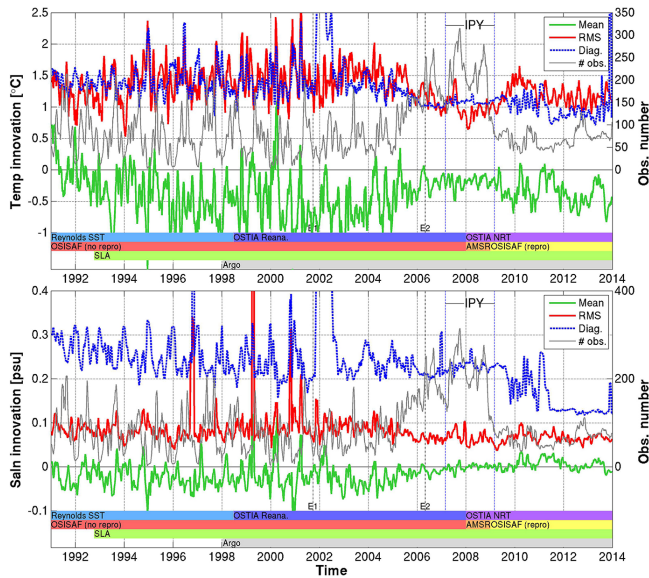
In September, TOPAZ4 shows a negative bias in the Greenland Sea. At that time of the year, the sea ice flows southwards and TOPAZ4 tends to underestimate the southern extension of the sea ice tongue along Greenland. This indicates that the dynamical forcing is biased or that the drag coefficients are incorrect as the ice is in free drift there.

The RMSD is approximately 5 % in most of Arctic region except close to the sea ice edge where the RMSD exceeds 25 %, which coincides with regions where the bias is high. Data assimilation does constrain the sea ice concentrations





**Figure 9.** The mean profiles of temperature (left) and salinity (right) and the corresponding bias and RMSD in each of the marginal seas of the pan-Arctic region. The green circles indicate the observations, the blue lines indicate the TOPAZ reanalysis and the pink lines are from the WOA13 climatology. The numbers in the first-column sub-panels are the minimal and maximal number of observations available in each of the 50 m depths; the upper numbers in the other-column sub-panels are the mean estimates in vertical for TOPAZ reanalysis, and the lower numbers are for WOA13.



**Figure 10.** Time series of innovation statistics for temperature (top) and salinity (bottom) observed at the depth between 300 and 800 m depths. The bias is plotted with a green line, the RMSD is in red and the number of assimilated observations is plotted with a grey line. The blue dashed line indicates  $\sigma_{\text{tot}}$  as defined in Eq. (8). The time series are filtered with a 28-day moving window. The vertical dashed lines indicate the change events tuning the bias correction in the course of the TOPAZ reanalysis.

but the model biases (lack of resolution of ocean currents, biases of ice drift or ice thickness) still cause locally high residual errors of ice concentrations.

In order to assess the interannual variability of the performance of TOPAZ4, we have decided to use the standard sea ice extent (SIE) metric. SIE is calculated as the surface area in which the ice concentration is larger than 15 %.

As the variability in the decadal trend of SIE in the Arctic is large, we present the interannual evolution in the whole Arctic and in two subregions: the Greenland Sea and Barents Sea (Fig. 12). TOPAZ4 shows good agreement with the OSISAF observations in the pan-Arctic region and the mean SIE in the 23 years is  $8.03 \times 10^6$  instead of  $7.96 \times 10^6$  km<sup>2</sup> in the observations. The decreasing trend of SIE during the period 1991–2013 is  $-6.16 \times 10^4$  km<sup>2</sup> yr<sup>-1</sup>, which compares well to the trend of the observations ( $-6.34 \times 10^4$  km<sup>2</sup> yr<sup>-1</sup>).

In the Greenland Sea, the SIE in TOPAZ4 is underestimated, which clearly relates to the bias in the southern extent of the sea ice tongue along the coast of Greenland. The bias in TOPAZ4 is on average  $-3.6 \times 10^4$  km<sup>2</sup> and the decreasing trend in TOPAZ4 is  $-3.1 \times 10^3$  km<sup>2</sup> yr<sup>-1</sup>, which is larger than observed ( $-2.3 \times 10^3$  km<sup>2</sup> yr<sup>-1</sup>). In the Barents Sea, the variability agrees well, although TOPAZ4 underestimates the SIE slightly. The decreasing trend is comparable.

The seasonality of the SIE in OSISAF and TOPAZ4 is investigated in Fig. 13. It is clear that the seasonal cycle of the

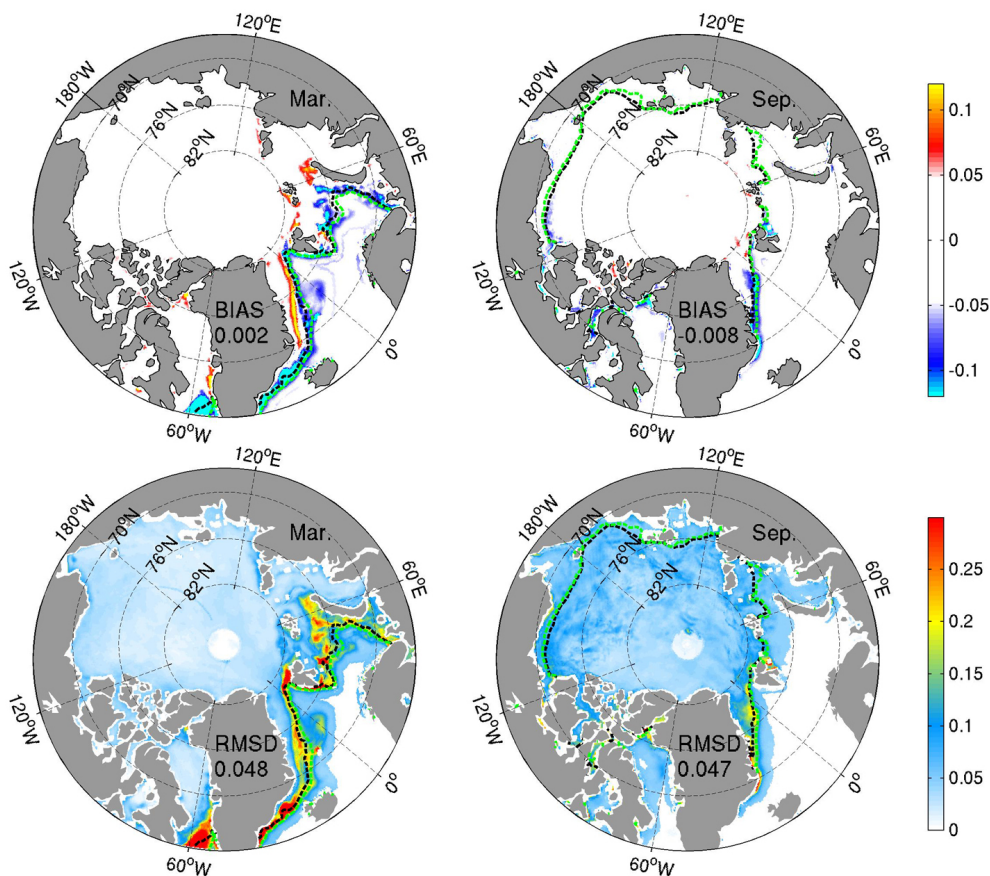
ice extent is generally well simulated by the reanalysis in the pan-Arctic area. In the summer months from June to August, a slight underestimation of the ice extent is apparent, and the minimal ice extent comes a little too early compared to the observations. In the Greenland Sea, the underestimation of sea ice extent is larger. The underestimation of sea ice extent starts in February and increases during the sea ice melt, reaching a maximum (of about  $1 \times 10^5$  km<sup>2</sup>) in July. In the Barents Sea, the seasonal cycle is well simulated but some differences are noticeable there in the beginning of the year, reaching a maximum in April, and returning to zero in August and September when there is no ice.

