AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2015JC011005

Special Section:
Forum for Arctic Modeling and Observational Synthesis (FAMOS): Results and Synthesis of Coordinated Experiments

Sea-surface temperature and salinity product comparison against external in situ data in the Arctic Ocean

J. N. Stroh[1], Gleb Panteleev[1,2], Sergey Kirillov[3], Mikhail Makhotin[4], and Natalia Shakhova[1,2]

[1]International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, Alaska, USA, [2]Institute of Natural Resources, National Research Tomsk, Polytechnic University, Tomsk, Russia, [3]Center for Earth Observation Science, University of Manitoba, Winnipeg, Manitoba, Canada, [4]Arctic and Antarctic Research Institute, Saint Petersburg, Russia

Abstract Sea-surface temperature and salinity (SST/S) in the Arctic Ocean (AO) are largely governed by sea-ice and continental runoff rather than evaporation and precipitation as in lower latitude oceans, and global satellite analyses and models which incorporate remotely observed SST/S may be inaccurate in the AO due to lack of direct measurements for calibrating satellite data. For this reason, we are motivated to validate several satellite sea-surface temperature (SST) data products and SST/S models by comparing gridded data in the AO with oceanographic records from 2006 to 2013. Statistical analysis of product-minus-observation differences reveals that the satellite SST products considered have a temperature bias magnitude of less than 0.5°C compared to ship-based CTD measurements, and most of these biases are negative in sign. SST/S models also show an overall negative temperature bias, but no common sign or magnitude of salinity bias against CTD data. Ice tethered profiler (ITP) near-surface data span the seasons of several years, and these measurements reflect a sea-ice dominated region where the ocean surface cannot be remotely observed. Against this data, many of the considered models and products show large errors with detectable seasonal differences in SST bias. Possible sources of these errors are discussed, and two adjustments of product SST on the basis of sea-ice concentration are suggested for reducing bias to within less than 0.01°C of ITP near-surface temperatures.

Key Points:
- SST products and models agree well with CTD data in the Arctic Ocean away from freshwater sources
- Satellite-derived products and model SST poorly match ITP near-surface data
- Product SST and freezing temperature may be corrected from sea-ice concentration and ITP data

Correspondence to:
J. N. Stroh,
jnstroh@alaska.edu

Citation:
Stroh, J. N., G. Panteleev, S. Kirillov, M. Makhotin, and N. Shakhova (2015), Sea-surface temperature and salinity product comparison against external in situ data in the Arctic Ocean, *J. Geophys. Res. Oceans*, *120*, 7223–7236, doi:10.1002/2015JC011005.

Received 28 MAY 2015
Accepted 13 OCT 2015
Accepted article online 16 OCT 2015
Published online 6 NOV 2015

1. Introduction

Sea-surface temperature and salinity (SST and SSS, respectively, or SST/S collectively) are designated by the World Meteorological Organization as "essential climate variables" [*GCOS*, 2011]. At the interface of the ocean and atmosphere, they play a role in coupling highly dynamical components of the global climate system and are diagnostic of the present ocean state. Tropical and midlatitude SST/S are driven largely by precipitation and evaporation, poleward heat transfer via ocean currents, and large basin-scale processes. In contrast, the SST/S of Arctic Ocean (AO) are strongly influenced by sea-ice and related melt waters, brine rejection, continental runoff onto broad continental shelves, and the upward flux of heat from the deeper warm ocean.

The AO is also marked by a cold, fresh surface layer in contrast to the monotonic decrease of temperature with depth common to most world oceans [*Comiso*, 2010]. Surface salinity therefore plays a significant role in the dynamics of the near-surface waters; density of cold water is more sensitive to salinity than temperature. While the AO is small compared to other oceans, a comprehensive understanding of its dynamics is inhibited by the cost and difficulty of directly measuring this extreme environment. At present, observations are dominated by remotely sensed surface data [cf. *Emery et al.*, 2001; *Kawai and Wada*, 2007; *Donlon et al.*, 2012]. Satellite-derived observations provide much greater spatial coverage than sparse in situ methods, albeit at lower resolution.

The spectral data which compose satellite SST/S observations are restricted to the very upper-most ocean layers. Infra-red (IR) sensors measure "skin" SST at micrometer depths where thermal transfer cools the interface. Below the skin layer lies a meters-thick region where temperatures may exhibit fluctuations due to daily insolation and atmospheric wind. Microwave (MW) sensors which penetrate to millimeter depths are able to measure SST/S at the top of this sub-surface layer where direct solar influences are most pronounced. The base temperature of this layer, independent of diurnal variability, defines foundational SST and is sought as the best near-surface representative of temperature.

