Responses to reviewers' comments

We would like to thank once more the reviewers for their comments. The major changes in the revision are listed as following:

- Add Section 3 to assess the ensemble reliability in the TOPAZ4 reanalysis.
- Shortened the model and assimilation description to reduce overlap with Sakov et al. (2012).
- Improving the uncertainty analysis of the reanalysis with respect to in situ profiles in Section 4.
- Add figures 3, 4 and 8 into the revision
- Changes of the text in agreement with the recommendations of the reviewers.

The detailed responses are listed one by one with blue as following:

Referee #1

1. The main purpose of the paper is the assessment of the reanalysis using all available observations. However, to compare the reanalysis to observations, the authors just compute the average and RMS difference between the ensemble mean and observations. This method looks very crude to me, and does not make justice to the advanced method that is used to perform data assimilation. The ensemble data assimilation system provides a probability distribution for the reanalysis, which is described by an ensemble of model states. Why then assessing the reanalysis using the ensemble mean only? Probabilistic tools exist to perform an objective comparison between ensemble simulations and observations (see for instance Toth et al., 2003, or Candille et al., 2007). Why performing an ensemble reanalysis if the probabilistic information is discarded to study the performance of the system? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method?

Reply: We would like to thank the reviewer for this constructive comment and suggestion. Our main purpose is to present and validate the official product of Copernicus CMEMS for the Arctic region, which is provided as a deterministic reanalysis product based on the ensemble mean, for consistency with other CMEMS reanalyses. However, we fully agree that validation of the quality of the ensemble is crucial to prove the ability of our reanalysis to make the best use of the heterogeneous observational network (spatially, temporally and various data sources); for example that we do not overfit one observational data set at the expense of the others. The reliability for an EnKF-based data assimilation system like ours is even rather important, since the efficiency of the system relies on adequate assumptions for model and observation errors. Unfortunately, our storage facility is insufficient to store the full ensemble of the daily averaged fields, and we only have at our disposal the ensemble statistics of the variables assimilated at each assimilation time (every week).

In order to address the reviewer comment, we have extended our validation work with a reliability analysis (e.g. Candille et al. 2007) of the observation network assimilated according to the all assimilated variables (SST, SSH, Ice concentration, T-S, and sea ice drift).

2. In assessing the performance by computing the difference with observations, the paper implicitly (and sometimes explicitly) assumes that the closer to the observations, the better the reanalysis. This amount to completely neglecting observation errors in the assessment of the reanalysis, which is usually not an appropriate approximation. This incorrect assumption is for instance made explicitly in:

- *p. 13, l. 4, where the misfit to observations is called "error" on the reanalysis;*

Reply: Thanks. It is corrected with “of the misfit”.

- *p. 14, l. 21, where the reanalysis is said to be improved if difference to observations is smaller;*

Reply: Thanks. We have noted that the increased accuracy is an improvement because the reliability remained equal in the meantime.

- *p. 16, l. 6-7, where it is said that an RMSD with observations of 5% is good; whereas the accuracy of the observations is said to be about 10%. In my view, this just mean that the reanalysis is excessively close to observations.*

Reply: Thank you, it is corrected. The reliability analysis in Section 3 revealed - in the contrary - an underdispersion. The sentence now concentrates on the qualitative message (errors concentrated near the ice edge).

I think that it would be important to better explain the limitations of this simple approach for assessing the performance of the reanalysis; to explain why more sophisticated comparison metrics were not applied (see my previous comment) and avoid the misleading expressions listed above.

Reply: We agree with the reviewer and the above statements will be revised according to the reliability analysis. We added the reliability analyses of the modified RCRV to ensure that we are not over-fitting observations and that the ensemble does not collapse, and also use the innovation budget (Rodwell et al., 2016) to investigate the uncertainty variability in time.

3. In the introduction, the authors provide several arguments to support the idea that ensemble methods are an appropriate way to apply the dynamical model constraint in the estimation process. However, this is not discussed anymore in the assessment of the performance of the reanalysis. Only quantitative difference to observations are provided and analysed. I think that the quality of the paper would be enhanced if more explicit evidence of what is stated in the introduction was provided in addition to the simple description of the distance between reanalysis and observations.

Reply: This is now extensively discussed in the manuscript, both during the reliability section and in the conclusion. Despites some discontinuities causes by the change of observational data set and change in the data assimilation setting, the statistic remains relatively stable through the course of the reanalysis. The reliability budget analysis exemplifies the challenge of providing a balances reanalysis with relative contributions from various data sources.

Referee #2

The manuscript appears more like a report than a scientific paper tackling a scientific or methodological issue. The model system is described elsewhere and has undergone very little changes with respect to previously published information. The assessment of the quality of the products uses a rather elementary approach.

Reply: The paper by Sakov et al. (2012) was a proof of concept that an EnKF-based assimilation system can be used with a coupled ocean and sea ice for long reanalysis. This study does not propose new methodological development but it verifies that the proof of concept holds when applied for a longer period (23 years are more relevant to the community than 6 years) with a more heterogeneous observation network (spatially, temporally and various data sources). The main purpose of the manuscript is to present and validate the official Copernicus CMEMS product for the Arctic region. The proposed reanalysis is unique (see table below extracted from Chevalier et al. 2016) as 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. We have tried to present this achievement in a concise manner, with a primary focus to inform the end-user about the strength and weaknesses of our data set. As a response to the recommendation of the first reviewer (and your following comment), we have extended the current validation with the analysis of the ensemble reliability, and asses whether our system manage to provide a dynamical reconstruction that falls within the uncertainty of the different observational data sets that are assimilated. We believe it has increased the scientific value of our manuscript.

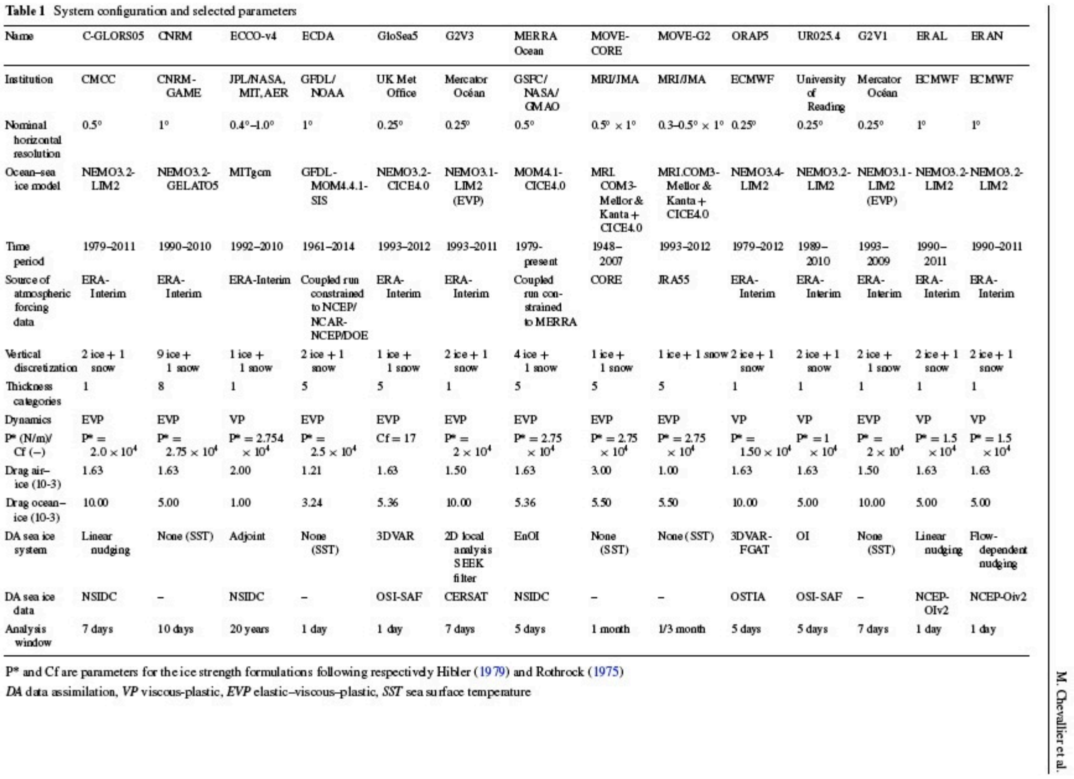


Table 1 System configuration and selected parameters

Name	C-GLORS05	CNRM	ECCO-v4	ECDA	GloSea5	G2V3	MERRA Ocean	MOVE-CORE	MOVE-G2	ORAP5	UR025.4	G2V1	ERAL	ERAN
Institution	CMCC	CNRM-GAME	JPL/NASA, MIT, AER	GFDL/NOAA	UK Met Office	Mercator Océan	GSFC/NASA/GMAO	MRI/JMA	MRI/JMA	ECMWF	University of Reading	Mercator Océan	ECMWF	ECMWF
Nominal horizontal resolution	0.5°	1°	0.4°–1.0°	1°	0.25°	0.25°	0.5°	0.5° × 1°	0.3–0.5° × 1°	0.25°	0.25°	0.25°	1°	1°
Ocean–sea ice model	NEMO3.2-LIM2	NEMO3.2-GELATO5	MITgcm	GFDL-MOM4.4.1-SIS	NEMO3.2-CICE4.0	NEMO3.1-LIM2 (EVP)	MOM4.1-CICE4.0	MRI.COM3-Mellor & Kanta + CICE4.0	MRI.COM3-Mellor & Kanta + CICE4.0	NEMO3.4-LIM2	NEMO3.2-LIM2	NEMO3.1-LIM2 (EVP)	NEMO3.2-LIM2	NEMO3.2-LIM2
Time period	1979–2011	1990–2010	1992–2010	1961–2014	1993–2012	1993–2011	1979-present	1948–2007	1993–2012	1979–2012	1989–2010	1993–2009	1990–2011	1990–2011
Source of atmospheric forcing data	ERA-Interim	ERA-Interim	ERA-Interim	Coupled run constrained to NCEP/NCAR-NCEP/DOE	ERA-Interim	ERA-Interim	Coupled run constrained to MERRA	CORE	JRA55	ERA-Interim	ERA-Interim	ERA-Interim	ERA-Interim	ERA-Interim
Vertical discretization	2 ice + 1 snow	9 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	4 ice + 1 snow	1 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow
Thickness categories	1	8	1	5	5	1	5	5	5	1	1	1	1	1
Dynamics	EVP	EVP	VP	EVP	EVP	EVP	EVP	EVP	EVP	VP	VP	EVP	VP	VP
P* (N/m)/ Cf (–)	$P^* = 2.0 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.754 \times 10^4$	$P^* = 2.5 \times 10^4$	Cf = 17	$P^* = 2 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 1.50 \times 10^4$	$P^* = 1 \times 10^4$	$P^* = 2 \times 10^4$	$P^* = 1.5 \times 10^4$	$P^* = 1.5 \times 10^4$
Drag air–ice (10-3)	1.63	1.63	2.00	1.21	1.63	1.50	1.63	3.00	1.00	1.63	1.63	1.50	1.63	1.63
Drag ocean–ice (10-3)	10.00	5.00	1.00	3.24	5.36	10.00	5.36	5.50	5.50	10.00	5.00	10.00	5.00	5.00
DA sea ice system	Linear nudging	None (SST)	Adjoint	None (SST)	3DVAR	2D local analysis SEEK filter	EnOI	None (SST)	None (SST)	3DVAR-FGAT	OI	None (SST)	Linear nudging	Flow-dependent nudging
DA sea ice data	NSIDC	–	NSIDC	–	OSI-SAF	CERSAT	NSIDC	–	–	OSTIA	OSI-SAF	–	NCEP-OIv2	NCEP-Oiv2
Analysis window	7 days	10 days	20 years	1 day	1 day	7 days	5 days	1 month	1/3 month	5 days	5 days	7 days	1 day	1 day

P* and Cf are parameters for the ice strength formulations following respectively Hibler (1979) and Rothrock (1975)

DA data assimilation, *VP* viscous-plastic, *EVP* elastic–viscous–plastic, *SST* sea surface temperature

M. Chevallier et al.

The results discussed in the manuscript can be useful as a support of further studies using the reanalysed fields but, as it stands, the manuscript is merely descriptive. Also, little information is given about the ensemble and this information is not used to assess the quality of the reanalysis: only the ensemble mean are used for this purpose.

Reply: We agree and as we answered to the other reviewer, the main objective is to present and validate the official product of Copernicus CMEMS for the Arctic region, which is provided as a deterministic reanalysis product based on the ensemble mean, for consistency with other CMEMS reanalyses. Unfortunately, our storage facility is insufficient to store the full ensemble of the daily averaged fields, and we only have at our disposal the ensemble statistics of the variables assimilated at each assimilation time (every week). We have extended our validation work with a reliability analysis (e.g. Candille et al. 2007) of the observation network assimilated according to the all assimilated variables (SST, SSH, Ice concentration, T-S, and sea ice drift).

The quality of the reanalysis obtained using TOPAZ4 could also be compared with the quality of similar other products.

Reply: We think that such comparison is beyond the scope of our paper and, for the sake of diplomatic correctness, is better undertaken in a separate collaborative initiative (The ongoing Ocean Synthesis COST action, a follow-up of the ORA-IP Arctic paper by Chevallier et al.).

A primary comparison of the ocean part of our analysis has been compared with other existing systems (Lien et al. 2016, cited in the manuscript).

Responses to reviewers' comments

We would like to thank once more the reviewers for their comments. The major changes in the revision are listed as following:

- Add Section 3 to assess the ensemble reliability in the TOPAZ4 reanalysis.
- Shortened the model and assimilation description to reduce overlap with Sakov et al. (2012).
- Improving the uncertainty analysis of the reanalysis with respect to in situ profiles in Section 4.
- Add figures 3, 4 and 8 into the revision
- Changes of the text in agreement with the recommendations of the reviewers.