#### 4.5 Sea ice drift

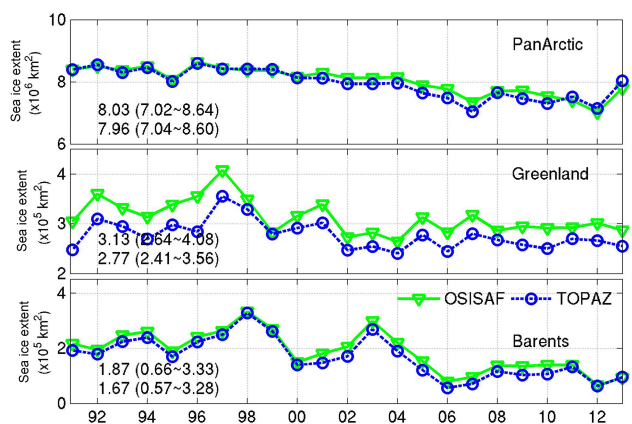
The sea ice drifts from the buoy data of the International Arctic Buoy Program (IABP) are available at 12 h frequency from 1991 to 2011. It is an independent data set and is used here for validation. To avoid the “survival bias” caused by the retreat of sea ice from the marginal seas and unresolved coastal effects, the buoy drift vectors are limited to the central Arctic, as shown with the red line in the right panel of Fig. 1. The waters shallower than 30 m and closer than 50 km to the coastline are excluded. This data set has been gridded to be compared with the model. Each grid cell is filled (i.e., considered reliable) if the calculation involves at least 30 buoys within a day. A coarser grid than the model resolution is used (four grid cells which correspond to approximately  $60 \times 60$  km<sup>2</sup>) to avoid having too many empty cells. The daily average from the measurement is the mean of the 12 h drifting speed. For comparison, the model drifting speed is calculated from the daily average of eastward and northward velocity. Several approximations are made during this comparison; we compare Eulerian to Lagrangian drift which is expected to be faster; the model ice drift is calculated from daily averages of  $u$  and  $v$  instead of daily ice drift, which is faster by approximately 0.5 km per day (not shown).

On average, over the period 1991–2011, the mean drift fields of sea ice are presented in Fig. 14. As the resulting drift estimate appeared noisy, a smoothing with the neighboring grid cells has been applied. Both observations and TOPAZ4 show a similar pattern with a pronounced Beaufort Gyre, although the center of the gyre is slightly shifted. We can also notice that TOPAZ4 globally overestimates the ice drift with a bias of 1.7 km d<sup>-1</sup>. In the Chukchi Sea, TOPAZ4 underestimates the drift by approximately  $-2$  km d<sup>-1</sup>.

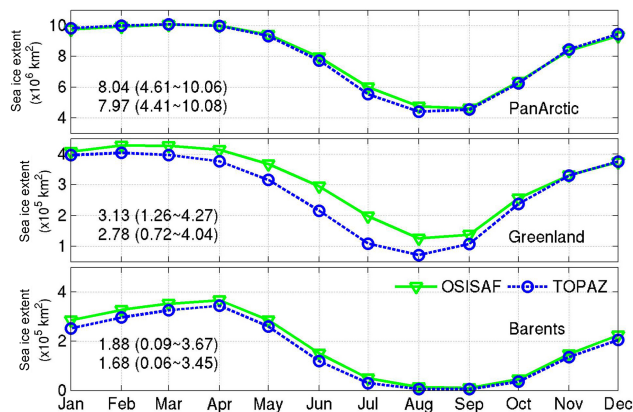
Over the period 1991–2011, the monthly time series of the ice drift speeds are compared in Fig. 15. They are averaged in the central Arctic from the reanalysis and the buoy data, respectively. On average, the drift speed is about 7 km d<sup>-1</sup> in buoy data, and about 9.4 km d<sup>-1</sup> in the TOPAZ4 reanalysis. The fast bias is clear until the end of 2010. From that time onward, the drag coefficient of the atmosphere on sea ice has been reduced from  $2.14 \times 10^{-3}$  to  $1.6 \times 10^{-3}$ . We can see that the bias is much reduced during the last year.



**Figure 11.** Spatial bias (upper) and RMSD (lower) of sea ice concentration in the TOPAZ reanalysis for March (left) and September (right), calculated from the daily averages for the period 1991–2013. The dashed black (green) lines delimit the monthly mean sea ice edges (at 15%) in the TOPAZ reanalysis (OSISAF).

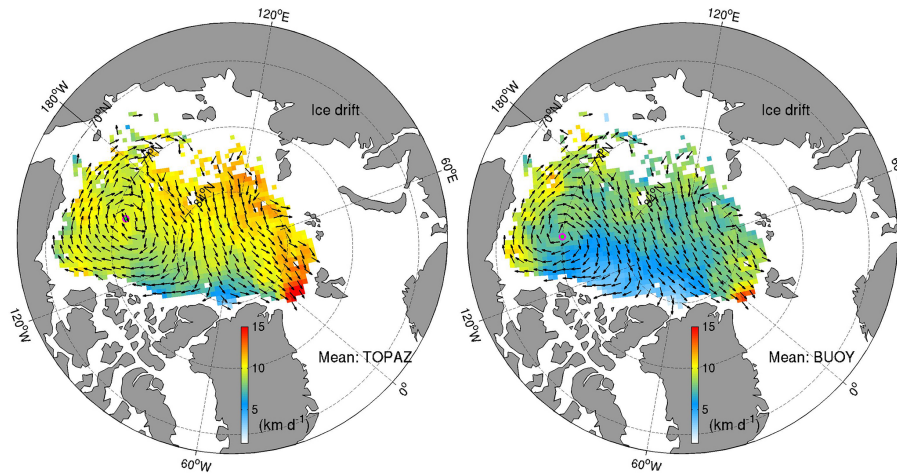


**Figure 12.** Yearly time series of the sea ice extent in the pan-Arctic region, the Greenland Sea and the Barents Sea from TOPAZ reanalysis (dashed) and OSISAF (solid).

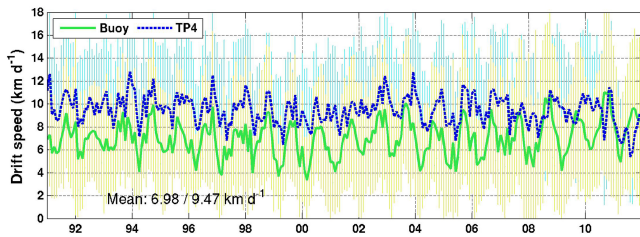


**Figure 13.** Seasonality of the sea ice extents in the TOPAZ reanalysis (blue line) and OSISAF (green line) in the pan-Arctic Ocean, Greenland Sea and Barents Sea regions.