The spatial resolution of IR sensors is on the order of 1 km, while that of MW sensors is an order larger. Satellite data products may incorporate observations from several sensors with different resolutions and penetration depths to produce accurate representations of SST. Remotely sensed emissions spectra must therefore be adjusted to foundational SST/S by accounting for cool-skin effects, diurnal insolation, wind, precipitation, and sensor inaccuracies [*Fairall et al*., 1996]. Such techniques for empirical calibration of satellite measurements are not discussed here but may be found in *Robinson* [2004] and product documentations. Limited in situ data available for tuning and calibrating these algorithms may lead to SST/S inaccuracy in the AO where the mechanisms influencing the vertical temperature structure differ from those in data-abundant regions.

Validation and inter-comparison studies of gridded global satellite SST analyses use in situ data such as Argo drifters [*Martin et al*., 2012] and other inventories of global in situ measurements [*Dash et al*., 2012] with sparse AO coverage. *Chen et al*. [2002], *Corlett et al*. [2006], and *Høyer et al*. [2012] use in situ buoy data for AO satellite SST calibration and validation, with the latter studies augmenting their comparison sets with ship and radiometer records, respectively. Recent modeling studies by *Gammelsrød et al*. [2009] and *Nguyen et al*. [2009, 2011] include in situ T/S for comparison of hydrographic transport and verification of calibrated model parameters, respectively. *Sakov et al*. [2012] compare model T/S to in situ profiles to validate a data-assimilated Arctic model and assess its sensitivity to sub-ice ITP data.

Concern for the uncertainties in transforming satellite-data into foundational SST/S estimates motivates this study which validates selected global satellite SST products and SST/S models in the Arctic region. The relationship of gridded data dependent on satellite temperature measurements to external in situ SST/S measurements is explored statistically in this paper. The investigation relies heavily on the contents of an online-accessible International Polar Year 2007–2008 (IPY) measurement database (http://oregon.iarc.uaf.edu/dbaccess.html) containing more than 12,000 ship-cast CTD profiles in the Arctic region spanning 2006–2011. Section 2 summarizes the satellite analyses, model outputs, and in situ data used in this comparison. Section 3 describes the comparison methodology, section 4 discusses results of the product-to-data correspondences, and the final section summarizes results. Use of SST, SSS, and SST/S throughout the text refer to foundational values unless otherwise noted.

2. Data

Two types of gridded SST/S are discussed here: SST analysis and SST/S model products. The former consist of SST maps and other fields synthesized from algorithmic processing of satellite data and typically in situ data when available. The latter products are generated by programs driven by primitive-equation models which use observational data for parameter calibration or state adjustment (i.e., data assimilation). The terms "analyses" and "models" hereafter succinctly refer to satellite-derived data sets and data-conditioned model output, respectively, and "product" is used to describe data from either source. Several criteria aided in the selection of the product data sets. The data needed to be freely available online, include gridded daily coverage of the Arctic regions between 65°N and 82°N for the majority of years 2006–2009 or 2010–2013.

2.1. Analysis Products

Briefly discussed in this section are the included satellite-derives SST products which meet the above-listed requirements. Satellite SSS analyses for the AO meeting the desired criteria were not available at the time of this study. Three included satellite products target at-depth SST: the NOAA 1/4° Optimal Interpolation SST version 2 (OISSTv2) [*Reynolds et al*., 2007] available for 1981-present at http://www.ncdc.noaa.gov/oisst, the NCEP Marine Modeling and Analysis Branch 1/12° high-resolution Real-Time Global SST (RTG_HR) [*Thiébaux et al*., 2003; *Gemmill et al*., 2007] available for 2006-present at ftp://polar.ncep.noaa.gov/pub/history/sst/ophi/, and the Naval Oceanographic Office 1/10° SST analysis (K10_SST) available for 2008-present at http://podaac.jpl.nasa.gov/dataset/NAVO-L4HR1m-GLOB-K10_SST. Included products which target foundational SST are: the UK Met Office 1/20° Operational SST and Sea Ice Analysis (OSTIA) [*Stark et al*., 2007; *Donlon et al*., 2012] available for April 2006-present at http://www.myocean.eu, the NASA Jet Propulsion Laboratory 1/100° Multi-scale Ultra-high Resolution SST (MURSST) [http://mur.jpl.nasa.gov/multi_resolution_analysis.php] available for 2002-present at http://podaac.jpl.nasa.gov/Multi-scale_Ultra-high_Resolution_MUR-SST, the Australian Bureau of Meteorology 1/4° Global Australian Multi-Sensor SST Analysis (GAMSSA)

[*Zhong and Beggs*, 2008] available for late 2008-present at ftp://podaac-ftp.jpl.nasa.gov/allData/ghrsst/data/L4/GLOB/ABOM/GAMSSA_28km, and the Remote Sensing Systems 1/4° MW-IR SST version 4 (RSS_SST) available for 2002-present at http://www.remss.com/measurements/sea-surface-temperature/.