The detailed responses are listed one by one with blue as following:

Referee #1

1. *The main purpose of the paper is the assessment of the reanalysis using all available observations. However, to compare the reanalysis to observations, the authors just compute the average and RMS difference between the ensemble mean and observations. This method looks very crude to me, and does not make justice to the advanced method that is used to perform data assimilation. The ensemble data assimilation system provides a probability distribution for the reanalysis, which is described by an ensemble of model states. Why then assessing the reanalysis using the ensemble mean only? Probabilistic tools exist to perform an objective comparison between ensemble simulations and observations (see for instance Toth et al., 2003, or Candille et al., 2007). Why performing an ensemble reanalysis if the probabilistic information is discarded to study the performance of the system? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method? Would it be possible to include some kind of probabilistic assessment, or at least explain better why using such a crude assessment method?*

Reply: We would like to thank the reviewer for this constructive comment and suggestion. Our main purpose is to present and validate the official product of Copernicus CMEMS for the Arctic region, which is provided as a deterministic reanalysis product based on the ensemble mean, for consistency with other CMEMS reanalyses. However, we fully agree that validation of the quality of the ensemble is crucial to prove the ability of our reanalysis to make the best use of the heterogeneous observational network (spatially, temporally and various data sources); for example that we do not overfit one observational data set at the expense of the others. The reliability for an EnKF-based data assimilation system like ours is even rather important, since the efficiency of the system relies on adequate assumptions for model and observation errors. Unfortunately, our storage facility is insufficient to store the

JipingMac xie 22/12/2016 18:10
Style Definition: Normal (Web)

JipingMac xie 22/12/2016 18:10
Deleted: ... [1]

JipingMac xie 22/12/2016 18:10
Formatted: Font:Bold

JipingMac xie 22/12/2016 18:10
Formatted: Pattern: Clear

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Formatted: Indent: Left: 0 cm, First line: 0 cm, Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0,63 cm + Indent at: 1,27 cm

JipingMac xie 22/12/2016 18:10
Deleted: meke

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted:
Reply: Yes, thank you for your comment. The comment. The modified RCRV (Talagrand et al., 1999; Candille et al., 2007) and the method of Desroziers et al. (2005) will be used to assess the concerned statistic features.

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted:
Reply: Yes, thank you for your comment. The comment. The modified RCRV (Talagrand et al., 1999; Candille et al., 2007) and the method of Desroziers et al. (2005) will be used to assess the concerned statistic features.

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: Reply: Yes, thank you for your comment. The modified RCRV (Talagrand et al., 1999; Candille et al., 2007) and the method of Desroziers et al. (2005) will be used to assess the concerned statistic features.

full ensemble of the daily averaged fields, and we only have at our disposal the ensemble statistics of the variables assimilated at each assimilation time (every week).

In order to address the reviewer comment, we have extended our validation work with a reliability analysis (e.g. Candille et al. 2007) of the observation network assimilated according to the all assimilated variables (SST, SSH, Ice concentration, T-S, and sea ice drift).

2. In assessing the performance by computing the difference with observations, the paper implicitly (and sometimes explicitly) assumes that the closer to the observations, the better the reanalysis. This amount to completely neglecting observation errors in the assessment of the reanalysis, which is usually not an appropriate approximation. This incorrect assumption is for instance made explicitly in:

- *p. 13, l. 4, where the misfit to observations is called "error" on the reanalysis;*

Reply: Thanks. It is corrected with "of the misfit".

- *p. 14, l. 21, where the reanalysis is said to be improved if difference to observations is smaller;*

Reply: Thanks. We have noted that the increased accuracy is an improvement because the reliability remained equal in the meantime.

- *p. 16, l. 6-7, where it is said that an RMSD with observations of 5% is good; whereas the accuracy of the observations is said to be about 10%. In my view, this just mean that the reanalysis is excessively close to observations.*

Reply: Thank you, it is corrected. The reliability analysis in Section 3 revealed - in the contrary - an underdispersion. The sentence now concentrates on the qualitative message (errors concentrated near the ice edge).

I think that it would be important to better explain the limitations of this simple approach for assessing the performance of the reanalysis; to explain why more sophisticated comparison metrics were not applied (see my previous comment) and avoid the misleading expressions listed above.

Reply: We agree with the reviewer and the above statements will be revised according to the reliability analysis. We added the reliability analyses of the modified RCRV to ensure that we are not over-fitting observations and that the ensemble does not collapse, and also use the innovation budget (Rodwell et al., 2016) to investigate the uncertainty variability in time.

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Formatted: Indent: Left: 0 cm, First line: 0 cm

JipingMac xie 22/12/2016 18:10
Deleted: completetly

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**:

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**:

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Formatted: Indent: Hanging: 1,79 cm, Outline numbered + Level: 2 + Numbering Style: Bullet + Aligned at: 1,9 cm + Tab after: 2,54 cm + Indent at: 2,54 cm

JipingMac xie 22/12/2016 18:10
Deleted:
Reply:

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**:

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**: Thanks for this comment. The additional illustration about the probabilistic information will be discussed.

JipingMac xie 22/12/2016 18:10
Formatted: Line spacing: 1.5 lines

3. In the introduction, the authors provide several arguments to support the idea that ensemble methods are an appropriate way to apply the dynamical model constraint in the estimation process. However, this is not discussed anymore in the assessment of the performance of the reanalysis. Only quantitative difference to observations are provided and analysed. I think that the quality of the paper would be enhanced if more explicit evidence of what is stated in the introduction was provided in addition to the simple description of the distance between reanalysis and observations.

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

Reply: This is now extensively discussed in the manuscript, both during the reliability section and in the conclusion. Despites some discontinuities causes by the change of observational data set and change in the data assimilation setting, the statistic remains relatively stable through the course of the reanalysis. The reliability budget analysis exemplifies the challenge of providing a balances reanalysis with relative contributions from various data sources.

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**:

Referee #2

The manuscript appears more like a report than a scientific paper tackling a scientific or methodological issue. The model system is described elsewhere and has undergone very little changes with respect to previously published information. The assessment of the quality of the products uses a rather elementary approach.

JipingMac xie 22/12/2016 18:10
Deleted: The manuscript provides a detailed description of the results of the 23-year reanalysis of the Arctic computed with the TOPAZ4 model system.

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

Reply: The paper by Sakov et al. (2012) was a proof of concept that an EnKF-based assimilation system can be used with a coupled ocean and sea ice for long reanalysis. This study does not propose new methodological development but it verifies that the proof of concept holds when applied for a longer period (23 years are more relevant to the community than 6 years) with a more heterogeneous observation network (spatially, temporally and various data sources). The main purpose of the manuscript is to present and validate the official Copernicus CMEMS product for the Arctic region. The proposed reanalysis is unique (see table below extracted from Chevalier et al. 2016) as 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. We have tried to present this achievement in a concise manner, with a primary focus to inform the end-user about the strength and weaknesses of our data set. As a response to the recommendation of the first reviewer (and your following comment), we have extended the current

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**: Thanks for this comment. However, we disagree this point and the reasons are list as follow: 1) The TOPAZ4 reanalysis is an important new reanalysis product describing the Arctic Ocean (delivered by the Arctic component of the MyOcean system). Although there are some documents (like Sakov et al., 2012) showing the part of validation results, the complete and detailed assessment is still unknown for the public. 2) The reanalysis products in the Arctic with long time period are very few, and their qualities about the dataset self have not been revealed clearly for both ocean and sea ice states. The assessment of the TOPAZ4 reanalysis can partly answer how about the present reanalysis quality in the Arctic. 3) Based on the present assessment results in this paper, we can find some weaknesses which are useful for the concerned system developer. 4) The model system has been simply described and is basic background for understanding the reanalysis system, although it had been described elsewhere.

validation with the analysis of the ensemble reliability, and asses whether our system manage to provide a dynamical reconstruction that falls within the uncertainty of the different observational data sets that are assimilated. We believe it has increased the scientific value of our manuscript.

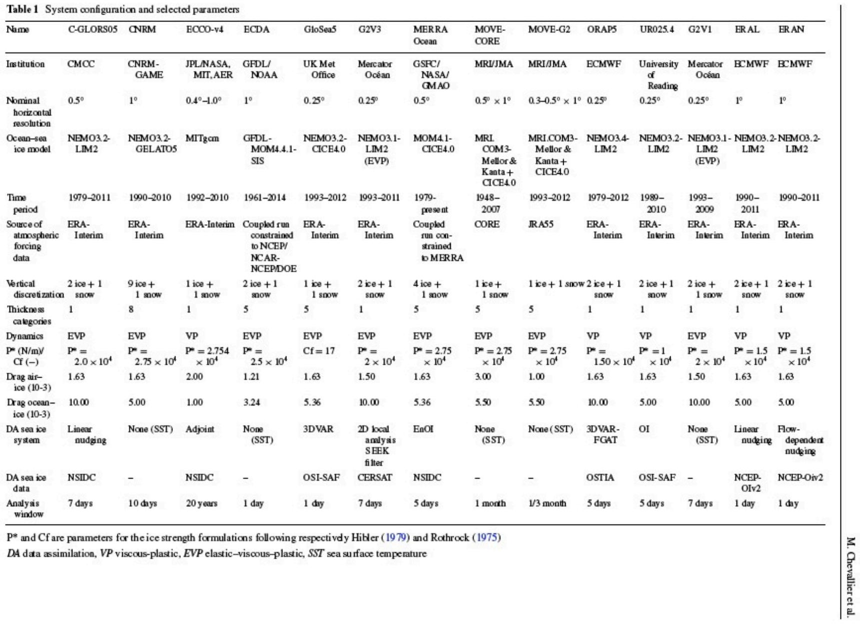


Table 1 System configuration and selected parameters

Name	C-GLORS05	CNRM	ECCO-v4	ECDA	GloSea5	G2V3	MERRA Ocean	MOVE-CORE	MOVE-G2	ORAP5	UR025.4	G2V1	ERAL	ERAN
Institution	CMCC	CNRM-GAME	JPL/NASA, MIT, AER	GFDL/NOAA	UK Met Office	Mercator Océan	GSFC/NASA/GMAO	MRI/JMA	MRI/JMA	ECMWF	University of Reading	Mercator Océan	ECMWF	ECMWF
Nominal horizontal resolution	0.5°	1°	0.4°–1.0°	1°	0.25°	0.25°	0.5°	0.5° × 1°	0.3–0.5° × 1°	0.25°	0.25°	0.25°	1°	1°
Ocean–sea ice model	NEMO3.2-LIM2	NEMO3.2-GELATO5	MITgcm	GFDL-MOM4.4.1-SIS	NEMO3.2-CICE4.0	NEMO3.1-LIM2 (EVP)	MOM4.1-CICE4.0	MRI.COM3-Mellor & Kanta + CICE4.0	MRI.COM3-Mellor & Kanta + CICE4.0	NEMO3.4-LIM2	NEMO3.2-LIM2	NEMO3.1-LIM2 (EVP)	NEMO3.2-LIM2	NEMO3.2-LIM2
Time period	1979–2011	1990–2010	1992–2010	1961–2014	1993–2012	1993–2011	1979-present	1948–2007	1993–2012	1979–2012	1989–2010	1993–2009	1990–2011	1990–2011
Source of atmospheric forcing data	ERA-Interim	ERA-Interim	ERA-Interim	Coupled run constrained to NCEP/NCAR-NCEP/DOE	ERA-Interim	ERA-Interim	Coupled run constrained to MERRA	CORE	JRA55	ERA-Interim	ERA-Interim	ERA-Interim	ERA-Interim	ERA-Interim
Vertical discretization	2 ice + 1 snow	9 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	4 ice + 1 snow	1 ice + 1 snow	1 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow	2 ice + 1 snow
Thickness categories	1	8	1	5	5	1	5	5	5	1	1	1	1	1
Dynamics	EVP	EVP	VP	EVP	EVP	EVP	EVP	EVP	EVP	VP	VP	EVP	VP	VP
P* (N/m)/ Cf (–)	$P^* = 2.0 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.754 \times 10^4$	$P^* = 2.5 \times 10^4$	$Cf = 17$	$P^* = 2 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 2.75 \times 10^4$	$P^* = 1.50 \times 10^4$	$P^* = 1 \times 10^4$	$P^* = 2 \times 10^4$	$P^* = 1.5 \times 10^4$	$P^* = 1.5 \times 10^4$
Drag air–ice (10-3)	1.63	1.63	2.00	1.21	1.63	1.50	1.63	3.00	1.00	1.63	1.63	1.50	1.63	1.63
Drag ocean–ice (10-3)	10.00	5.00	1.00	3.24	5.36	10.00	5.36	5.50	5.50	10.00	5.00	10.00	5.00	5.00
DA sea ice system	Linear nudging	None (SST)	Adjoint	None (SST)	3DVAR	2D local analysis SEEK filter	EnOI	None (SST)	None (SST)	3DVAR-FGAT	OI	None (SST)	Linear nudging	Flow-dependent nudging
DA sea ice data	NSIDC	–	NSIDC	–	OSI-SAF	CERSAT	NSIDC	–	–	OSTIA	OSI-SAF	–	NCEP-OIv2	NCEP-Oiv2
Analysis window	7 days	10 days	20 years	1 day	1 day	7 days	5 days	1 month	1/3 month	5 days	5 days	7 days	1 day	1 day

P* and Cf are parameters for the ice strength formulations following respectively Hibler (1979) and Rothrock (1975)

DA data assimilation, *VP* viscous-plastic, *EVP* elastic–viscous–plastic, *SST* sea surface temperature

M. Chevallier et al.

The results discussed in the manuscript can be useful as a support of further studies using the reanalysed fields but, as it stands, the manuscript is merely descriptive. Also, little information is given about the ensemble and this information is not used to assess the quality of the reanalysis: only the ensemble mean are used for this purpose.

Reply: We agree and as we answered to the other reviewer, the main objective is to present and validate the official product of Copernicus CMEMS for the Arctic region, which is provided as a deterministic reanalysis product based on the ensemble mean, for consistency with other CMEMS reanalyses. Unfortunately, our storage facility is insufficient to store the full ensemble of the daily averaged fields, and we only have at our disposal the ensemble statistics of the variables assimilated at each assimilation time (every week). We have extended our validation work with a reliability analysis (e.g. Candille et al. 2007) of the observation network assimilated according to the all assimilated variables (SST, SSH, Ice concentration, T-S, and sea ice drift).

JipingMac xie 22/12/2016 18:10
Formatted: Font:11 pt, Italic

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**:

The quality of the reanalysis obtained using TOPAZ4 could also be compared with the quality of similar other products.

Reply: We think that such comparison is beyond the scope of our paper and, for the sake of diplomatic correctness, is better undertaken in a separate collaborative initiative (The ongoing Ocean Synthesis COST action, a follow-up of the ORA-IP Arctic paper by Chevallier et al.).

A primary comparison of the ocean part of our analysis has been compared with other existing systems (Lien et al. 2016, cited in the manuscript).