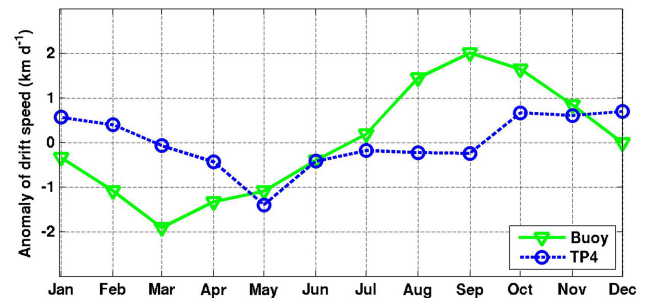




**Figure 14.** Sea ice drift vectors (arrows) and speeds (color shading) averaged over the period 1991–2011 for (left) TOPAZ reanalysis and (right) IABP buoys. The center of the anticyclonic Beaufort Gyre is marked with a magenta circle in the TOPAZ reanalysis ( $155^{\circ}$  W,  $78.1^{\circ}$  N) and in the observations ( $145^{\circ}$  W,  $77^{\circ}$  N), respectively.



**Figure 15.** Monthly time series of the daily averaged sea ice drift speeds in the central Arctic from the TOPAZ reanalysis (blue line) and the IABP buoys (green line) during 1991–2011. The error bars represent the standard deviations of the daily estimates for each month.



**Figure 16.** Seasonality of the sea ice drift velocities from the reanalysis and the buoy during 1991–2011.

The RMSD is on average  $5.1 \text{ km d}^{-1}$ , of which  $2.5 \text{ km d}^{-1}$  can be attributed to the bias. The correlation between the two curves is about 0.6.

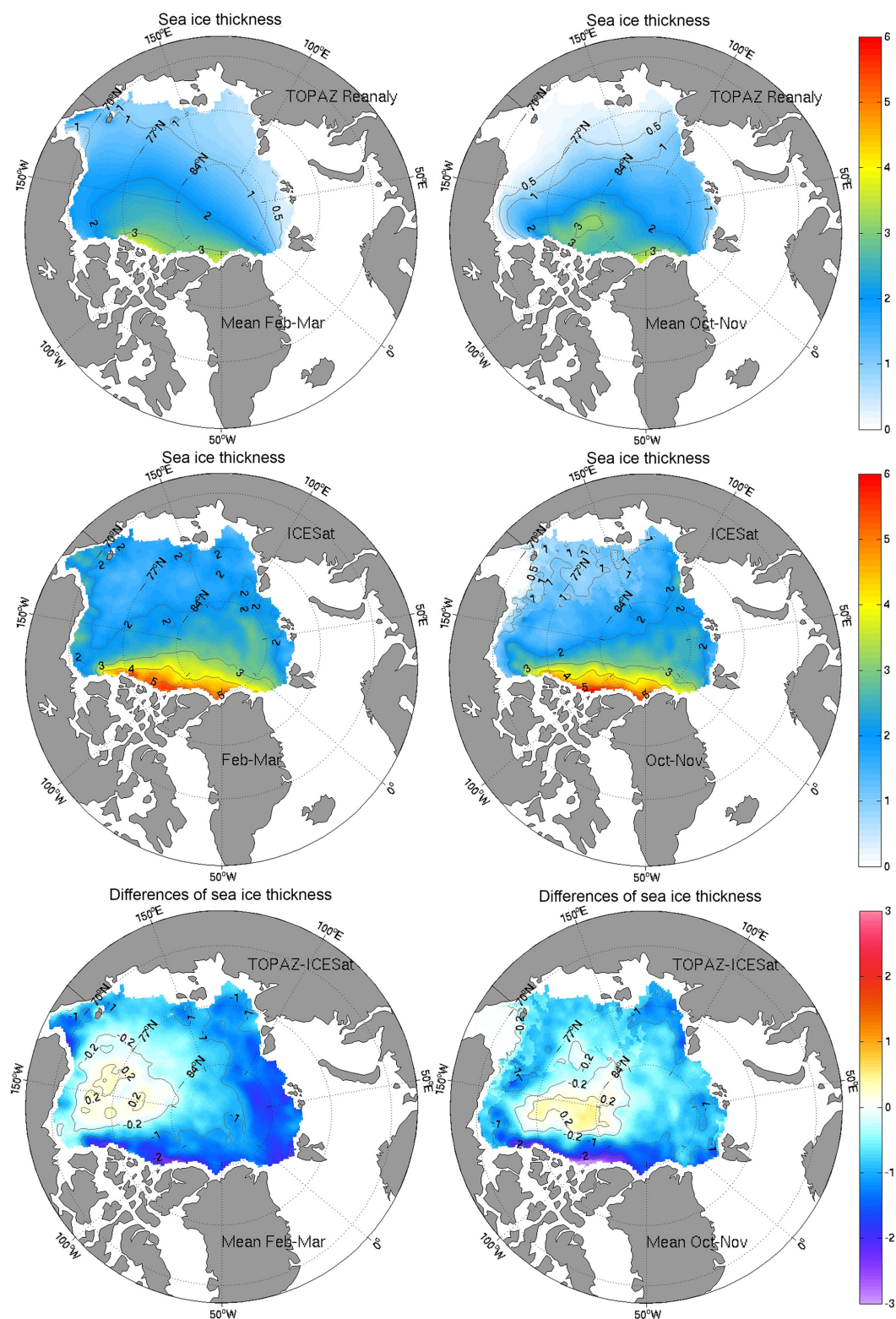
In addition, the monthly seasonality cycle of the ice drift over the period 1991–2011 is plotted in Fig. 16. While the buoys show a clear seasonality in the ice drift, being slowest in March and fastest in September, the seasonality in the TOPAZ4 reanalysis is weaker and reaches a minimum in May (delayed by 2 months).

#### 4.6 Sea ice thickness

The sea ice thickness in Arctic has attracted much attention in recent years because it has been found to be sensitive to global warming (Kwok et al., 2009; Zygmontowska et al., 2014). In this study, sea ice thickness is an independent data set, as it has not been assimilated. The observations of ice thickness with basin scale are still very few. A satellite-derived product for the Arctic Ocean ice provides the estimations of sea ice thickness for February–March and

October–November between 2003 and 2008 (ICESat, Kwok et al., 2009). Figure 17 shows the spatial distributions of the mean sea ice thicknesses and their differences. The spatial correlations are 0.74 and 0.87 for spring and fall, respectively. On average, TOPAZ4 is too thin compared to ICESat with a bias of  $-0.79$  and  $-0.64$  m in spring and in fall. In spring, TOPAZ4 is too thin, in particular north of Ellesmere Island by approximately 2 m. There is a positive bias centered in the Beaufort Gyre in spring. In fall, this bias is wider and displaced slightly to the east.

Another source of validation is the Unified Sea Ice Thickness Climate Data Record (Lindsay, 2013) resulting from a concerted effort to collect as many observations as possible of Arctic sea ice draft, freeboard and thickness. The sea ice draft is measured by the sonar of US Navy Submarines from National Snow and Ice Data Center (USSUB-DG and USSUB-AN; Wadhams and Horne, 1980; Wensnahan and Rothrock, 2005; Rothrock and Wensnahan, 2007), and the sea ice thickness by flight campaigns from NASA Operation IceBridge (IceBridge; Kurtz et al., 2013), as shown



**Figure 17.** Mean sea ice thicknesses from TOPAZ (upper) and ICESat (middle), and their difference (bottom) for February–March (left column) and October–November (right column) averaged over the period 2003–2008.

in Fig. 18a. The sea ice draft data have been diagnosed in TOPAZ4 as proposed by Eq. (4) of Alexandrov et al. (2010):

$$D_i = H_i \cdot \frac{\rho_i}{\rho_w} + H_{sn} \cdot \frac{\rho_{sn}}{\rho_w}, \quad (9)$$

where  $D_i$  is ice draft,  $H_i$  is ice thickness and  $H_{sn}$  is the snow thickness. The  $\rho_i$ ,  $\rho_w$  and  $\rho_{sn}$  are the densities for sea ice, water and snow (respectively, 900, 1000 and 300 kg m<sup>-3</sup>).