All analyses are based on MW and IR satellite data except for RTG_HR, which incorporates no MW data, and OISSTv2, which incorporated MW only until October 2011. Algorithms used by the analyses to combine satellite observations and other data into complete gridded fields generally are based either on optimal interpolation or related variational methods. In addition to satellite data, analyses often use in situ data from ships and buoys in their optimization; however, K10_SST and RSS_SST analyses do not use in situ data, and MURSST uses this data for quality control rather than in its variational optimization.

2.2. Model Products

The first SST/S model output to be compared comes from the NASA JPL Estimating the Circulation and Climate of the Ocean, Phase II project. The MITgcm-based model generates full-depth global ocean output together with an interactive sea-ice component at 1/4° grid spacing, then assimilates satellite and in situ observations via kernel and adjoint methods [*Menemenlis et al*., 2005]. The data set used in this work is the geographically interpolated NASA/JPL Cube92 model output available at ftp://ecco2.jpl.nasa.gov/data1/cube/cube92. The two-dimensional daily averaged SST and SSS fields from 2006 to 2012 of that data set are herein referred to as ECCO2 data, and represent quantities averaged over the top 10 m.

The second model output is produced as part of the US Global Ocean Data Assimilation Experiment. This production system data-optimizes the model state calculated by the Hybrid Coordinate Ocean Model (HYCOM) using the Navy Coupled Ocean Data Assimilation [*Cummings*, 2005]. This three-dimensional variational system assimilates satellite and in situ observations including altimeter data, SST, and sea-ice. This study specifically uses the 0 m vertical level of the three-dimensional temperature and salinity (T/S) from HYCOM + NCODA Global analysis GLBa0.08 experiments 90.6–90.1 (https://hycom.org/dataserver/glb-analysis). These fields correspond to SST and SSS from late 2008–present and are referred to herein as HYCOM data for brevity.

2.3. In Situ Observations

The direct observational data used for validation include: a rich collection of hydrographical measurements from the 2007–2008 IPY collaboration which extends from 2006 through 2012; hydrographical profiles from Woods Hole Oceanographic Institute (WHOI) cruises to the Beaufort Sea during 2006–2013 and from international cruises to the Eurasian Arctic shelf during 2011–2013; and WHOI ice-tethered profilers (ITPs) adrift starting in 2006 [*Krishfield et al*., 2008; *Toole et al*., 2011]. This data is organized into three groups: sets of hydrographical profiles collected during 2006–2009 and 2010–2013, and one set of near-surface measurements from ITPs during 2006–2013. The first groups consist of ship-based rosette CTD profiles collected either in the years around the recent IPY or in the years after 2009, supplemented by a small number of expendable CTD (xCTD) readings. For brevity, these groups are identified as "IPY CTD" and "post-IPY CTD" hereafter. The third group comprises T/S data obtained from shallow microCAT sensors attached to ITPs as well as the upper-most profiler readings, and is referred to as "ITP."

The IPY CTD set includes the 2006–2009 ship-based CTD and xCTD profiles from the IPY database as well as Beaufort Sea and Siberian Shelf rosette CTD and xCTD profiles from cruises conducted during 2006–2009. Ship-based CTD observations from 2010 to 2013 contained in the IPY database are included in the post-IPY CTD set. Much of the IPY and post-IPY CTD data were obtained in preprocessed form with profiles resolved to less than 1 m depth using undocumented methods. Therefore, any data between 0 m and 3 m are ignored as its origins are often uncertain. The ITP data set was obtained online in the form of roughly 30,000 time-averaged profiles from the WHOI website (http://www.whoi.edu/website/itp), and this data was filtered to retain only data above 10 m depth.