JipingMac xie 22/12/2016 18:10
Formatted: Font:11 pt, Italic

JipingMac xie 22/12/2016 18:10
Formatted: Line spacing: 1.5 lines

JipingMac xie 22/12/2016 18:10
Formatted: Font:Italic

JipingMac xie 22/12/2016 18:10
Deleted: **Reply**: Yes, the comparison with other similar products also is important.

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

Jiping Xie[1], **Laurent Bertino**[1], **Francois Counillon**[1], **Knut A. Lisæter**[1], **and Pavel Sakov**[2]

[1]Nansen Environmental and Remote Sensing Center, Bergen N5006, Norway
[2]Bureau of Meteorology, Melbourne VIC3001, Australia

E-mail: jiping.xie@nersc.no

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-years 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

JipingMac xie 22/12/2016 18:03
Formatted: Font:12 pt

JipingMac xie 22/12/2016 18:03
Deleted: the Arctic

JipingMac xie 22/12/2016 18:03
Deleted: .

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 (RCRV) 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 datasets 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.

Keywords: Arctic Ocean, EnKF, Reanalysis, Reliability analysis, Quality assessment.

JipingMac xie 22/12/2016 18:03
Deleted: HYCOM, EnKF

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); 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

JipingMac xie 22/12/2016 18:03
Deleted: and in particular

JipingMac xie 22/12/2016 18:03
Deleted: in ice-covered regions.

hampered by the sparseness of the concerned observations, which should not be improved dramatically in a 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 time scales.

The latter activities thus necessitate a state-of-the-art reanalysis system able to honour accurately 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 optimization 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), or 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

JipingMac xie 22/12/2016 18:03
Deleted: of sea ice and ocean

JipingMac xie 22/12/2016 18:03
Deleted: is

JipingMac xie 22/12/2016 18:03
Deleted: expected to improve

JipingMac xie 22/12/2016 18:03
Deleted: the mislocation

JipingMac xie 22/12/2016 18:03
Deleted: the

JipingMac xie 22/12/2016 18:03
Deleted: , of

JipingMac xie 22/12/2016 18:03
Deleted: in an unconstrained simulation.

JipingMac xie 22/12/2016 18:03
Deleted: then recurs

JipingMac xie 22/12/2016 18:03
Deleted: studying

JipingMac xie 22/12/2016 18:03
Deleted: databases

JipingMac xie 22/12/2016 18:03
Deleted: ,

JipingMac xie 22/12/2016 18:03
Deleted: . 2016

account for flow-dependent coupled ice-ocean data assimilation updates had 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 Ocean (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 (Samuelsen 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 the European Center for Medium-Range Weather Forecast (ECMWF) ERA-Interim reanalysis (Dee et al., 2011), TOPAZ4 assimilates most available measurements including along-track sea level anomalies (SLA) from satellite altimeters, sea surface temperatures (SST), sea ice concentrations (SIC) and sea ice drift (SID) from satellites as well as in situ temperature and salinity profiles. The proposed reanalysis is four 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

JipingMac xie 22/12/2016 18:03
Deleted: was mitigating

JipingMac xie 22/12/2016 18:03
Deleted: ,

JipingMac xie 22/12/2016 18:03
Deleted: MyOcean

JipingMac xie 22/12/2016 18:03
Deleted:),

JipingMac xie 22/12/2016 18:03
Deleted: demonstrated

JipingMac xie 22/12/2016 18:03
Deleted: found that

JipingMac xie 22/12/2016 18:03
Deleted: can avoid

JipingMac xie 22/12/2016 18:03
Deleted: and that the ensemble statistics provide

JipingMac xie 22/12/2016 18:03
Deleted: that data assimilation

JipingMac xie 22/12/2016 18:03
Deleted: model

JipingMac xie 22/12/2016 18:03
Deleted: TOPAZ4 system is forced by

JipingMac xie 22/12/2016 18:03
Deleted:

JipingMac xie 22/12/2016 18:03
Deleted:) and

JipingMac xie 22/12/2016 18:03
Deleted: altimetry data

JipingMac xie 22/12/2016 18:03
Deleted: ,

JipingMac xie 22/12/2016 18:03
Deleted: data

JipingMac xie 22/12/2016 18:03
Deleted: observations

JipingMac xie 22/12/2016 18:03
Deleted: observations

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 about 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 subareas, its seasonal cycle and its trend in the course of the reanalysis.

The outline of this paper is as follows: In section 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 section 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 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 the version 3.0 of the Polar Science Center Hydrographic Climatology (PHC, Steele et al., 2001). HYCOM is

JipingMac xie 22/12/2016 18:03
Deleted: of the reanalysis for ocean and sea ice variables

JipingMac xie 22/12/2016 18:03
Deleted: performance

JipingMac xie 22/12/2016 18:03
Deleted: space

JipingMac xie 22/12/2016 18:03
Deleted: the implementing set-up.

JipingMac xie 22/12/2016 18:03
Deleted: presents

JipingMac xie 22/12/2016 18:03
Deleted: result of

JipingMac xie 22/12/2016 18:03
Deleted: of the system with

JipingMac xie 22/12/2016 18:03
Deleted: with

JipingMac xie 22/12/2016 18:03
Deleted: In

JipingMac xie 22/12/2016 18:03
Deleted: 4, we summarize the results,

JipingMac xie 22/12/2016 18:03
Deleted: discuss the results and

JipingMac xie 22/12/2016 18:03
Deleted: ways to improve

JipingMac xie 22/12/2016 18:03
Deleted: The model is mainly unchanged compared to that used in Sakov et al. (2012).

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 out flow 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*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 **H** is the observation operator denoting the projection from the model state variables to the measurements. The superscripts “a” and “f” refer to the analyzed and the forecast state 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

JipingMac xie 22/12/2016 18:03
Deleted: The model is spin up from 1973 using the European Center for Medium-Range Weather Forecast (ECMWF) ERA-Interim reanalysis data (Simmons et al., 2007).

JipingMac xie 22/12/2016 18:03
Deleted: of

JipingMac xie 22/12/2016 18:03
Deleted: state and its uncertainty

JipingMac xie 22/12/2016 18:03
Deleted: computed

JipingMac xie 22/12/2016 18:03
Deleted: $\mathrm{K_i}$

JipingMac xie 22/12/2016 18:03
Deleted: $\mathbf{d}$

JipingMac xie 22/12/2016 18:03
Deleted: $\mathbf{d}_i$

JipingMac xie 22/12/2016 18:03
Deleted: observation vector

JipingMac xie 22/12/2016 18:03
Formatted: Font:Not Italic

JipingMac xie 22/12/2016 18:03
Deleted: vector

JipingMac xie 22/12/2016 18:03
Deleted: measurement

JipingMac xie 22/12/2016 18:03
Deleted: respectively

JipingMac xie 22/12/2016 18:03
Deleted: states.

JipingMac xie 22/12/2016 18:03
Deleted:). Compared

observations, named innovations. Differed from 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 (are 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 Section 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 reanalysis, the SST data assimilated is the 1° resolution Reynolds SST from NOAA (Reynolds and Smith, 1994), which is replaced in June 1998 by the high-resolution OSTIA data (Stark et al, 2007) from the UK Metoffice. The SLA data assimilated is the delayed-time product (vxxc), which is validated, unfiltered and not sub-sampled from Collecte

JipingMac xie 22/12/2016 18:03
Deleted: only modification is the removal of the ...% multiplicative inflation ne ... [1]

JipingMac xie 22/12/2016 18:03
Deleted: TOPAZ4 system...eanaly ... [2]

JipingMac xie 22/12/2016 18:03
Deleted: The ...t the beginning of t ... [3]

Localisation Satellites (CLS). The SIC from the Ocean & Sea Ice Satellite Application Facility (OSISAF) are assimilated into. Before the 19th June 2002, this assimilated product is derived from SSM/I at 25 km resolution, and later is derived from AMSR-E 89 GHz brightness temperature at 12.5 km resolution. In the last three years, this product has been upgraded to at 10 km resolution. The temperature and salinity profiles include Argo floats, Ice-Tethered Profiles (ITP) from the Damocles project and a large collection of hydrographic cruise data. At 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 Equation (1) is modified as:

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

where $\bar{\mathbf{x}}_{\mathrm{i}}$ is the ensemble mean of the model state vector at the analysis time *i*, $\mathbf{y}_{\mathrm{i}}$ is the vector of observations, and $\mathbf{c}_{\mathrm{i}}^{\mathrm{f}}$ 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 **H** operator).

The initial biases for each ensemble member are random values, homogeneous in space and uniformly distributed. The initial SST biases are sampled in the interval [-4, 4] °C and within [-0.6, 0.6] m for the MSSH.

JipingMac xie 22/12/2016 18:03
Deleted: Furthermore, the assimilated ICEC are…he SIC from the Ocean ... [4]

JipingMac xie 22/12/2016 18:03
Deleted: As described in Sakov et al. (2012), two… Two bias fields (for ... [5]

JipingMac xie 22/12/2016 18:03
Deleted: $\begin{pmatrix}\mathbf{X}_{\mathrm{i}}^{\mathrm{a}}\\ \mathbf{B}_{\mathrm{i}}^{\mathrm{a}}\end{pmatrix}$

JipingMac xie 22/12/2016 18:03
Deleted: $\begin{pmatrix}\mathbf{X}_{\mathrm{i}}^{\mathrm{f}}\\ \mathbf{B}_{\mathrm{i}}^{\mathrm{f}}\end{pmatrix}\begin{pmatrix}\overline{\mathbf{X}} \ {}_{\mathrm{i}}^{\mathrm{f}}\\ \mathbf{c} \ {}_{\mathrm{i}}^{\mathrm{f}}\end{pmatrix} +$ ... [6]

JipingMac xie 22/12/2016 18:03
Deleted: $\mathbf{B}_{\mathbf{i}}^{\mathbf{a}}$ represents the bias estimates

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Deleted: ,…and uniformly distribut ... [7]

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 observations 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 of order one, 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 end of the reanalysis for the SSH and the SST. The bias patterns compare well with those obtained in Sakov et al. (2012)[1]. 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 $\mathbf{c}_i^a$ are used to offset the reanalysis state.

JipingMac xie 22/12/2016 18:03
Deleted: and do… but are not evo ... [8]

JipingMac xie 22/12/2016 18:03
Deleted: from that date.

JipingMac xie 22/12/2016 18:03
Deleted: In order to avoid inconsistencies between assimilation of SST and temperature profile, the SST bias is injected downwards into the model mixed layer and decays exponentially. ... [9]

JipingMac xie 22/12/2016 18:03
Deleted: while we show the SSH bias

---

[1] Sakov et al. (2012) present the mean SSH bias of 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 of 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 Section 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 $\bar{\mathbf{x}}^{\mathrm{f}}$, the scalar variable q can be defined as the innovation normalized by the observation and model uncertainties:

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

where σ_o is the observation error and σ_{en} is the standard deviation of the corresponding forecast ensemble, including the 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:

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

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

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

measures the ensemble dispersion with respect to the normalized misfits.

The first two moments of the RCRV, *b* and *d*, provide simple diagnostics 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 of 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 the *b* and *d* in the 23 years are shown in Fig. 3 and Fig. 4.

The dispersion and seasonal bias of SLA increase after the launch of ENVISAT in 2002, when previously unobserved areas at high latitude get to be included in the calculation of the statistics. We can notice that the bias stabilizes 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, improves initially 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 of land mask. The warm bias is dominant in summer. During the last three years of the reanalysis, the summer warm bias b is reduced but the dispersion shrinks dramatically. This coincides with the time when the observation error was increased and the quality control of the observations (based on observation uncertainty) was softened, which results in assimilating more observations in the Gulf Stream and near the ice edge. Although it is somewhat counter-intuitive 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 spatio-temporal interpolation of SST may have degraded the reliability (this will be further discussed in Section 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 (Lisæter et al.

2003). A bias degradation and a dispersion improvement are jointed with clear seasonality during the last three 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 by 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 (consistently 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 (~d=1.36) when the SID data source is switched from 3-days drifts on 35 km resolution to 2-days drifts on 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 Section 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 type is severely less dispersed than the others. For most of the assimilated observation datasets, 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 marched 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 (RMSD) of the misfits are calculated as described in Equations of (6) and (7):

$$\mathrm{Bias} = \frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i=1}}^{\mathbf{N}}(\mathbf{H_i\bar{x}_i^f - y_i - Hc_i^f}) \quad (6)$$

$$\mathrm{RMSD} = \sqrt{\frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i=1}}^{\mathbf{N}}(\mathbf{H_i\bar{x}_i^f - y_i - Hc_i^f})^2} \quad , \quad (7).$$

Where $\bar{x}_i^f$ is the forecasted daily average from the ensemble mean, which is compared to the observations y_i on the same day. N is the number of time sampling over the diagnostic period (like either 365 or 366 for 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

JipingMac xie 22/12/2016 18:03
Formatted: Indent: Hanging: 0,77 cm, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 3 + Alignment: Left + Aligned at: 0,63 cm + Indent at: 1,27 cm

JipingMac xie 22/12/2016 18:03
Deleted: model

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Deleted: model error is

JipingMac xie 22/12/2016 18:03
Deleted: from

JipingMac xie 22/12/2016 18:03
Deleted: misfits of

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: 4

JipingMac xie 22/12/2016 18:03
Deleted: $\sum_{\mathbf{i=1}}^{\mathbf{N}}(\mathbf{H_i\bar{X}_i^f - d_i - HB_i^f})$

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: $\sqrt{\frac{1}{N}\sum_{\mathbf{i=1}}^{\mathbf{N}}(\mathbf{H_i\bar{X}_i^f - d_i - HB_i^f})^2}$

JipingMac xie 22/12/2016 18:03
Deleted: 4

JipingMac xie 22/12/2016 18:03
Deleted: $\bar{\mathbf{X}}_i^f$

JipingMac xie 22/12/2016 18:03
Deleted: $\mathbf{d}_i$

JipingMac xie 22/12/2016 18:03
Deleted: daily averages available

JipingMac xie 22/12/2016 18:03
Deleted: B_i^f

JipingMac xie 22/12/2016 18:03
Deleted: bias estimate (B_i^f

JipingMac xie 22/12/2016 18:03
Deleted: B_{i-1}^a).

quantities - i.e. the variability 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 to 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 up 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 in left of Fig. 6. The number of assimilated observations evolves with the launch or completion 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 in 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 1cm but remains positive. The RMSD gradually reduces with the introduction of new

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: variability…LA accuracy ... [10]

JipingMac xie 22/12/2016 18:03
Deleted: the …eft column…of Fig ... [11]

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 results of improved SST or temperature and salinity profiles. Meanwhile, the temporal standard deviation of the RMSD during the year (shown as the-half-error 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 in right 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 is ranged from 5 to 7 cm as a consequence of the seasonal spatial coverage. The residual bias is positive throughout in one year but reaching 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 in 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 Section 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 °C) north of Denmark Strait. It is known where the recirculation of Atlantic Water inflow in

JipingMac xie 22/12/2016 18:03
Deleted: 2008-…009-2010 does r ... [12]

JipingMac xie 22/12/2016 18:03
Deleted: the …ight column …f Fig ... [13]

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: distribution…ariability of ... [14]

TOPAZ4 is excessive as identified in Lien et al. (2016). This pattern was also visible in the estimated bias shown in Fig. 2, suggesting that the estimated bias account for most of the bias but that it still underestimates the true bias. An additional stripe of the cold residual bias and higher RMSD is clear along Mohn’s 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 °C. On averaged in the whole Arctic region, the SST RMSD is about 0.44 °C during the period 1999-2013.