The IceBridge ice thickness covers the period of 2009–2011. TOPAZ4 reanalysis is too thin with a bias of 1.1 m, a RMSD of 1.4 m and a correlation of 0.5. The bias against the sea ice draft is smaller with 0.3–0.4 m, and a RMSD about 0.6–0.7 m. The correlation coefficients are relatively good with 0.86 and 0.69, which is higher than for the IceBridge data. These discrepancies are likely to be related to the spatial distribution of the different data sets. Hence, IceBridge data are concentrated around the northern coast of Greenland where TOPAZ4 showed largest bias in the comparison with ICESat.

As another diagnostic of interest, the daily time series of sea ice volume from TOPAZ4 in the Arctic in 1991–2013 is shown by the blue curve in the left panel of Fig. 19. Before 2001, the sea ice volume varies stably around  $1.4 \times 10^4$  km<sup>3</sup>, with a significant seasonal variability between  $8 \times 10^3$  km<sup>3</sup> and  $1.9 \times 10^4$  km<sup>3</sup>. Afterwards, in the period 2001–2010, the sea ice volume decreases dramatically. This reduction of sea ice volume is qualitatively consistent with the limited satellite records. First, the estimate from Kwok et al. (2009), derived from the ICESat record from 2003 to 2008, shows a similar trend. After revising the uncertainties of input data (snow depth, sea ice density and ice concentrations), Zygmuntowska et al. (2014) corrected the estimates of the mean sea ice volume, shown as the starred line in Fig. 18. With respect to these sea ice volume estimates, TOPAZ4 still has too little ice. In the right panel of Fig. 19, the seasonal cycles of sea ice volume from TOPAZ4 and the standard deviation in the 23 years are shown by the blue curve and the cyan error bars, respectively. In May, the maximum sea ice volume is about  $1.5 \times 10^4$  km<sup>3</sup>, and in September is less than  $5 \times 10^3$  km<sup>3</sup>. The sea ice volumes from Zygmuntowska et al. (2014) are plotted on top of the averaged TOPAZ4 seasonal cycle in the period 1991–2013. These correspond well to the model climatology, but still betray an underestimation because the measurements are representative of a period of lower ice volume.

The TOPAZ4 seasonal cycle of ice volume seems to change in amplitude during different time eras, although the reasons lie in two successive changes of the settings of the EnKF. In December 2001, the variance of precipitation errors is increased from  $1.10^{-17}$  to  $1.10^{-12}$  m<sup>2</sup> s<sup>-2</sup>, as an adjustment for a slow decrease of ensemble spread. These perturbations being truncated Gaussian, the truncation resulted in excessive snow precipitations. The excessive snow depths have then isolated the ice from the atmosphere and reduced the amplitude of the yearly cycle from 1.08 to 0.74 m (see

Fig. 20); this also delayed the phase of the cycle. In January 2011, an unbiased log-normal law replaced the truncated Gaussian perturbations with an amplitude of 30 %. The amplitude and phase of the seasonal cycle returned to more correct values. The sensitivity experiments in Finck et al. (2013) verified the above-mentioned issue.

## 5 Summary and discussions

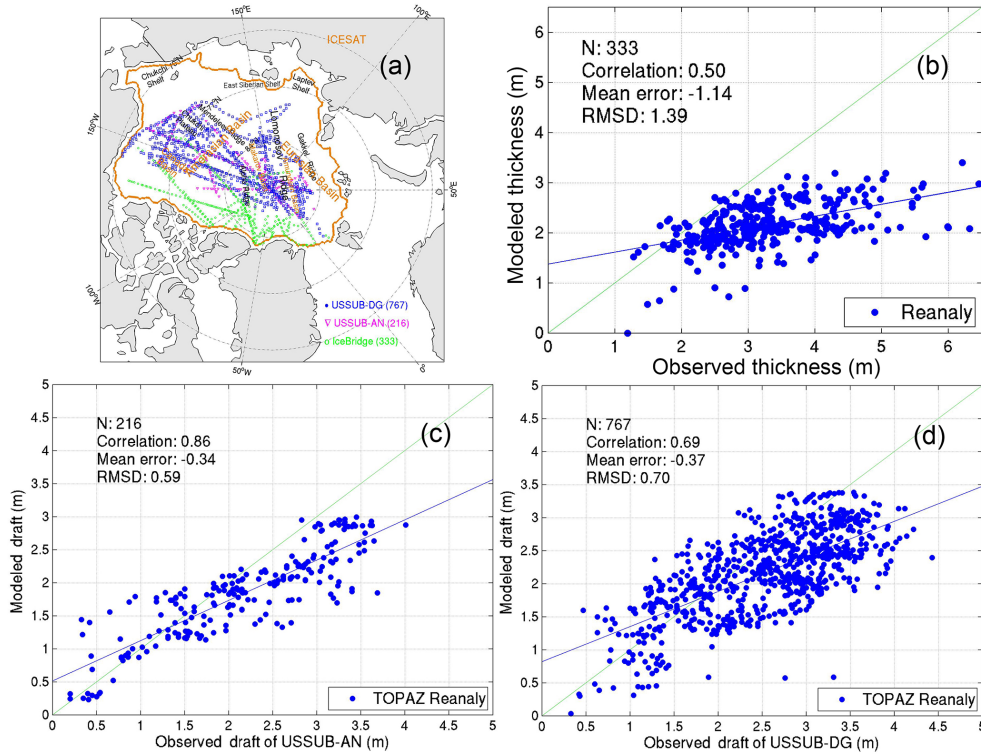
This study is conducted to present and validate the official physical multi-year CMEMS product for the Arctic region. The proposed reanalysis is unique compared to other reanalysis products (see Table 1 of Chevallier et al., 2016). It proposes a long high-resolution dynamical reconstruction of the ocean and sea ice, and assimilates a complete set of observations available in the Arctic region with an advanced ensemble data assimilation method and with strongly coupled data assimilation between ocean and sea ice. The above results present a concise account of the strengths and weaknesses of the resulting data set. The above findings can be summarized variable by variable:

**SLA** In the period 1993–2013, the RMSD of daily SLA in the reanalysis is gradually decreased from over 9 cm to less than 6 cm in the pan-Arctic region. The introduction of a bias estimation scheme proves very efficient in constraining the bias. The largest RMSDs over 9 cm are found around the Lofoten Basin. There is also a patch of larger misfit near the ice edge, but observations are also less accurate there. There is a weak seasonality in the performance of the system with the best results in the summer. The system is slightly overdispersive mostly due to bias estimation.