The evolution of SST accuracy of the TOPAZ4 reanalysis is shown in left 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 °C, and the bias -0.08 °C. In the first years, the SST RMSDs are initially about 1 °C but decrease gradually down to 0.8 °C before 1998. During this period, the model has a cold SST bias around -0.3 °C with 0.1 °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°C as a direct consequence of the bias reduction and the more abundant observations. In 2010, the RMSD reaches the minimum below 0.4°C. At that time, the ensemble spread was getting too small, and the system performance was too constrained by SST as can be seen on the standard deviation of RMSD. It was thus decided to increase artificially the SST observation errors, which resulted in a small increase of the misfit up to 0.5 °C. It is clear from the above that the transition to high-resolution SST in our system has led to a higher SST accuracy.

JipingMac xie 22/12/2016 18:03
Formatted ... [15]

JipingMac xie 22/12/2016 18:03
Deleted: (…hown in Fig. 2., with a ... [16]

JipingMac xie 22/12/2016 18:03
Deleted: on the…n left column … ... [17]

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 smaller than 0.4 °C owing to more observations, and is maximum at 0.6 °C in June and July when the bias is as well maximum. 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 1.1x10^5 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 too weak northward Atlantic Water flow across the Fram Strait and a too weak southward flow of Arctic Water East of Svalbard. There are too saline biases 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 together their misfits are shown in Fig. 9. The analysis is

JipingMac xie 22/12/2016 18:03
Deleted:

JipingMac xie 22/12/2016 18:03
Deleted: variability of the SST

JipingMac xie 22/12/2016 18:03
Deleted: 5

JipingMac xie 22/12/2016 18:03
Deleted: : the

JipingMac xie 22/12/2016 18:03
Deleted: less

JipingMac xie 22/12/2016 18:03
Deleted: when there are

JipingMac xie 22/12/2016 18:03
Deleted: observation

JipingMac xie 22/12/2016 18:03
Deleted: over

JipingMac xie 22/12/2016 18:03
Deleted: unclear

JipingMac xie 22/12/2016 18:03
Deleted: is possibly related

JipingMac xie 22/12/2016 18:03
Deleted: errors in

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: distribution of these profiles is

JipingMac xie 22/12/2016 18:03
Deleted: The averaged assimilated

JipingMac xie 22/12/2016 18:03
Deleted: , the corresponding profiles

JipingMac xie 22/12/2016 18:03
Deleted: ,

JipingMac xie 22/12/2016 18:03
Deleted:)

JipingMac xie 22/12/2016 18:03
Deleted: all shown in Fig.6. The stratification in the Arctic varies regionally, depending on the influence of the Atlantic Water inflow, the Pacific Water infl ... [18]

JipingMac xie 22/12/2016 18:03
Deleted: into

separated in four sub-regions: 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). In the near 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 Section 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 having some skills near the surface but deteriorating at depths where the AW is present but it 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 (Ilicak et al., 2016).

In the Barents Sea, the RMSD for temperature and salinity can be reduced near surface, even compared to that of the climatology. But the AW (temperature > 3°C and salinity > 35 psu, Blindheim and Østerhus, 2003) of

JipingMac xie 22/12/2016 18:03
Deleted: warmer water and

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Deleted: ,

JipingMac xie 22/12/2016 18:03
Deleted: is doing

JipingMac xie 22/12/2016 18:03
Deleted: better

JipingMac xie 22/12/2016 18:03
Deleted: for salinity

JipingMac xie 22/12/2016 18:03
Deleted: RMSD (

JipingMac xie 22/12/2016 18:03
Deleted:)

JipingMac xie 22/12/2016 18:03
Deleted: An area of larger

JipingMac xie 22/12/2016 18:03
Deleted: better

JipingMac xie 22/12/2016 18:03
Deleted: 200m

JipingMac xie 22/12/2016 18:03
Deleted: 800m

JipingMac xie 22/12/2016 18:03
Deleted: . The Atlantic Water (AW)

JipingMac xie 22/12/2016 18:03
Deleted: is improved

the TOPAZ4 is too warm and saline, which suggests there is too much AW inflow or too weak vertical mixing.

Furthermore, we investigate the time evolution of the misfits throughout the reanalysis. Figure 10 shows the time series of the Root Mean Square innovations (RMSI) 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 σ_{tot} from Rodwell et al. (2016), which assume that for a perfect reliable system RMSI is equal to σ_{tot}, with bias included:

$$\sigma_{tot}^2 = \mathrm{BIAS}^2 + \sigma_{en}^2 + \sigma_o^2 \qquad (8)$$

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 Section 3), the enhanced accuracy can be considered a performance improvement, directly caused by the intensive observation efforts. The diagnosed uncertainty σ_{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 σ_{tot} slightly underestimates the RMSI, which may results from the removal of the multiplicative inflation.

JipingMac xie 22/12/2016 18:03
Deleted: In Fig. 7,

JipingMac xie 22/12/2016 18:03
Deleted: performance of TOPAZ4 for temperature and salinity is presented by the

JipingMac xie 22/12/2016 18:03
Deleted: innovation diagnostics

JipingMac xie 22/12/2016 18:03
Deleted: . The system performance is relatively stable until 2006, with a cold and fresh bias and RMSD

JipingMac xie 22/12/2016 18:03
Deleted: about 1.5°C and 0.1 psu. From 2006, the performance is greatly improved, supposedly by the combined effect

JipingMac xie 22/12/2016 18:03
Deleted: change

JipingMac xie 22/12/2016 18:03
Deleted: multiplicative

JipingMac xie 22/12/2016 18:03
Deleted: additive inflation and

JipingMac xie 22/12/2016 18:03
Deleted: large increase

JipingMac xie 22/12/2016 18:03
Deleted: observation number

JipingMac xie 22/12/2016 18:03
Deleted:

JipingMac xie 22/12/2016 18:03
Deleted: Once

JipingMac xie 22/12/2016 18:03
Deleted: period is finished, the number of observation drastically reduces. It

JipingMac xie 22/12/2016 18:03
Deleted: in a clear increase

JipingMac xie 22/12/2016 18:03
Deleted: RMSD and bias for temperature to a level intermediate to prior the

For salinity, the model seems too saline until the start of the IPY. The bias is not reemerging 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 σ_{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 is calculated (300-800 m here against full observation depth in Fig3). The cause of the overestimation stems from a too large 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 are 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 sides 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 region (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

JipingMac xie 22/12/2016 18:03
Deleted: period, but the error

JipingMac xie 22/12/2016 18:03
Deleted: remains to a lower level than before the IPY

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Formatted: Indent: First line: 0,5 cm

JipingMac xie 22/12/2016 18:03
Deleted: Copernicus

JipingMac xie 22/12/2016 18:03
Deleted: daily

JipingMac xie 22/12/2016 18:03
Deleted: of sea-ice concentration is

JipingMac xie 22/12/2016 18:03
Deleted: 8

JipingMac xie 22/12/2016 18:03
Deleted: there is

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 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.

JipingMac xie 22/12/2016 18:03
Deleted: the whole

JipingMac xie 22/12/2016 18:03
Deleted: , which is good in view the accuracy of the data set (~10%). There are regions

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 sub-regions: 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 are 8.03x10^6 instead of 7.96x10^6 km^2 in the observations. The decreasing trend of SIE during the period 1991-2013, is -6.16x10^4 km^2 y^{-1}, which compares well to the trend of the observations (-6.34x10^4 km^2 y^{-1}).

JipingMac xie 22/12/2016 18:03
Deleted: 9

JipingMac xie 22/12/2016 18:03
Deleted: the

In the Greenland Sea, the SIE in TOPAZ4 is underestimated, which clearly relate to the bias in the southern extent of the sea-ice tongue along the coast of Greenland. The bias in TOPAZ4 is in averaged -3.6x10^4 km^2 and the decreasing trend in TOPAZ4 is -3.1x10^3 km^2 y^{-1}, which is larger than observed (-2.3x10^3 km^2 y^{-1}). In the Barents Sea, the variability agrees well, although TOPAZ4 underestimates slightly the SIE. The decreasing trend is comparable.

The seasonality of the SIE in OSISAF and TOPAZ4 are 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 1x10^5 km^2) 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 back 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 12h 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 from 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 (4 grid cells which corresponds to approximately 60x60 km^2) to avoid having too many empty cells. The daily averaged from the measurement is the mean of the 12h drifting speed. For comparison, the model drifting speed is calculated from daily averaged of eastward and

JipingMac xie 22/12/2016 18:03
Deleted: 10

JipingMac xie 22/12/2016 18:03
Deleted: little

JipingMac xie 22/12/2016 18:03
Deleted: in the reanalysis

JipingMac xie 22/12/2016 18:03
Deleted: is

JipingMac xie 22/12/2016 18:03
Deleted: feature

JipingMac xie 22/12/2016 18:03
Deleted: large

JipingMac xie 22/12/2016 18:03
Deleted: but close

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: drift

JipingMac xie 22/12/2016 18:03
Deleted: Probram

JipingMac xie 22/12/2016 18:03
Deleted: ,

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 overestimates globally the ice drift with a bias of 1.7 km d^{-1}. In the Chukchi Sea, TOPAZ4 underestimates the drift by approximately -2 km d^{-1}.

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^{-1} in buoy data, and about 9.4 km d^{-1} 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 2.14x10^{-3} to 1.6x10^{-3}. We can see that the bias is much reduced during the last year. The RMSD is on average 5.1 km d^{-1}, of which 2.5 km d^{-1} can be attributed to the bias. The correlation between the 2 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).

JipingMac xie 22/12/2016 18:03
Deleted: with lagrangian

JipingMac xie 22/12/2016 18:03
Deleted: for

JipingMac xie 22/12/2016 18:03
Deleted: the drift

JipingMac xie 22/12/2016 18:03
Deleted: larger

JipingMac xie 22/12/2016 18:03
Deleted: average

JipingMac xie 22/12/2016 18:03
Deleted: for

JipingMac xie 22/12/2016 18:03
Deleted: the drift

JipingMac xie 22/12/2016 18:03
Deleted: higher

JipingMac xie 22/12/2016 18:03
Deleted: averaged

JipingMac xie 22/12/2016 18:03
Deleted: 11

JipingMac xie 22/12/2016 18:03
Deleted: we have applied

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Deleted: differs

JipingMac xie 22/12/2016 18:03
Deleted: Ice

JipingMac xie 22/12/2016 18:03
Deleted: 12

JipingMac xie 22/12/2016 18:03
Deleted: anomaly

JipingMac xie 22/12/2016 18:03
Deleted: for

JipingMac xie 22/12/2016 18:03
Deleted: 13

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; Zygmuntowska 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 yet 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 2003-2008 (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 m 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 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 in Fig. 18(a). The sea-ice draft data has been diagnosed in TOPAZ4 as proposed by the equation (4) of Alexandrov et al. (2010):

JipingMac xie 22/12/2016 18:03
Deleted: 3

JipingMac xie 22/12/2016 18:03
Deleted: been paid

JipingMac xie 22/12/2016 18:03
Deleted: 14

JipingMac xie 22/12/2016 18:03
Deleted: 15

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

Where D_i is ice draft, H_i is ice thickness, and H_{sn} is the snow thickness. The ρ_i, ρ_w, and ρ_{sn} are the densities for sea ice, water, and snow (respectively 900 kg m^{-3}, 1000 kg m^{-3}, and 300 kg m^{-3}).

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 .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 set. Hence, IceBridge data is concentrated around the Northern coast of Greenland where TOPAZ4 showed largest bias in the comparison with ICESat.

As another diagnostics 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 1.4x10^4 km^3, with a significant seasonal variability between 8x10^3 km^3 and 1.9x10^4 km^3. 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. (2013) 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.

JipingMac xie 22/12/2016 18:03
Deleted: 5

JipingMac xie 22/12/2016 18:03
Deleted: ICESAT

JipingMac xie 22/12/2016 18:03
Deleted: of

JipingMac xie 22/12/2016 18:03
Deleted: 16

JipingMac xie 22/12/2016 18:03
Deleted: 16

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 1.5x10^4 km^3, and in September is less than 5x10^3 km^3. The sea ice volumes from Zygmuntowska et al. (2013) 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^2.s^{-2}, 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 has then isolated the ice from the atmosphere and reduced the amplitude of the yearly cycle from 1.08 m to 0.74 m (see Figure 20), this also delayed the phase of the cycle. In January 2011, an unbiased log-normal law replaces the truncated Gaussian perturbations with an amplitude of 30%. The amplitude and phase of the seasonal cycle return to more correct values. The sensitivity experiments in Finck et al. (2013) verified the above-mentioned issue.