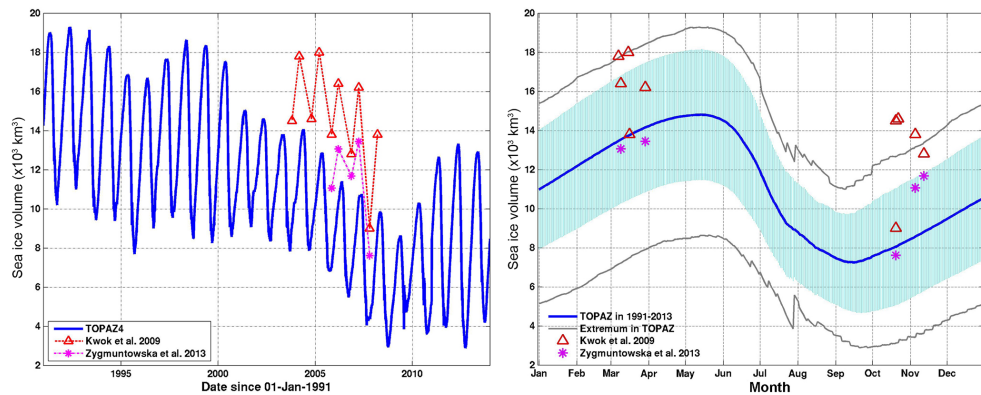
**SST** The SST RMSD is about 0.63 °C over the period 1991–2013, and after 1999 it is reduced to about 0.44 °C with a smaller bias around  $-0.02$  °C. The transition to high-resolution OSTIA SST is highly beneficial for constraining the bias and the RMSD, but an overestimation of the observation error from the provider was needed to avoid a collapse of the ensemble spread. The performance of the system varies seasonally following the observational amounts and a larger bias during summer months. The system dispersion is close to 1 in most of the years but can be over- or underdispersive depending on the settings of observation errors and bias estimation.

**Temperature and salinity profiles** The misfits of the reanalysis are small near the surface (in the top 100 to 200 m), even compared to those of the WOA13 climatology. Below this depth, the model shows large biases and performs poorer (RMSD > 1 °C and about 0.1 psu). Some of the biases relate to the limitations of the model to maintain the Atlantic water (as expected from Ilıcak et al., 2016) and others relate to a degradation intro-





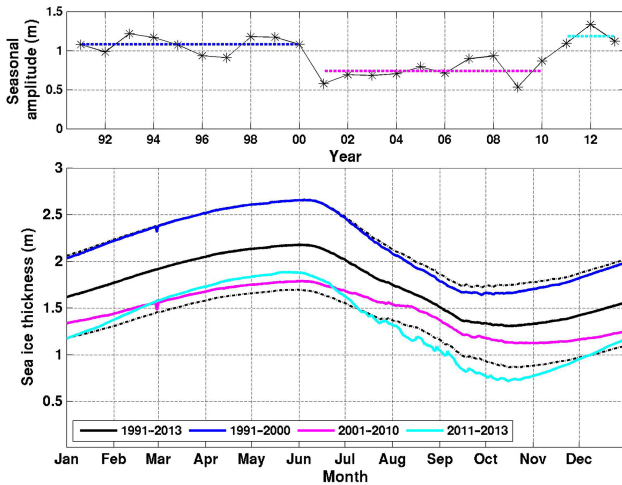
**Figure 18.** Validation the sea ice thickness in the TOPAZ reanalysis versus available in situ observations. (a) Locations of in situ observations available from IceBridge, USSUB-AN and USSUB-DG in the central Arctic. Regression analysis of TOPAZ reanalysis (b) vs. IceBridge; (c) vs. USSUB-AN; (d) vs. USSUB-DG.



**Figure 19.** Left: time series of the daily averaged sea ice volume in the Arctic from the TOPAZ4 (blue line) and the observations from Kwok et al. (2009) and from Zyguntowska et al. (2014). Right: daily time series of the averaged sea ice volume in the Arctic from the TOPAZ4 for the period 1991–2013 (blue line) and the standard deviation shown as the cyan error bar. The grey lines represent the extreme volumes in the 23 years. The triangle and star markers are the observations estimated by Kwok et al. (2009) and Zyguntowska et al. (2014), respectively.

duced by data assimilation (a flat multiplicative inflation). A large improvement occurred at the times when the inflation method was upgraded and when there were more available observations during the IPY. The system reliability is overall stable in time, in spite of the very inhomogeneous data sampling over the past 23 years.

**Sea ice concentration and extent** TOPAZ4 agrees well with the OSI-SAF sea ice concentrations. On average, the RMSDs are lower than 5% and the biases close to zero. The misfits are larger close to the ice edge, and poorest in the Greenland Sea. The errors are related to biases in the thermodynamics and dynamics of the sea ice model. The bias is largest during the



**Figure 20.** Top: yearly time series of the seasonal amplitudes of the mean sea ice thickness in the central Arctic, shown with the solid black line. The dashed lines represent the averaged estimate for 1991–2000, 2001–2010 and 2011–2013 (1.08, 0.74 and 1.18 m, respectively). Bottom: daily time series of the mean sea ice thickness in the central Arctic for three different time periods. The black dashed lines denote the standard deviation for the 23 yearly estimates.

summer season. The performance is stable throughout the reanalysis but the dispersion is consistently too low ( $d = 0.55$ ), probably due to a too rudimentary thermodynamical sea ice model.

**Sea ice drift** The averaged drift in TOPAZ4 shows comparable patterns to independent observation from IABP buoys with the classical Beaufort Sea gyre and transpolar drift. However, the center of the gyre is slightly misplaced. The RMSD of drift speed in the reanalysis is about  $5.1 \text{ km d}^{-1}$ , and has a fast bias by about  $2.5 \text{ km d}^{-1}$ . The monthly time variability compares well, but TOPAZ4 has too weak a seasonal cycle and shifted by 2 months. From 2011 onwards, the atmospheric drag coefficient was adjusted and the ice drift speed agrees better with observations after the change. Still, with RMSDs of  $5 \text{ km d}^{-1}$  close to the signal itself, improving the performance of ice drift appears a priority for future operational use. The dispersion is also low but becomes too large after switching to a different observational product.

**Sea ice thickness** TOPAZ4 shows some large biases (approximately  $-1.1 \text{ m}$ ) compared to ice thickness from ICESat and IceBridge as well as compared to ice draft data, although the thick ICESat ice draft may have been overestimated (Khvorostovsky and Rampal, 2016). The thickness bias is largest north of Ellesmere Island with bias up to 2 m. The spatial pattern and regression compare reasonably well. The ice is too thin in the period

2001–2010 due to excessive snow depths and the seasonal cycle is too small during that time.

RCRV diagnostics have shown a good balance between the different data types assimilated: none of the data types are severely less dispersed than the others. The results from the 23-year reanalysis show overall a reasonable stability over time and good agreements with observations. However, some clear discontinuities are caused by transitions from one data set to other new observations in areas that were completely unobserved, and also by changes in the data assimilation settings. Assessing the system for such a long period also reveals some limitations that are either inherent to the data assimilation implementation or due to model flaws. In the following, we list the possible reasons and the means to tackle these in the future version of the ARC MFC system.

- The Atlantic waters have a signature that is too diffused. In order to improve their advection, we will double the horizontal and vertical resolution (50 hybrid layers and 5 km horizontal resolution). The parameterization of diapycnal mixing will be reduced under sea ice as proposed in Morison et al. (1985). We also foresee that increasing the resolution will be well useful for resolving the circulation in the Nordic Seas and reduce the seasonal biases of SST and SSH.
- The system has too sharp an ice edge. The current thermodynamic model does not account for the heat capacity of the sea ice. TOPAZ will be upgraded to the community sea ice mode CICE (Hunke et al., 2010), which uses a complex thermodynamic parameterization.
- Observations detect melt ponds as open water, whereas melt ponds are not simulated in the current TOPAZ4. This creates bias in sea ice during summer months that is transferred to the interior of the ocean via coupled data assimilation. In the future, we will choose the best alternative between using an existing melt pond model or detect and remove the signature of the melt ponds from the observations.
- Comparisons against sea ice drift and ice thickness highlighted more severe limitations: ice that is too thin, a thickness gradient that is too smooth from Greenland into the Beaufort Gyre; the center of the Beaufort Gyre being slightly misplaced, the sea ice drift being too fast. These biases can be reduced by optimizing the sea ice strength ( $P^*$ ) and the drag parameters both in ocean and atmospheric (Massonnet et al., 2014). However, optimal values of these parameters are moving targets in view of their limited physical realism. The methodology proposed by Barth et al. (2015), to estimate biases in atmospheric wind from ice drift will also be considered. But the RMSDs of ice drift are relatively high ( $5 \text{ km d}^{-1}$  for an ice drift generally inferior to  $10 \text{ km d}^{-1}$ ) although

comparable to short-term forecasts in Schweiger and Zhang (2015). These fluctuating misfits are less likely to be reduced by model tuning.