JipingMac xie 22/12/2016 18:03
Deleted: 16

JipingMac xie 22/12/2016 18:03
Deleted: 17

JipingMac xie 22/12/2016 18:03
Formatted: Font:Bold

JipingMac xie 22/12/2016 18:03
Formatted: Left, Indent: Left: 1,5 cm

JipingMac xie 22/12/2016 18:03
Formatted: Indent: Left: 0,75 cm, Hanging: 0,75 cm, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, ... + Start at: 3 + Alignment: Left + Aligned at: 0,63 cm + Indent at: 1,27 cm

5. Summary and discussions

This study is 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 of 100 to 200 m), even compared to that

JipingMac xie 22/12/2016 18:03
Formatted: Indent: Left: 0,12 cm

JipingMac xie 22/12/2016 18:03
Deleted: near 8

JipingMac xie 22/12/2016 18:03
Deleted: 10

JipingMac xie 22/12/2016 18:03
Deleted: in

JipingMac xie 22/12/2016 18:03
Deleted: error

JipingMac xie 22/12/2016 18:03
Deleted: in summer months

JipingMac xie 22/12/2016 18:03
Deleted: 66

JipingMac xie 22/12/2016 18:03
Deleted: the bias

JipingMac xie 22/12/2016 18:03
Deleted: 04

JipingMac xie 22/12/2016 18:03
Deleted: . Some biases along the sea-ice edge can be related to errors in the observations. Other biases and RMSD relate to errors in the circulation within the Nordic Sea.

JipingMac xie 22/12/2016 18:03
Deleted: OSTIA

JipingMac xie 22/12/2016 18:03
Deleted: observation

JipingMac xie 22/12/2016 18:03
Deleted: amount of observations

JipingMac xie 22/12/2016 18:03
Deleted: errors

JipingMac xie 22/12/2016 18:03
Deleted: lower than the WOA13 climatology

JipingMac xie 22/12/2016 18:03
Deleted:).

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 Ilicak et al. 2016) and others relate to a degradation introduced by data assimilation (a flat multiplicative inflation). A large improvement occurs at the times when the inflation method was upgraded and when 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 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 IAPB 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 km d^{-1}, and has a fast bias by about 2.5 km d^{-1}. The monthly time variability compares well, but TOPAZ4 has a too weak seasonal cycle and shifted by two 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 km d^{-1} close to the signal itself, improving

JipingMac xie 22/12/2016 18:03
Deleted: time

JipingMac xie 22/12/2016 18:03
Deleted: replaced

JipingMac xie 22/12/2016 18:03
Deleted: numerous

JipingMac xie 22/12/2016 18:03
Deleted: were available from

JipingMac xie 22/12/2016 18:03
Deleted: errors

the performance of ice drift appears as 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 m) compared to ice thickness from ICESat and IceBridge as well as compare 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 type is severely less dispersed than the others. The results from the 23-years reanalysis show overall a reasonable stability over time and good agreements with observations. However, some clear discontinuities are caused by transitions from one dataset 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 too diffuse signature. 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).

JipingMac xie 22/12/2016 18:03
Deleted: . The

JipingMac xie 22/12/2016 18:03
Deleted:

JipingMac xie 22/12/2016 18:03
Deleted: good

JipingMac xie 22/12/2016 18:03
Deleted: assessing

JipingMac xie 22/12/2016 18:03
Deleted: method

JipingMac xie 22/12/2016 18:03
Deleted: illustrate

JipingMac xie 22/12/2016 18:03
Deleted: TOPAZ5

JipingMac xie 22/12/2016 18:03
Formatted: Indent: Left: 0,12 cm

JipingMac xie 22/12/2016 18:03
Deleted: 58

We also foresee that increasing the resolution will be well to resolve the circulation in the Nordic Seas and reduce the seasonal biases of SST and SSH.

- The system has a too sharp 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: Too thin ice, a too smooth thickness gradient 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 km d^{-1} for an ice drift generally inferior to 10 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.

JipingMac xie 22/12/2016 18:03
Deleted: improve

JipingMac xie 22/12/2016 18:03
Deleted: associated

JipingMac xie 22/12/2016 18:03
Deleted: TOPAZ5

JipingMac xie 22/12/2016 18:03
Deleted: transit

JipingMac xie 22/12/2016 18:03
Deleted: an improved

JipingMac xie 22/12/2016 18:03
Deleted: atmospheric and ocean

JipingMac xie 22/12/2016 18:03
Deleted: errors

- 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 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.

JipingMac xie 22/12/2016 18:03
Deleted: present

JipingMac xie 22/12/2016 18:03
Deleted: highly spatially

- The EnKF has proven capable to assimilate a large variety of observations, but more observations should be added. 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).

JipingMac xie 22/12/2016 18:03
Deleted: 2009

- Although efforts were made to freeze as much as possible the assimilation setting; 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 optimising 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 type of reanalysis prioritizes consistency at the expenses of accuracy, which is not the purpose of TOPAZ system. In a future reanalysis production, consistently reprocessed data sets from the ESA Climate Change Initiatives (ESA CCI) will be assimilated over the whole period (these were not available yet 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 a time and to provide input data to probabilistic weather or stand-alone sea ice models.

Acknowledgements

Thanks to Dr. P. Rampal for processing the buoy dataset for sea ice drift and 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 numbers 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.

References

Alexandrov, V., Sandven, S., Wåhlin, J., and Johannessen, O.M.: The relation between sea ice thickness and freeboard in the Artic. The Cryosphere 4: 378-380. doi: 10.5194/tc-4-373-2010, 2010.

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation. Mon. Wea. Rev. 129, 2884-2903, DOI: http://dx.doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2, 2001.

Barth, A., Canter, M., Schaeybroeck, B. V., Vannitsem, S., Massonnet, F., Zunz, V., Mathiot, P., Alvera-Azcárate, A., Beckers, J. –M.: Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the Southern Ocean, *Ocean Modelling*, 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(2), 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, in: Drange, H., Dokken, T., Furevik, T., Gerdes, R., Berger, W. (Eds.), Amer. Geophys. Union Mono. Ser. 158, 11-37, 2003

Brasseur, P., Bahurel, 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 Verron, J.: Data assimilation for marine monitoring and prediction: The

JipingMac xie 22/12/2016 18:03
Deleted: S. …andven, J.…., Wåh ... [19]

JipingMac xie 22/12/2016 18:03
Deleted: ., 2001:…: An ensemble ... [20]

JipingMac xie 22/12/2016 18:03
Deleted: M. …anter, B. V.…., ... [21]

JipingMac xie 22/12/2016 18:03
Formatted: Font:Not Italic

JipingMac xie 22/12/2016 18:03
Deleted: (9),… 22-39, ... [22]

JipingMac xie 22/12/2016 18:03
Deleted: K. A. …isæter, 2008:… ... [23]

JipingMac xie 22/12/2016 18:03
Deleted: ., 2002:…: An oceanic g ... [24]

JipingMac xie 22/12/2016 18:03
Deleted: .,… and S. …sterhus, 2 ... [25]

JipingMac xie 22/12/2016 18:03
Deleted: P. …ahurel, L.…., Berti ... [26]

MERCATOR operational assimilation systems and the MERSEA developments, Q. J. R. Meteorol. Soc., 131(613), 3561– 3582, doi: 10.1256/qj.05.142, 2005.

Bouillon, S. and Rampal, P.: Presentation of the dynamical core of neXtSIM, a new sea ice model. *Ocean Modelling*, 91(7), 23-37, doi:/10.1016/j.ocemod.2015.04.005, 2015.

Candille, G., Côté, C., Houtekamer, P. L. and Pellerin, G.: Verification of an Ensemble Prediction system against observations. Mon. Wea. Rev., 135, 2688-2699, DOI: http://dx.doi.org/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: http://dx.doi.org/10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO:2, 2003.

Chevallier, M., Salas-Mélia, D., Voldoire, 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: http://dx.doi.org/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. *Climate Dynamics*, *Special Issue : Ocean Reanalysis*, 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, http://dx.doi.org/10.3402/tellusa.v68.32437, 2016.

Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M.,

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Deleted: .,... and P. ...ampal, 20 ... [27]

JipingMac xie 22/12/2016 18:03
Deleted: L. T. ...mith, L. T. and C ... [28]

JipingMac xie 22/12/2016 18:03
Formatted: No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

JipingMac xie 22/12/2016 18:03
Formatted: Font:12 pt

JipingMac xie 22/12/2016 18:03
Deleted: D. ...alas-Mélia, A....., ... [29]

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Formatted ... [30]

JipingMac xie 22/12/2016 18:03
Formatted: Font:Helvetica

JipingMac xie 22/12/2016 18:03
Deleted: G. C.

JipingMac xie 22/12/2016 18:03
Formatted ... [31]

JipingMac xie 22/12/2016 18:03
Deleted: .

JipingMac xie 22/12/2016 18:03
Formatted ... [32]

JipingMac xie 22/12/2016 18:03
Deleted: J. -F. Lemieux et al., 2016

JipingMac xie 22/12/2016 18:03
Deleted: –

JipingMac xie 22/12/2016 18:03
Formatted: Font:Helvetica

JipingMac xie 22/12/2016 18:03
Formatted ... [33]

JipingMac xie 22/12/2016 18:03
Deleted: L. ...ertino, P....., Brass ... [34]

Mogensen, K., Oke, P., Testut, C. E., Verron, J. and Weaver, A.: Ocean data assimilation systems for GODAE, Oceanography, 22 (3), 96–109, http://dx.doi.org/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, L., 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, Quart. 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 of Nansen Environmental and Remote Sensing Center, 23pp, 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.

—, G.: The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics, 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 of Nansen Environmental and Remote Sensing Center, 30pp, 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 and Planetary Change, 125, 13-35, doi: 10.1016/j.gloplacha.2014.11.013, 2015.

JipingMac xie 22/12/2016 18:03
Deleted: P....., Oke, C. E....., Te ... [35]

JipingMac xie 22/12/2016 18:03
Deleted: Dee, D.P., S. M. Uppala, A. J. Simmons, P. Berrisford, et al., 2011:

JipingMac xie 22/12/2016 18:03
Deleted: .,... and K. ...imonsen, ... [36]

JipingMac xie 22/12/2016 18:03
Deleted: ., 1994:

JipingMac xie 22/12/2016 18:03
Deleted: 2003:

JipingMac xie 22/12/2016 18:03
Deleted: F. ...ounillon, L....., Ber ... [37]

JipingMac xie 22/12/2016 18:03
Deleted: B. ...urry, R....., Gerdes ... [38]

Hunke, E. C. and Dukowicz, J. K.: An elastic-viscous-plastic model for sea ice dynamics, J. Phys. Oceanogr., 27, 1849–1867, Doi: http://dx.doi.org/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,* 76pp, 2010.

Ilıcak, M., Drange, H., Wang, Q., Gerdes, R., Aksenov, Y., Bailey, D., Bentsen, M., Biastoch, A., Bozec, A., Böning, C., Cassou, 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., Voldoire, 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 Modelling. 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., et al.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations. The Cryosphere, 9(5), doi: 10.5194/tcd-9-1269-2015, 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 (4), 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-

JipingMac xie 22/12/2016 18:03
Deleted: J. K. …ukowicz, 1997:. ... [39]

JipingMac xie 22/12/2016 18:03
Deleted: W. H. …ipscomb, W. H ... [40]

JipingMac xie 22/12/2016 18:03
Deleted: Ilicak, M., H. Drange, H. Wang et al., 2016:

JipingMac xie 22/12/2016 18:03
Formatted: Font:10 pt

JipingMac xie 22/12/2016 18:03
Deleted: . et al, 2015. Ivanova, N ... [41]

JipingMac xie 22/12/2016 18:03
Deleted: L. …engtsson, M. W.… ... [42]

JipingMac xie 22/12/2016 18:03
Deleted: X. …ian-Kunze, N.…., M ... [43]

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. Oceanic Technol, 17, 1421– 1438, DOI: http://dx.doi.org/10.1175/1520-0426(2000)017<1421:EAABPO>2.0.CO;2, 2000.

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., Peterson, T. C., Vose, R. S. and Zhang, H. -M.: Possible artifacts of data biases in the recent global surface warming hiatus. Science, 348 (6242), 1469-1472, doi: 10.1126/science.aaa5632, 2015.

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., 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.

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.

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.,

JipingMac xie 22/12/2016 18:03
Deleted: P. A. ...ochford, P. A. a ... [44]

JipingMac xie 22/12/2016 18:03
Deleted: A. ...rguez, B....., Huan ... [45]

JipingMac xie 22/12/2016 18:03
Formatted: pb_toc_link

JipingMac xie 22/12/2016 18:03
Deleted: S. L. ...arrell, M..... L., ... [46]

JipingMac xie 22/12/2016 18:03
Deleted: G. F. ...unningham, M.. ... [47]

JipingMac xie 22/12/2016 18:03
Deleted: .,... and D. ...othrock, 2 ... [48]

JipingMac xie 22/12/2016 18:03
Deleted: ...., K. A. ...iles, K. A. L ... [49]

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:sup1, s12-s27, DOI: 10.1080/1755876X.2015.1022050, 2015.

Lien, V. S., Hjøllo, 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 Modelling, 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. http://dx.doi.org/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. In S. Levitus and A. Mishonov, editors, 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. R., 119(7), 4168-4184, doi:10.1002/2013JC009705, 2014.

JipingMac xie 22/12/2016 18:03
Deleted: S. S. ...jøllo, M. D..... S ... [50]

JipingMac xie 22/12/2016 18:03
Deleted: ., 2013:...: Unified sea i ... [51]

JipingMac xie 22/12/2016 18:03
Deleted: J. ...osanova, J. and G. ... [52]

JipingMac xie 22/12/2016 18:03
Deleted: J. ...ntonov, J. and H. . ... [53]

JipingMac xie 22/12/2016 18:03
Deleted: A. V. ...ishonov, J. I..... ... [54]

JipingMac xie 22/12/2016 18:03
Deleted: H. ...oosse, T....., Fiche ... [55]

Mathiot, P., Beatty, C. K., 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, www.geosci-model-dev.net/5/1501/2012/, 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(C6), 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 Modelling, 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. R., 114(C5), 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, 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 Trans. Geosci. Remote

JipingMac xie 22/12/2016 18:03
Deleted: C. K. ...eatty, T..... K., ... [56]

JipingMac xie 22/12/2016 18:03
Deleted: C. E. ...ong, C. E. and I ... [57]

JipingMac xie 22/12/2016 18:03
Deleted: D. ...enemenlis, D. and ... [58]

JipingMac xie 22/12/2016 18:03
Deleted: —, 2011:

JipingMac xie 22/12/2016 18:03
Deleted: G. B. ...rassington, D. A ... [59]

JipingMac xie 22/12/2016 18:03
Deleted: J. ...eiss, J. and D. ...a ... [60]

JipingMac xie 22/12/2016 18:03
Deleted: .,... and Campin, J.M., ... [61]

JipingMac xie 22/12/2016 18:03
Deleted: .,... and Morlighem, M., ... [62]

JipingMac xie 22/12/2016 18:03
Deleted: J. ...enerelli, J...., Bouti ... [63]

Sens., 50(5), 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: http://dx.doi.org/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. Quart. J. Roy. Meteor. Soc., 142, 443–454, Doi: 10.1002/qj.2663, 2016.

Roemmich, D., Church, J, Gilson, J., Monselesan, D., Sutton, P. and Wijffels, S.: Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change* 5, 240–245. doi:10.1038/nclimate2513, 2015.

Rothrock, D.A. and Wensnahan, M.: The accuracy of sea-ice drafts measured from U. S. Navy submarines. J. Atmos. Oceanic Technol. 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(2):361-371, doi: 10.1111/j.1600-0870.2007.00299.x, 2008.

Sakov, P., Counillon, F., Bertino, L., Lisæther, K. A., Oke, P. R. and Korablev, 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.

Samuelsen, 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(7), 2187-2202, DOI: 10.5194/gmd-8-2187-2015, 2015.

Schweiger, A., Lindsay, R., Zhang, J. L., Steele, M., Stern, H. and Kwok, R.:

JipingMac xie 22/12/2016 18:03
Deleted: .,... and T. ...mith, 1994 ... [64]

JipingMac xie 22/12/2016 18:03
Deleted: J. ...hurch, J...., Gilson ... [65]

JipingMac xie 22/12/2016 18:03
Deleted: M. ...ensnahan, 2007: ... [66]

JipingMac xie 22/12/2016 18:03
Deleted: D. B. ...ercival, D. B. ar ... [67]

JipingMac xie 22/12/2016 18:03
Deleted: P. R. ...ke, 2008: ... [68]

JipingMac xie 22/12/2016 18:03
Deleted: F. ...ounillon, L....., Ber ... [69]

JipingMac xie 22/12/2016 18:03
Deleted: C. ...ansen, C. and H. ... [70]

JipingMac xie 22/12/2016 18:03
Deleted: R. ...indsay, J. L....., Zh ... [71]

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., 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(8), doi:10.1029/2005GL025624, 2006.

Simon, E., Samuelsen, 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 Sys., 152, 1-17, dio: 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, Geophysical Research Letters, 38(19), 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, in OCEAN 2007-Eurrope, pp. 1-4, IEEE, 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:http://dx.doi.org/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 sea ice thickness: algorithm baseline, product

JipingMac xie 22/12/2016 18:03
Deleted: .,… and J. …hang, 201 ... [72]

JipingMac xie 22/12/2016 18:03
Deleted: T. …amoshida, M.…., I ... [73]

JipingMac xie 22/12/2016 18:03
Deleted: Simmons, A., S. Uppala, D. Dee, and S. Kobayashi, 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, ECMWF Newsletter, 110, 25-35. ... [74]

JipingMac xie 22/12/2016 18:03
Deleted: R. …wok, R. and D. ... [75]

JipingMac xie 22/12/2016 18:03
Deleted: C. …onlon, M.…., Marti ... [76]

JipingMac xie 22/12/2016 18:03
Deleted: R. …orley,…R. and W. ... [77]

JipingMac xie 22/12/2016 18:03
Deleted: L. …aleschke, N.…., M ... [78]

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., Francis, C., Baker, S., Bouzinac, C., et al: 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. Geophysical Research Letters, 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 alongtrack 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.

JipingMac xie 22/12/2016 18:03
Deleted: .,… and …arotzke, J., … [79]

JipingMac xie 22/12/2016 18:03
Deleted: R. J. …orne. 1980: … [80]

JipingMac xie 22/12/2016 18:03
Deleted: D. A. …othrock, 2005: … [81]

JipingMac xie 22/12/2016 18:03
Deleted: C. …rancis, S.…., Bake … [82]

JipingMac xie 22/12/2016 18:03
Deleted: K. …agaard, K. and T. … [83]

JipingMac xie 22/12/2016 18:03
Formatted … [84]

JipingMac xie 22/12/2016 18:03
Deleted: F. …ounillon, J.…., Zhu … [85]

JipingMac xie 22/12/2016 18:03
Deleted: P. …ampal, N.…., Ivan … [86]

Table 1. Overview of assimilated observations per cycle, average numbers for the cycles during which the observations are present. (1) The resolution of ice concentration product increased to 10 km. Unless specified, all observations from http://marine.copernicus.eu

Type	Number	After SO	Spacing	Resolution	Period	Provider
SLA	9x10^4	5x10^4	Track	7 km	1992-2013	CLS
SST	6x10^3	6x10^3	Gridded	100 km	1990-1998	Reynolds SST from NCDC (http://www.nhc.noaa.gov/aboutsst.shtml)
SST	2x10^6	2.4x10^5	Gridded	5 km	1998-2013	OSTIA from UK Met Office
In-situ T/S	3x10^4	5x10^3	Point	-	1990-2013	Ifremer + other
ICEC (SSM/I)	9x10^4	5x10^4	Gridded	25 km	1990-2002	OSISAF
ICEC (AMSR-E)	1.6x10^5	5x10^4	Gridded	12.5 km(1)	2002-2013	OSISAF
ICEC (AMSR-E)	1.6x10^5	5x10^4	Gridded	12.5 km	2008-2009	AMSR-E (http://nsidc.org/data/amsre/)
Ice drift (CERSAT)	6x10^3	10^3	Gridded	35 km	2002-2010	Ifremer
Ice drift (OSISAF)	4x10^3	10^3	Gridded	62.5 km	2011-2013	OSISAF
Total	2.3x10^6	4x10^5				

xie 22/12/2016 17:59
Deleted: ... [1]

xie 22/12/2016 17:59
Deleted: 4x10^4

xie 22/12/2016 17:59
Deleted: 5x10^3

xie 22/12/2016 17:59
Deleted: 5x10^3

xie 22/12/2016 17:59
Deleted: 62.5

xie 22/12/2016 17:59
Deleted: 5x10^3

xie 22/12/2016 17:59
Deleted: 5x10^3

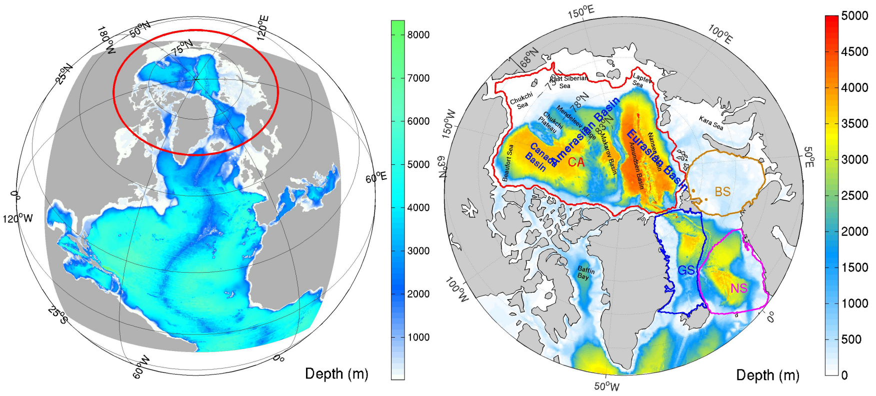


Fig 1. **Left**: Bottom topography in the whole TOPAZ4 domain. The red line delimits the Pan-Arctic region north of 63°N. **Right**: Definition of sub-basins and marginal seas. The domain is divided into the four sub-regions delimits by the colored lines: the Central Arctic in red (CA), the Greenland Sea in blue (GS), the Barents Sea in orange (BS), and the Norwegian Sea in magenta (NS).

xie 22/12/2016 17:59

Deleted: 2

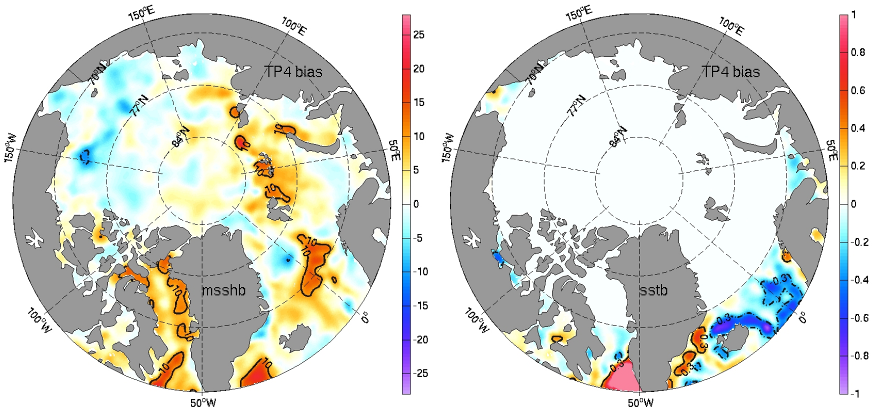


Fig 2. Estimates of the mean SSH bias (**Left**) and the SST bias (**Right**) obtained at last analyzed date by online parameter estimation. In the left panel, the solid (dashed) line indicates the 10 (-10) cm isolines. In the right panel, the solid (dashed) line indicates the 0.3 °C (-0.3 °C) isolines.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

xie 22/12/2016 17:59

Deleted:

Unknown

Formatted: Font:Arial

xie 22/12/2016 17:59

Deleted: 2

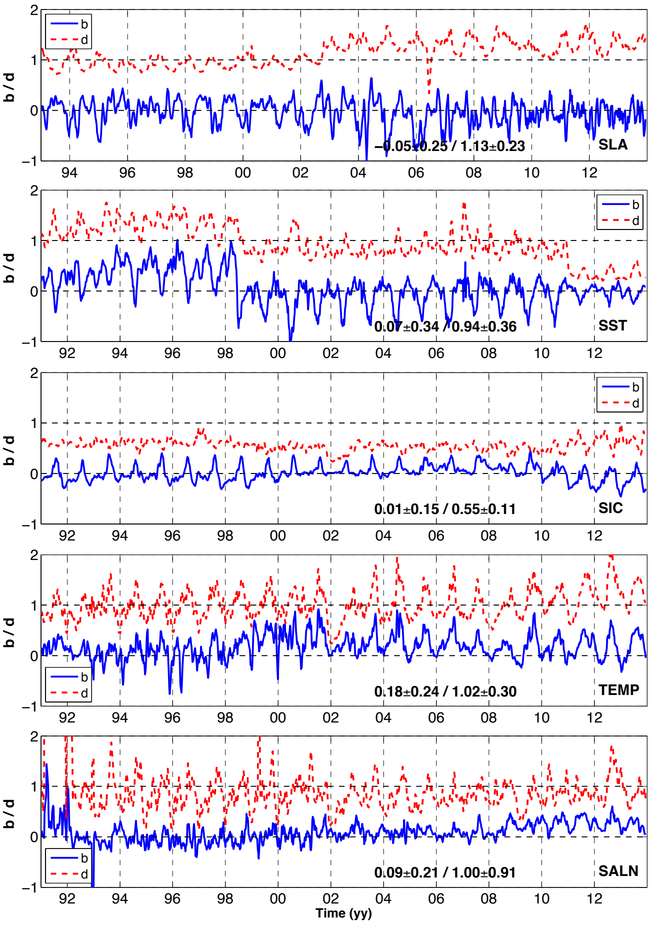


Fig. 3 Time series of *b* (blue line) and *d* (dashed red line) of SLA, SST, SIC, temperature and salinity from in situ respectively in the Arctic region. They are filtered by a smoothing average within 28 days. The averaged (standard deviation) of b and d are shown in the panels.

Unknown
Formatted: Font:Arial

xie 22/12/2016 17:59
Deleted: 2

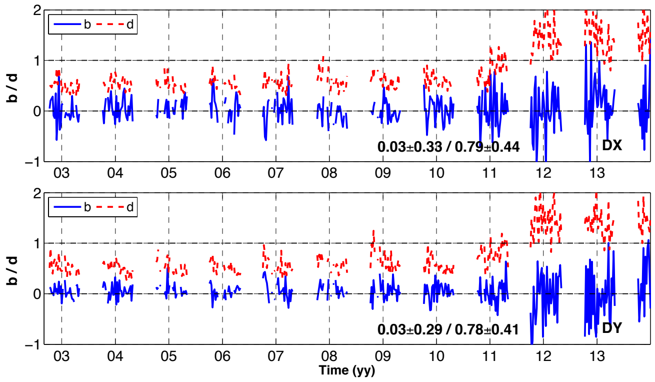


Fig. 4 Time series of *b* (blue line) and *d* (dashed red line) about the zonal (DX) and meridional (DY) drifts of sea ice in the Arctic. The averaged (standard deviation) of b and d are shown in the panels.