- There are further indications that the viscous–plastic and the related elastic–viscous–plastic rheologies have inherent limitations for simulating long-term properties of the ice drift – e.g., the acceleration of sea ice drift, the phase of its seasonal cycle (Rampal et al., 2011). A high-priority objective is therefore to couple TOPAZ to the neXtSIM sea ice model that is based on an elasto-brittle rheology. Recent studies with a forced version of neXtSIM (Bouillon and Rampal, 2015; Rampal et al., 2016) suggest that the model is capable of reproducing the sea ice deformations over a wide range of spatial and temporal scales and reduces the error of the sea ice drift. It is of interest to understand to which extent the coupling feedback will respond to this improved dynamical model.
- The online bias estimation appeared quite successful to limit bias in our model, but its implementation in the EnKF was very sensitive to the choice of inflation method used. The latest configuration that combined  $r$  factor inflation and autoregressive additive inflation for parameters is our recommendation in a realistic system with a strongly variable observation network.
- The EnKF has proven capable to assimilate a large variety of observations, but more observations should be assimilated, like the sea ice thickness of thin ice from the European Space Agency’s (ESA) Soil Moisture and Ocean Salinity (SMOS) in Kaleschke et al. (2012) and Tian-Kunze et al. (2014). Also the complementary thickness of thick ice from ICESat (Kwok et al., 2009; Khvorostovsky and Rampal, 2016) and CryoSat-2 (Wingham et al., 2006; Laxon et al., 2013), and SMOS sea surface salinity (Reul et al., 2012) will be tested in order to determine how to better assimilate into the system in the near future.
- Although efforts were made to freeze as much of the assimilation setting as possible, some change have been necessary: e.g., replacing the multiplicative inflation by additive inflation or changes of observation product. These have caused discontinuities in the accuracy and in the reliability of the system. These discontinuities may become problematic for the interpretation of mechanisms of variability in the Arctic. For optimizing its consistency, a reanalysis should limit its observation network to that available through the whole reanalysis period, as done in Counillon et al. (2016) with assimilation of SST only. However, such a type of reanalysis prioritizes consistency at the expense of accuracy, which is not the purpose of the TOPAZ system. In a future reanalysis production, consistently reprocessed data sets

from the ESA climate change initiatives (ESA CCIs) will be assimilated over the whole period (these were not available at the start of this reanalysis). The monitoring of reliability metrics can be automated and the results presented here indicate that the reliability should then remain stable.

- The next physical ARC MFC reanalysis will provide a stochastic product, in order to provide a natural framework for estimating the system accuracy in space and time, and to provide input data for probabilistic weather or stand-alone sea ice models.

## 6 Data availability

The reanalysis data used in this paper are freely available from CMEMS (<http://marine.copernicus.eu>) under the product named ARCTIC\_REANALYSIS\_PHYS\_002\_003. The ensemble statistics are estimated by the diagnosing files, which can be obtained from the authors. Assimilated observations are listed in the text; additional in situ ocean temperature and salinity profiles are from IOPAS, ICES, MMBI and AARI and quality-checked by Alexander Korablev.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* Thanks to P. Rampal for processing the buoy data set for sea ice drift and for the useful discussions. We thank to the US National Snow and Ice Data Center (NSIDC) for providing the IceBridge data. This study was supported by successive MyOcean projects from the European Commission (grant number 218812), the Arctic element of the Copernicus Marine Services and a grant of CPU time from the Norwegian Supercomputing Project (NOTUR II grant number nn2993k). We thank two anonymous reviewers for constructive suggestions that have improved this manuscript.

Edited by: E. J. M. Delhez

Reviewed by: two anonymous referees

## References

- Alexandrov, V., Sandven, S., Wahlin, J., and Johannessen, O. M.: The relation between sea ice thickness and freeboard in the Arctic, *The Cryosphere*, 4, 373–380, doi:10.5194/tc-4-373-2010, 2010.
- Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, 129, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2, 2001.
- Barth, A., Canter, M., Schaeybroeck, B. V., Vannitsem, S., Massonet, F., Zunz, V., Mathiot, P., Alvera-Azcárate, A., and Beckers,