Unknown
Formatted: Font:Bold

xie 22/12/2016 17:59
Deleted: 2

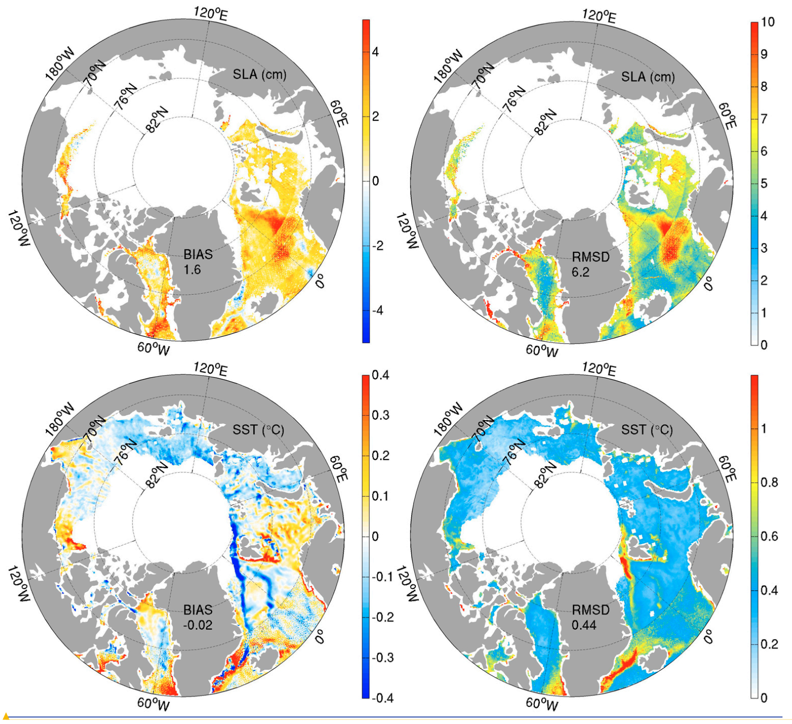


Fig 5. **Top**: Residual bias (left) and RMSD (right) between the daily average SLA from the reanalysis and the assimilated along-track SLA data averaged over the period 1993-2013 (unit: cm). **Bottom**: The corresponding residual bias (left) and RMSD (right) between the daily average SST from the reanalysis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

Unknown
Formatted: Font:Arial

xie 22/12/2016 17:59
Deleted: 3

xie 22/12/2016 17:59
Deleted:

... [2]

Unknown
Formatted: Font:Arial, Bold

xie 22/12/2016 17:59
Deleted: 2

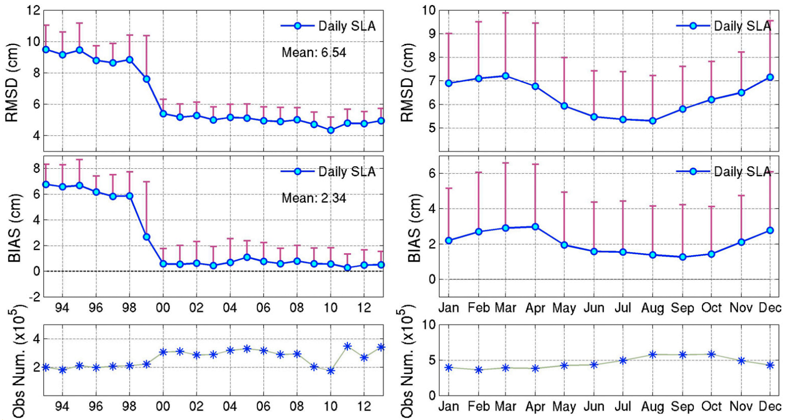


Fig 6. Left: Yearly averaged estimates of daily SLA RMSD (upper) and the residual bias (middle) of the TOPAZ reanalysis calculated against the along-track SLA available in the Pan-Arctic region (unit: cm). The error bars denote the standard deviations of the daily statistics within each year. The bottom 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.

Unknown
Formatted: Font:Arial, Bold

xie 22/12/2016 17:59
Deleted: 4

xie 22/12/2016 17:59
Deleted:

Unknown
Formatted: Font:Arial

xie 22/12/2016 17:59
Deleted: 2

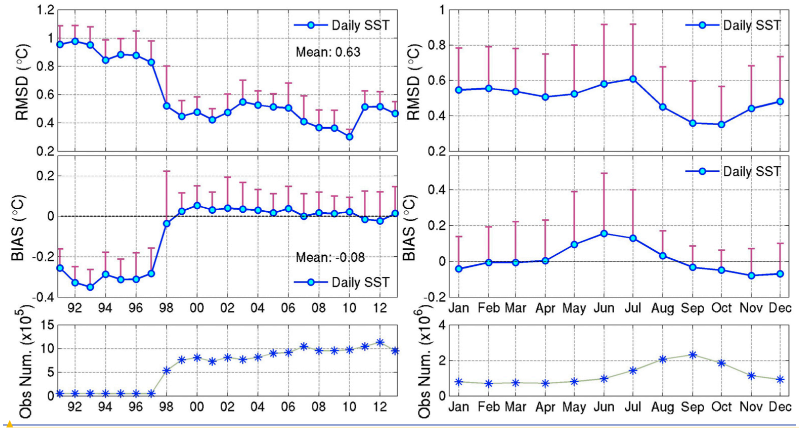


Fig 7. Same as the previous figure but for SST over the period 1991-2013 (unit: °C).

Unknown
Formatted: Font:Arial

xie 22/12/2016 17:59
Deleted: 5

xie 22/12/2016 17:59
Deleted: for

xie 22/12/2016 17:59
Formatted: Font:Not Bold

xie 22/12/2016 17:59
Formatted: List Paragraph, Centered, Indent: Left: 0 cm, Hanging: 1,98 cm

xie 22/12/2016 17:59
Formatted: List Paragraph, Indent: Left: 1 cm, Hanging: 0,98 cm

xie 22/12/2016 17:59
Formatted: List Paragraph, Centered, Indent: Left: 0 cm, Hanging: 1,98 cm

xie 22/12/2016 17:59
Deleted: 2

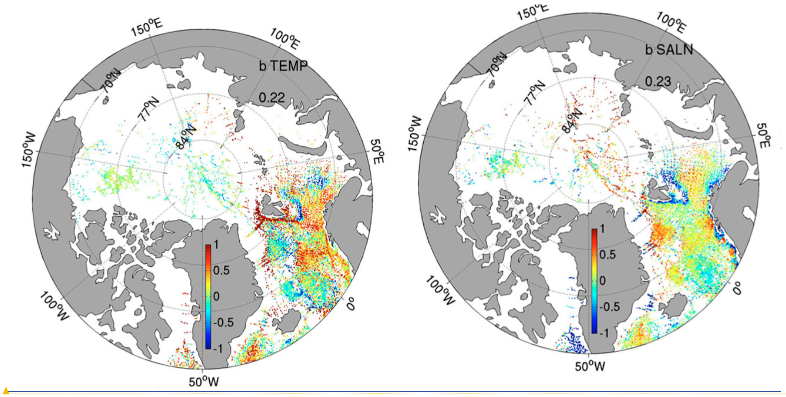
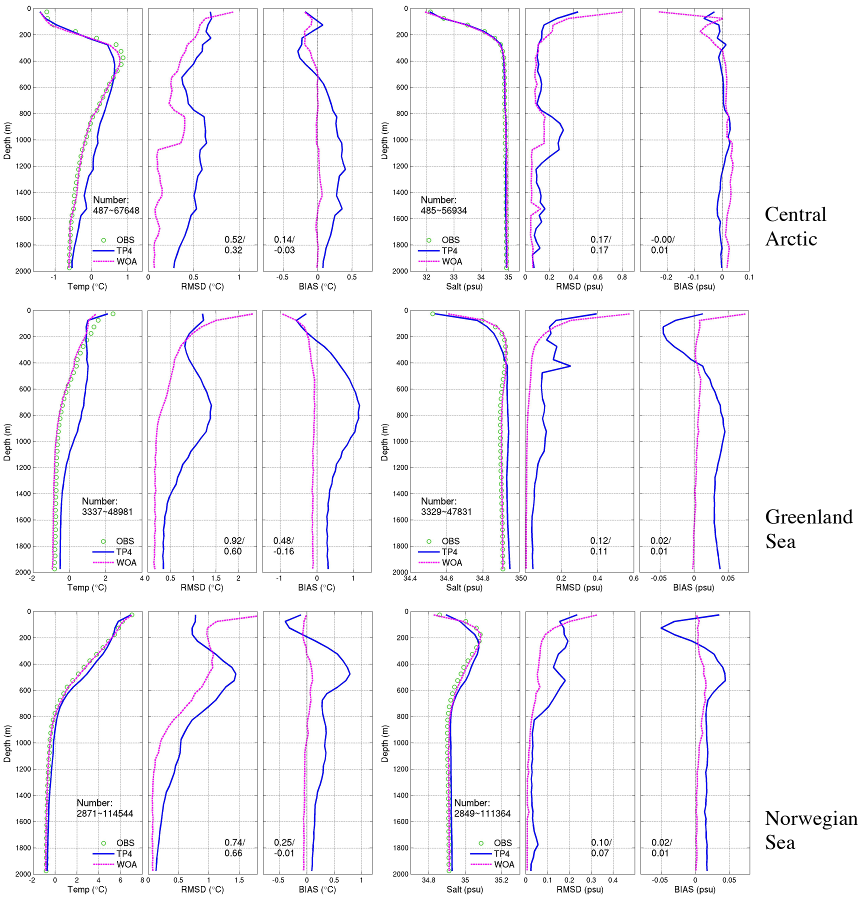


Fig. 8 Spatial distribution of *b* for temperature (left) and salinity (right) from in situ during the period from 1991 to 2013. The observation number in a grid is required more than 30. Note that profiles may end at different depths and cause spottiness.

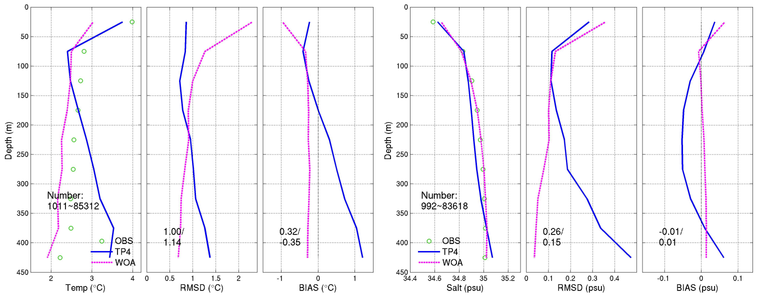
Unknown
Formatted: Font:Arial

xie 22/12/2016 17:59
Deleted: 2



xie 22/12/2016 17:59

Deleted: 2



Barents
Sea

Fig 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 circle is the observations, the blue lines are the TOPAZ reanalysis, and the pink lines are from the WOA13 climatology. The numbers in the first-column subpanels are the minimal and maximal number of observations available in each of 50 m depths; the upper numbers in the other-column subpanels are the mean estimate in vertical for TOPAZ reanalysis, and the lower numbers is for WOA13.

xie 22/12/2016 17:59

Deleted: 6

xie 22/12/2016 17:59

Deleted: 2

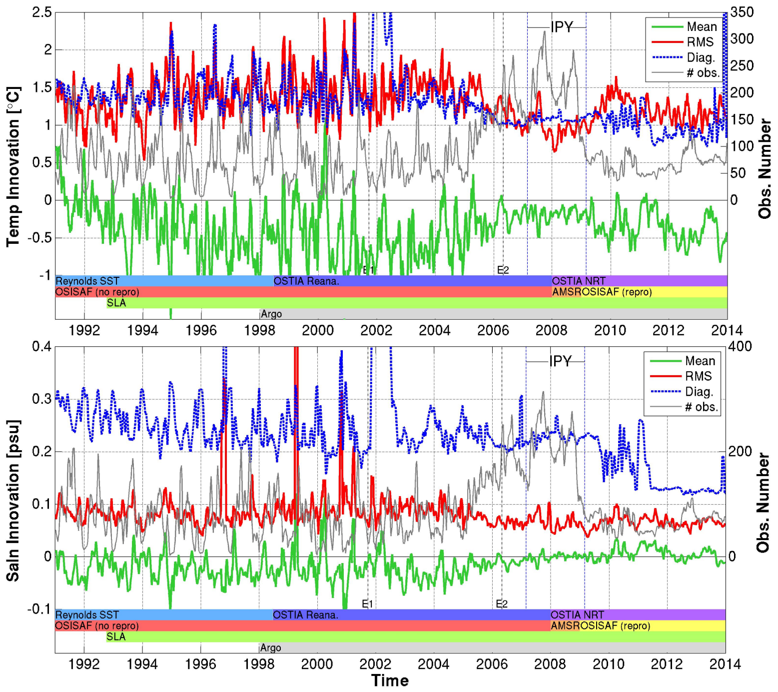


Fig 10. Time series of innovation statistics for temperature (**top**) and salinity (**bottom**) observed at the depth of between 300-800 m depths. The bias is plotted with a green line, the RMSD is in red and the number of assimilated observations is plotted with a grey line. The blue dashed line indicates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

Deleted:

Unknown

Formatted ... [3]

xie 22/12/2016 17:59

Deleted: 7…0. Time series of innov ... [4]

xie 22/12/2016 17:59

Deleted: major

xie 22/12/2016 17:59

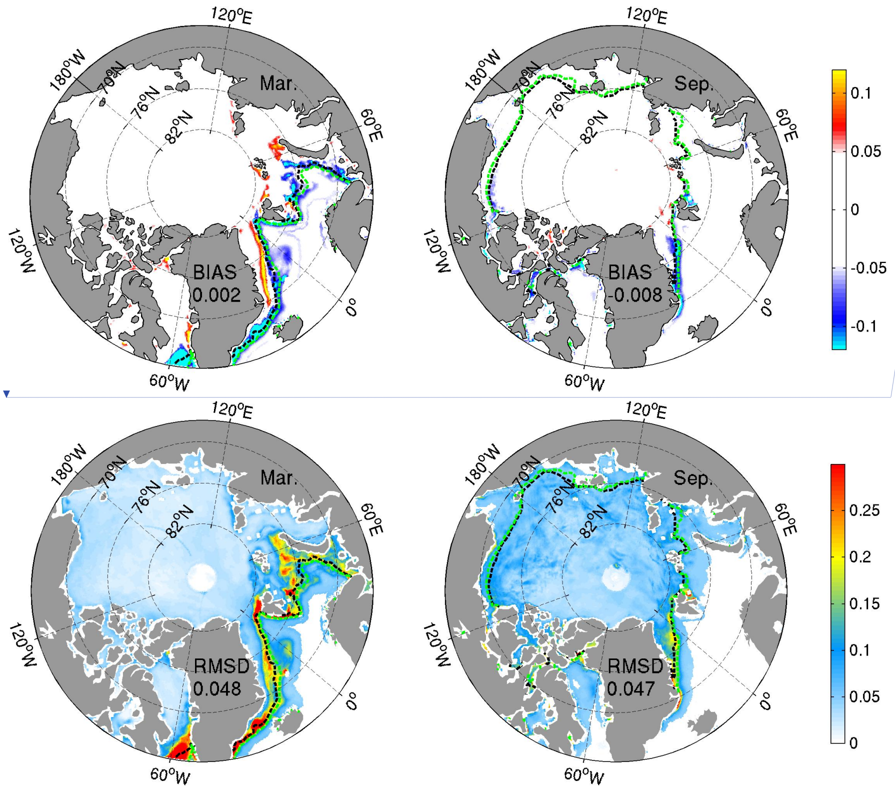
Formatted: Font:Arial, 12 pt, Bold

xie 22/12/2016 17:59

Formatted: Centered, Indent: Left: 0 cm, First line: 0 cm

xie 22/12/2016 17:59

Deleted: 2



xie 22/12/2016 17:59
Deleted:

Fig 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).

xie 22/12/2016 17:59
Deleted: 8

xie 22/12/2016 17:59
Deleted: 2

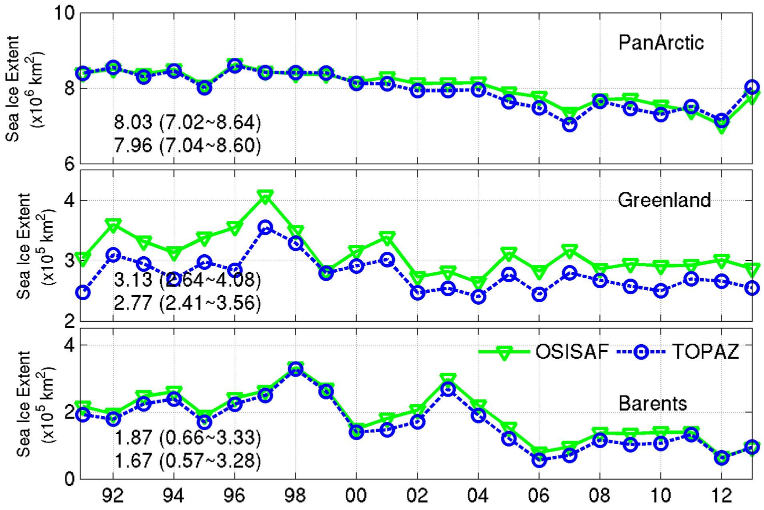


Fig 12. Yearly time series of the sea ice extent in the Pan-Arctic region, the Greenland Sea, and the Barents Sea from TOPAZ reanalysis (dashed) and OSISAF (solid).

xie 22/12/2016 17:59

Deleted: 9

xie 22/12/2016 17:59

Deleted: 2

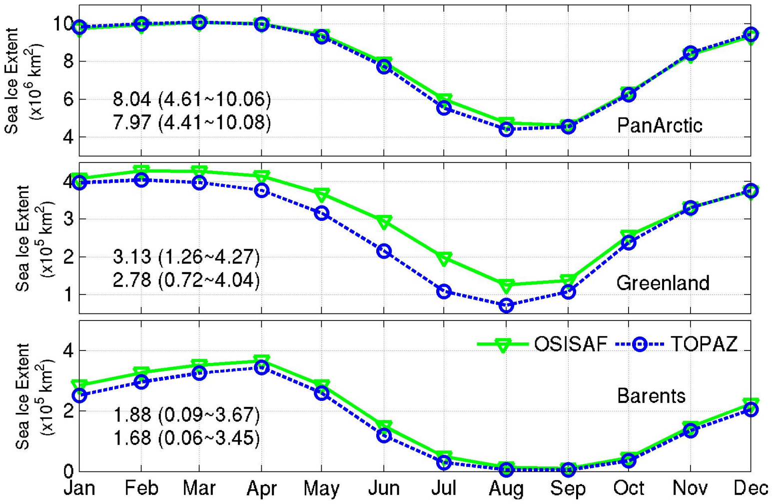


Fig 13. Seasonality of the sea ice extents in the TOPAZ reanalysis (blue line) and OSISAF (green line) in the Pan-Arctic Ocean, Greenland Sea, and Barents Sea regions.

xie 22/12/2016 17:59

Deleted: 10

xie 22/12/2016 17:59

Deleted: 2

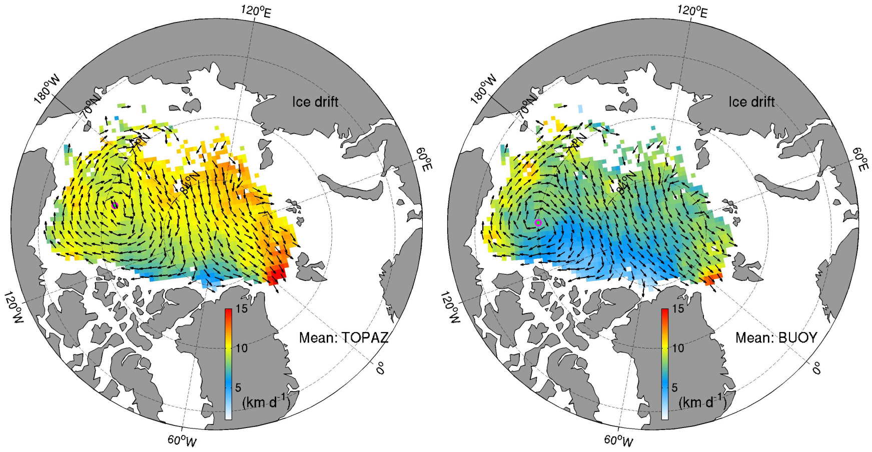


Fig 14. Sea ice drift vectors (*arrows*) and speeds (*color shading*) averaged over the period 1991-2011 for (a) TOPAZ reanalysis and (b) IABP buoys. The center of the anticyclonic Beaufort Gyre is marked with a magenta circle at (155°W, 78.1°N) in the TOPAZ reanalysis and (145°W, 77°N) in the observations respectively.

xie 22/12/2016 17:59
Deleted: 11

xie 22/12/2016 17:59
Deleted: 2

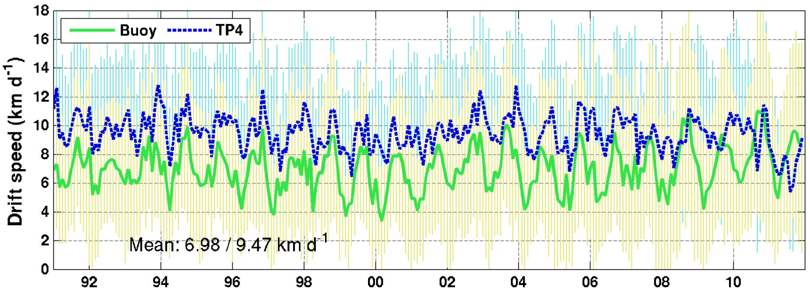


Fig 15. Monthly time series of the daily averaged sea ice drift speeds in the Central Arctic from the TOPAZ reanalysis (blue line) and the IABP buoys (green line) during 1991-2011. The error bars represent the standard deviations of the daily estimates for each month.

xie 22/12/2016 17:59
Deleted: 12

xie 22/12/2016 17:59
Deleted: 2

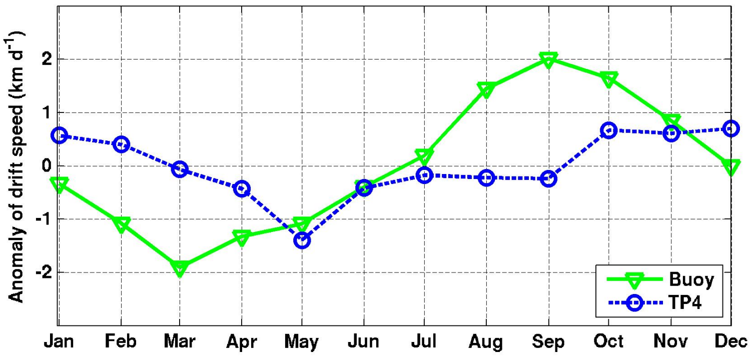


Fig 16. Seasonality of the sea ice drift velocities from the reanalysis and the buoy during 1991-2011.

xie 22/12/2016 17:59

Deleted: 13

xie 22/12/2016 17:59

Deleted: 2

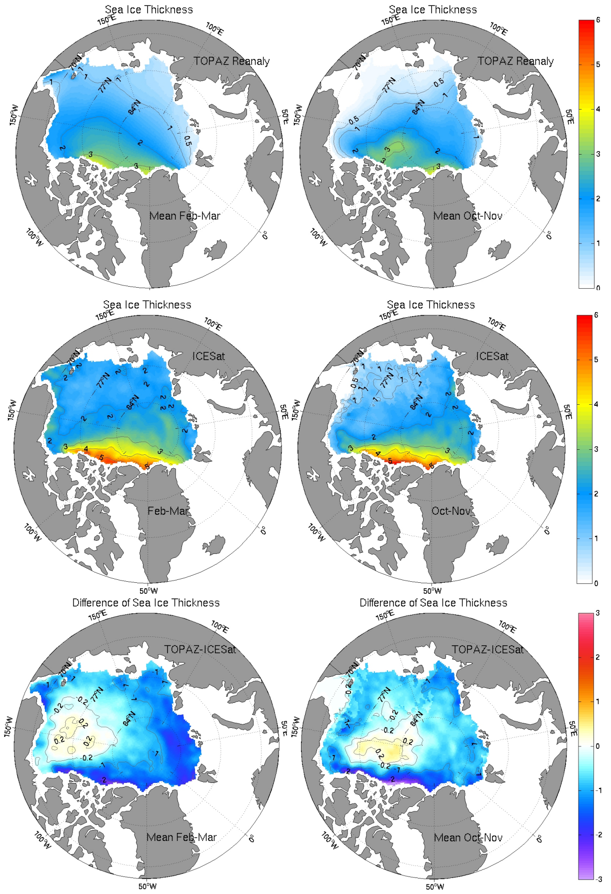


Fig 17. Mean sea ice thicknesses from TOPAZ (upper) and ICESat (middle), and their difference (bottom) for February-March (*in left column*) and October-November (*in right column*) averaged over the period 2003-2008.

xie 22/12/2016 17:59

Deleted: 14

xie 22/12/2016 17:59

Deleted: 2

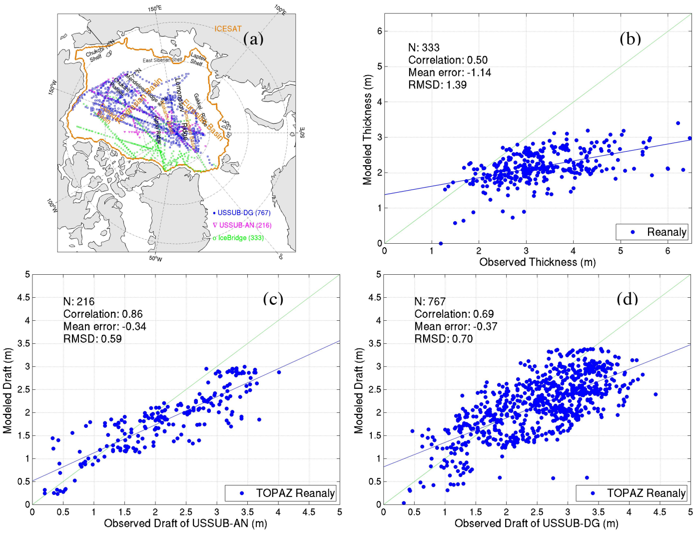


Fig 18. Validation the sea ice thickness in the TOPAZ reanalysis versus available in situ observations. (a) Locations of in situ observations available from IceBridge, USSUB-AN and USSUB-DG in the Central Arctic. Regression analysis of TOPAZ reanalysis (b) vs. IceBridge; (c) vs. USSUB-AN; (d) vs. USSUB-DG.

xie 22/12/2016 17:59
Deleted: 15

xie 22/12/2016 17:59
Deleted: 2

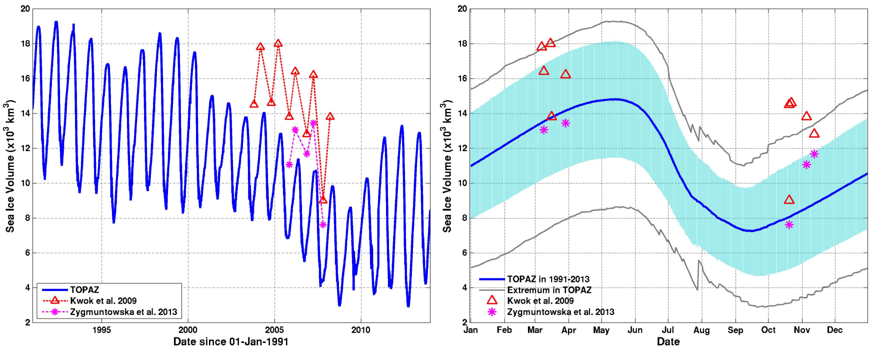


Fig 19. **Left**: Time series of the daily averaged sea ice volume in the Arctic from the TOPAZ4 (blue line) and the observations from Kwok et al. (2009) and from Zygmuntowska et al. (2013). **Right**: Daily time series of the averaged sea ice volume in the Arctic from the TOPAZ4 for the period 1991-2013 (blue line) and the standard deviation shown as the cyan error-bar. The gray lines represent the extreme volumes in the 23 years. The triangle and start markers are the observations estimated by Kwok et al. (2009) and Zygmuntowska et al. (2013) respectively.

xie 22/12/2016 17:59

Deleted: 16

xie 22/12/2016 17:59

Deleted: 2

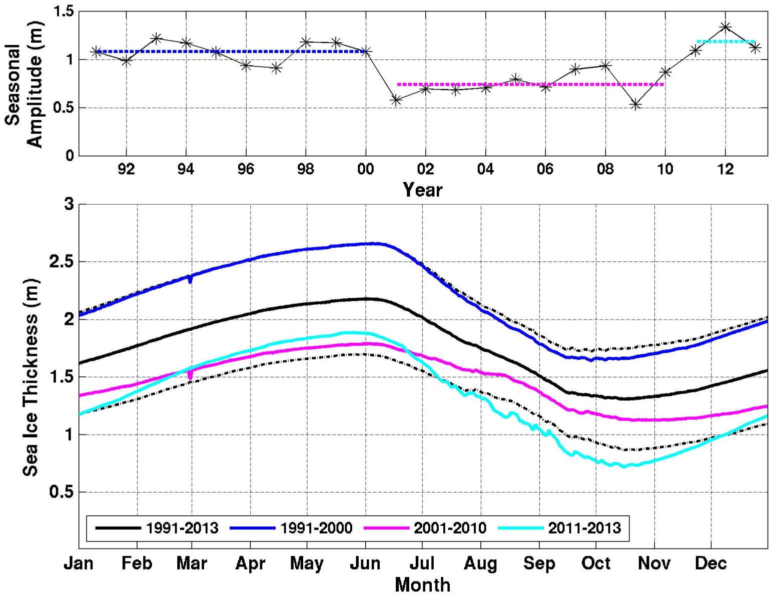


Fig 20. **TOP**: Yearly time series of the seasonal amplitudes of the mean sea ice thickness in the Central Arctic with the solid black line. The dashed lines represent the averaged estimate for: 1991-2000, 2001-2010, and 2011-2013 (1.08, 0.74, and 1.18 m respectively). **Bottom**: Daily time series 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.

xie 22/12/2016 17:59

Deleted: 17

xie 22/12/2016 17:59

Formatted: Font:+Theme Body, Not Bold

xie 22/12/2016 17:59

Formatted: Justified, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

xie 22/12/2016 17:59

Deleted: 2