- J.-M.: Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the Southern Ocean, *Ocean Model.*, 93, 22–39, doi:10.1016/j.ocemod.2015.07.011, 2015.
- Bertino, L. and Lisæter, K. A.: The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans, *Journal of Operational Oceanography*, 1, 15–19, doi:10.1080/1755876X.2008.11020098, 2008.
- Bleck, R.: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Model.*, 4, 55–88, doi:10.1016/S1463-5003(01)00012-9, 2002.
- Blindheim, J. and Østerhus, S.: The Nordic Seas, Main Oceanographic Features, in: *The Nordic Seas: An Integrated Perspective*, edited by: Drange, H., Dokken, T., Furevik, T., Gerdes, R., and Berger, W., Geoph. Monog. Series, 158, 11–37, 2003.
- Bouillon, S. and Rampal, P.: Presentation of the dynamical core of neXtSIM, a new sea ice model, *Ocean Model.*, 91, 23–37, doi:10.1016/j.ocemod.2015.04.005, 2015.
- Brasseur, P., Baharel, P., Bertino, L., Birol, F., Brankart, J.-M., Ferry, N., Losa, S., Remy, E., Schröter, J., Skachko, S., Testut, C.-E., Tranchant, B., Van Leeuwen, P. J., and Veron, J.: Data assimilation for marine monitoring and prediction: The MERCATOR operational assimilation systems and the MERSEA developments, *Q. J. Roy. Meteor. Soc.*, 131, 3561–3582, doi:10.1256/qj.05.142, 2005.
- Candille, G., Côté, C., Houtekamer, P. L., and Pellerin, G.: Verification of an Ensemble Prediction system against observations, *Mon. Weather Rev.*, 135, 2688–2699, doi:10.1175/MWR3414.1, 2007.
- Chassignet, E. P., Smith, L. T., and Halliwell, G. R.: North Atlantic Simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity, *J. Phys. Oceanogr.*, 33, 2504–2526, doi:10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO;2, 2003.
- Chevallier, M., Salas-Mélia, D., Voltaire, A., and Déqué, M.: Seasonal forecasts of the Pan-Arctic sea ice extent using a GCM-based seasonal prediction system, *J. Climate*, 26, 6092–6104, doi:10.1175/JCLI-D-12-00612.1, 2013.
- Chevallier, M., Smith, G., Lemieux, J.-F., Dupont, F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Zuo, H., Balmaseda, M., Chang, Y.-S., Ferry, N., Garric, G., Haines, K., Keeley, S., Kovach, R. M., Kuragano, T., Masina, S., Tang, Y., Tsujino, H., and Wang, X.: Intercomparison of the Arctic sea ice cover in global ocean-sea ice reanalyses from the ORA-IP project, *Clim. Dynam.*, 1–30, doi:10.1007/s00382-016-2985-y, 2016.
- Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., and Bentsen, M.: Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model, *Tellus A*, 68, 32437, doi:10.3402/tellusa.v68.32437, 2016.
- Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M., Mogensén, K., Oke, P., Testut, C. E., Verron, J., and Weaver, A.: Ocean data assimilation systems for GODAE, *Oceanography*, 22, 96–109, doi:10.5670/oceanog.2009.69, 2009.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Drange, H. and Simonsen, K.: Formulation of air-sea fluxes in the ESOP2 version of MICOM, Technical Report No. 125, Nansen Environmental and Remote Sensing Center, 23 pp., 1996.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10143–10162, doi:10.1029/94JC00572, 1994.
- Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation, *Ocean Dynam.*, 53, 343–367, doi:10.1007/s10236-003-0036-9, 2003.
- Finck, N., Counillon, F., Bertino, L., Bouillon, S., and Rampal, P.: Validation of sea ice quantities of TOPAZ for the period 1990–2010, Technical Report No. 332, Nansen Environmental and Remote Sensing Center, 30 pp., 2013.
- Haine, T., Curry, B., Gerdes, R., Hansen, E., Karcher, M., Lee, C., Rudels, B., Spreen, G., Steur, L., Stewart, K. D., and Woodgate R.: Arctic freshwater export: Status, mechanisms, and prospects, *Global Planet. Change*, 125, 13–35, doi:10.1016/j.gloplacha.2014.11.013, 2015.
- Hunke, E. C. and Dukowicz, J. K.: An elastic-viscous-plastic model for sea ice dynamics, *J. Phys. Oceanogr.*, 27, 1849–1867, doi:10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.
- Hunke, E. C., Lipscomb, W. H., and Turner, A. K.: CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual Version 4.1 LA-CC-06-012, T-3 Fluid Dynamics Group, Los Alamos National Laboratory, Los Alamos NM 87545, 76 pp., 2010.
- Ilicak, M., Drange, H., Wang, Q., Gerdes, R., Aksenov, Y., Bailey, D., Bentsen, M., Biastoch, A., Bozec, A., Böning, C., Casou, C., Chassignet, E., Coward, A. C., Curry, B., Danabasoglu, G., Danilov, S., Fernandez, E., Fogli, P. G., Fujii, Y., Griffies, S. M., Iovino, D., Jahn, A., Jung, T., Large, W. G., Lee, C., Lique, C., Lu, J., Masina, S., George Nurser, A., Roth, C., Salas y Mélia, D., Samuels, B. L., Spence, P., Tsujino, H., Valcke, S., Voltaire, A., Wang, X., and Yeager, S. G.: An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part III: Hydrography and fluxes, *Ocean Model.*, 100, 141–161, doi:10.1016/j.ocemod.2016.02.004, 2016.
- Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, *The Cryosphere*, 9, 1797–1817, doi:10.5194/tc-9-1797-2015, 2015.
- Johannessen, O. M., Bengtsson, L., Miles, M. W., Kuzmina, S. I., Semenov, V. A., Alekseev, G. V., Nagurny, A. P., Zakharov, V. F., Bobylev, L. P., Pettersson, L. H., Hasselmann, K., and Cattle, H. P.: Arctic climate change – observed and modelled temperature and sea-ice variability, *Tellus A*, 56, 328–341, doi:10.1111/j.1600-0870.2004.00060.x, 2004.

- Kaleschke, L., Tian-Kunze, X., Maaß, N., Mäkynen, M., and Drusch, M.: Sea ice thickness retrieval from SMOS brightness temperatures during the Arctic freeze-up period, *J. Geophys. Lett.*, 39, L05501, doi:10.1029/2012GL050916, 2012.
- Kara, A., Rochford, P. A., and Hurlburt, H. E.: Efficient and accurate bulk parameterizations of air-sea fluxes for use in general circulation models, *J. Atmos. Ocean. Tech.*, 17, 1421–1438, doi:10.1175/1520-0426(2000)017<1421:EAABPO>2.0.CO;2, 2000.
- Khvorostovsky, K. and Rampal, P.: On retrieving sea ice freeboard from ICESat laser altimeter, *The Cryosphere*, 10, 2329–2346, doi:10.5194/tc-10-2329-2016, 2016.
- Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, *The Cryosphere*, 7, 1035–1056, doi:10.5194/tc-7-1035-2013, 2013.
- Kwok, R. and Rothrock, D.: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008, *Geophys. Res. Lett.*, 36, L15501, doi:10.1029/2009GL039035, 2009.
- Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, *J. Geophys. Res.*, 114, C07005, doi:10.1029/2009JC005312, 2009.
- Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, *Geophys. Res. Lett.*, 40, 732–737, doi:10.1002/grl.50193, 2013.
- Le Traon, P.-Y., Antoine, D., Bentamy, A., Bonekamp, H., Breivik, L. A., Chapron, B., Corlett, G., Dibarboure, G., DiGiacomo, P., Donlon, C., Faugère, Y., Font, J., Girard-Ardhuin, F., Gohin, F., Johannessen, J. A., Kamachi, M., Lagerloef, G., Lambin, J., Larnicol, G., Le Borgne, P., Leuliette, E., Lindstrom, E., Martin, M. J., Maturi, E., Miller, L., Mingsen, L., Morrow, R., Reul, N., Rio, M. H., Roquet, H., Santoleri, R., and Wilkin, J.: Use of satellite observations for operational oceanography: recent achievements and future prospects, *Journal of Operational Oceanography*, 8, s12–s27, doi:10.1080/1755876X.2015.1022050, 2015.
- Lien, V. S., Hjøllø, S. S., Skogen, M. D., Svendsen, E., Wehde, H., Bertino L., Counillon, F., Chevallier, M., and Garric, G.: An assessment of the added value from data assimilation on modelled Nordic Seas hydrography and ocean transports, *Ocean Model.*, 99, 43–59, doi:10.1016/j.ocemod.2015.12.010, 2016.
- Lindsay, R. W.: Unified sea ice thickness climate data record collection spanning 1947–2012, Boulder, Colorado USA: National Snow and Ice Data Center, doi:10.7265/N5D50JXV, 2013.
- Lisæter, K., Rosanova, J., and Evensen, G.: Assimilation of ice concentration in a coupled ice-ocean model, using the Ensemble Kalman filter, *Ocean Dynam.*, 53, 368–388, doi:10.1007/s10236-003-0049-4, 2003.
- Locarnini, R., Antonov, J., and Garcia, H.: World Ocean Atlas 2005, Volume 1: Temperature, vol. 61, US Dept. of Commerce, National Oceanic and Atmospheric Administration, 2006.
- Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., and Seidov, D.: World Ocean Atlas 2013, Volume 1, Temperature, edited by: Levitus, S. and Mishonov, A., NOAA Atlas NESDIS, 40 pp., 2013.
- Massonnet, F., Goosse, H., Fichefet, T., and Counillon, F.: Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter, *J. Geophys. Res.*, 119, 4168–4184, doi:10.1002/2013JC009705, 2014.
- Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F., and Vancoppenolle, M.: Better constraints on the sea-ice state using global sea-ice data assimilation, *Geosci. Model Dev.*, 5, 1501–1515, doi:10.5194/gmd-5-1501-2012, 2012.
- Morison, J. H., Long, C. E., and Levine, M. D.: Internal wave dissipation under sea ice, *J. Geophys. Res.*, 90, 11959–11966, doi:10.1029/JC090iC06p11959, 1985.
- Nguyen, A., Menemenlis, D., and Kwok, R.: Improved modeling of the Arctic halocline with a subgrid-scale brine rejection parameterization, *J. Geophys. Res.*, 114, C11014, doi:10.1029/2008JC005121, 2009.
- Nguyen, A., Menemenlis, D., and Kwok, R.: Arctic ice-ocean simulation with optimized model parameters: Approach and assessment, *J. Geophys. Res.*, 116, C04025, doi:10.1029/2010JC006573, 2011.
- Oke, P. R., Brassington, G. B., Griffin, D. A., and Schiller, A.: The Bluelink ocean data assimilation system (BODAS), *Ocean Model.*, 21, 46–70, doi:10.1016/j.ocemod.2007.11.002, 2008.
- Rampal, P., Weiss, J., and Marsan, D.: Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979–2007, *J. Geophys. Res.*, 114, C05013, doi:10.1029/2008JC005066, 2009.
- Rampal, P., Weiss, J., Dubois, C., and Campin, J. M.: IPCC climate models do not capture Arctic sea ice drift acceleration: Consequences in terms of projected sea ice thinning and decline, *J. Geophys. Res.*, 116, C00D07, doi:10.1029/2011JC007110, 2011.
- Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: a new Lagrangian sea ice model, *The Cryosphere*, 10, 1055–1073, doi:10.5194/tc-10-1055-2016, 2016.
- Reul, N., Tenerelli, J., Boutin, J., Chapron, B., Paul, F., Brion, E., Gaillard, F., and Archer, O.: Overview of the first SMOS sea surface salinity products. Part I: Quality assessment for the second half of 2010, *IEEE T. Geosci. Remote*, 50, 1636–1647, doi:10.1109/TGRS.2012.2188408, 2012.
- Reynolds, R. and Smith, T.: Improved global sea surface temperature analyses using optimum interpolation, *J. Climate*, 7, 929–948, doi:10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2, 1994.
- Rodwell, M. J., Lang, S. T. K., Ingleby, N. B., Bormann, N., Hólm, E., Rabier, F., Richardson, D. S., and Yamaguchi, M.: Reliability in ensemble data assimilation, *Q. J. Roy. Meteor. Soc.*, 142, 443–454, doi:10.1002/qj.2663, 2016.
- Rothrock, D. A. and Wensnahan, M.: The accuracy of sea-ice drafts measured from U. S. Navy submarines, *J. Atmos. Ocean. Tech.*, 24, 1936–1949, doi:10.1175/JTECH2097.1, 2007.
- Rothrock, D. A., Percival, D. B., and Wensnahan, M.: The decline in arctic sea-ice thickness: Separating the spatial, annual, and interannual variability in a quarter century of submarine data, *J. Geophys. Res.*, 113, C05003, doi:10.1029/2007JC004252, 2008.
- Sakov, P. and Oke, P. R.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, *Tellus A*, 60, 361–371, doi:10.1111/j.1600-0870.2007.00299.x, 2008.
- Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korabiev, A.: TOPAZ4: an ocean-sea ice data assimilation

- system for the North Atlantic and Arctic, *Ocean Science*, 8, 633–656, doi:10.5194/os-8-633-2012, 2012.
- Samuelson, A., Hansen, C., and Wehde, H.: Tuning and assessment of the HYCOM-NORWECOM V2.1 biogeochemical modeling system for the North Atlantic and Arctic oceans, *Geosci. Model Dev.*, 8, 2187–2202, doi:10.5194/gmd-8-2187-2015, 2015.
- Schweiger, A., Lindsay, R., Zhang, J. L., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic Sea Ice volume, *J. Geophys. Res.*, 116, C00D06, doi:10.1029/2011JC007084, 2011.
- Schweiger, A. J. and Zhang, J.: Accuracy of short-term sea ice drift forecasts using a coupled ice-ocean model, *J. Geophys. Res.*, 120, 7827–7841, doi:10.1002/2015jc011273, 2015.
- Shimada, K., Kamoshida, T., Itoh, M., Nishino, S., Carmack, E., McLaughlin, F., Zimmermann, S., and Proshutinsky, A.: Pacific Ocean inflow: Influence on catastrophic reduction of sea ice cover in the Arctic Ocean, *Geophys. Res. Lett.*, 33, L08605, doi:10.1029/2005GL025624, 2006.
- Simon, E., Samuelson, A., Bertino, L., and Mouysset, S.: Experiences in multiyear combined state-parameter estimation with an ecosystem model of the North Atlantic and Arctic Oceans using the Ensemble Kalman Filter, *J. Marine Syst.*, 152, 1–17, doi:10.1016/j.jmarsys.2015.07.004, 2015.
- Spreen, G., Kwok, R., and Menemenlis, D.: Trends in Arctic sea ice drift and role of wind forcing: 1992–2009, *Geophys. Res. Lett.*, 38, L19501, doi:10.1029/2011GL048970, 2011.
- Stark, J., Donlon, C., Martin, M., and McCulloch, M.: OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system, *OCEAN 2007-Europe*, IEEE, 1–4, doi:10.1109/OCEANSE.2007.4302251, 2007.
- Steele, M., Morley, R., and Ermold, W.: PHC: A global ocean hydrography with a high-quality Arctic Ocean, *J. Climate*, 14, 2079–2087, doi:10.1175/1520-0442(2001)014<2079:PAGOHW>2.0.CO;2, 2001.
- Talagrand, O., Vautard, R., and Strauss, B.: Evaluation of probabilistic prediction system. *Proc. Workshop on Predictability*, Reading, United Kingdom, ECMWF, 1–25, 1999.
- Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, *The Cryosphere*, 8, 997–1018, doi:10.5194/tc-8-997-2014, 2014.
- Tietsche, S., Notz, D., Jungclaus, J. H., and Marotzke, J.: Assimilation of sea-ice concentration in a global climate model – physical and statistical aspects, *Ocean Sci.*, 9, 19–36, doi:10.5194/os-9-19-2013, 2013.
- Wadhams, P. and Horne, R. J.: An analysis of ice profiles obtained by submarine in the Beaufort Sea, *J. Glaciol.*, 25, 401–424, 1980.
- Wensnahan, M. and Rothrock, D. A.: Sea-ice draft from submarine-based sonar: Establishing a consistent record from analog and digitally recorded data, *Geophys. Res. Lett.*, 32, L11502, doi:10.1029/2005GL022507, 2005.
- Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., Chateau-Thierry, P., Laxon, S. W., Mallow, U., Mavrocordatos, C., Phalippou, L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, *Adv. Space Res.*, 37, 841–871, doi:10.1016/j.asr.2005.07.027, 2006.
- Woodgate, R., Aagaard, K., and Weingartner, T.: Monthly temperature, salinity, and transport variability of the Bering Strait through flow, *Geophys. Res. Lett.*, 32, L04601, doi:10.1029/2004GL021880, 2005.
- Xie, J., Counillon, F., Zhu, J., and Bertino, L.: An eddy resolving tidal-driven model of the South China Sea assimilating along-track SLA data using the EnOI, *Ocean Sci.*, 7, 609–627, doi:10.5194/os-7-609-2011, 2011.
- Zygmuntowska, M., Rampal, P., Ivanova, N., and Smedsrud, L. H.: Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends, *The Cryosphere*, 8, 705–720, doi:10.5194/tc-8-705-2014, 2014.