All in situ measurements used contain T/S data; records lacking either were omitted. Profiles with values identically zero or unrealistic temperatures beyond the range −2.2–30°C or salinities beyond 0–40 PSU in the top 10 m are omitted. The inherent instrument errors associated with these data sets are small. CTD sensors typically have accuracies on order 10^{-3} °C (10^{-2} °C for xCTD) for temperature and order 10^{-3} PSU (10^{-2} PSU for xCTD) for salinity [*Sy and Wright*, 2000], with similar initial accuracies established for microCAT sensors. The IPY CTD, post-IPY CTD, and ITP data sets consist of roughly 9800, 1230, and 14,000 near-surface profile representatives,

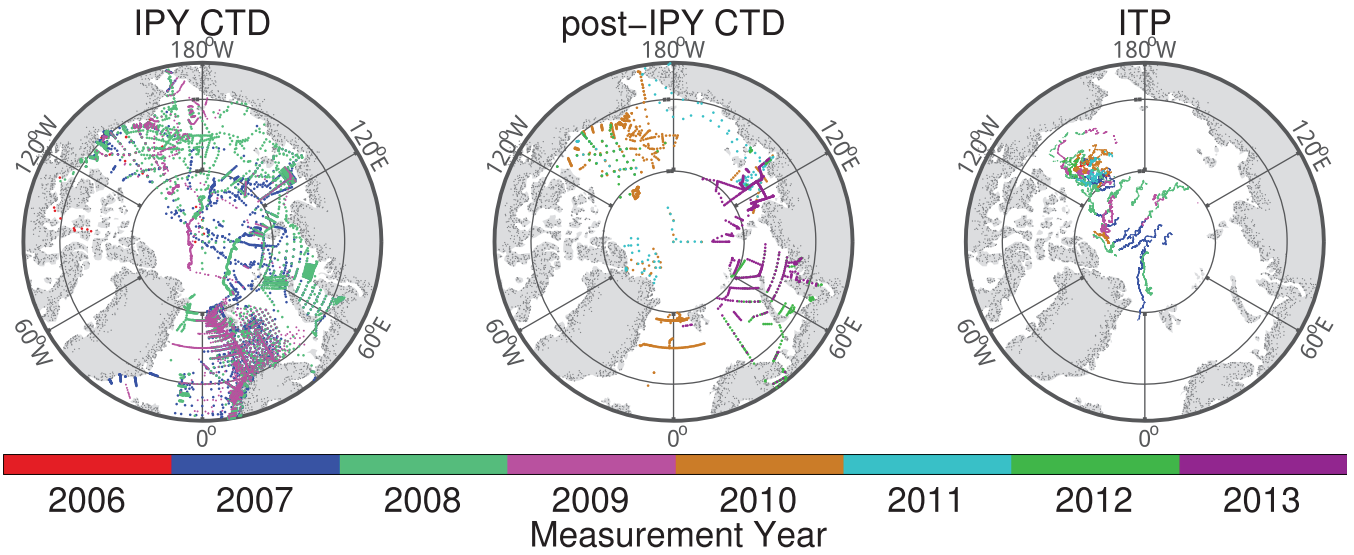

Figure 1. Locations of in situ data poleward of 65°N. Observational data referred to, from left to right, as "IPY CTD," "post-IPY CTD," and "ITP," respectively. Colors correspond to measurement year. Plotted points indicate measurements between 3 m and 7 m depth.

respectively, at locations illustrated in Figure 1. Use of "CTD" hereafter refers exclusively to the ship CTD and xCTD records collected in the IPY CTD and post-IPY CTD observation groups unless specified otherwise.

3. Methodology

Calculation of foundational SST/S from CTD and ITP profiles requires careful consideration. This study uses the measurement average over the range of 3–7 m to represent foundational SST/S for each profile, similar to *Lee et al.* [2010]. Exclusion of the upper 3 m is intended to avoid the most extreme variations in the diurnal temperature cycle and to avoid measurements of questionable origin. For non-ITP records where surface salinity remains above 20 PSU, mean density and temperature profiles show little vertical change over the range 3–7 m (Figure 2). Use of this range for averaging is analytically justified by way of a constrained variance minimization as follows. Use R and T to denote the mean density and temperature profiles, respectively, shown in Figure 2 between depths 1 m and 10 m, and use $\bar{R}(a,b)$ and $\bar{T}(a,b)$ to denote the averages of those profiles over depths a–b m. Ranges which minimize the total variance in R and T may be sought by finding argmin $\left[(R-\bar{R}(a,b))^2+(T-\bar{T}(a,b))^2\right]$ constrained by values such that $a<2$ m and $a<b<10$ m. Among the several near-integer solutions with roughly- equal minimum variances in R and T, the range 3–7 m is the optimal solution with the largest difference

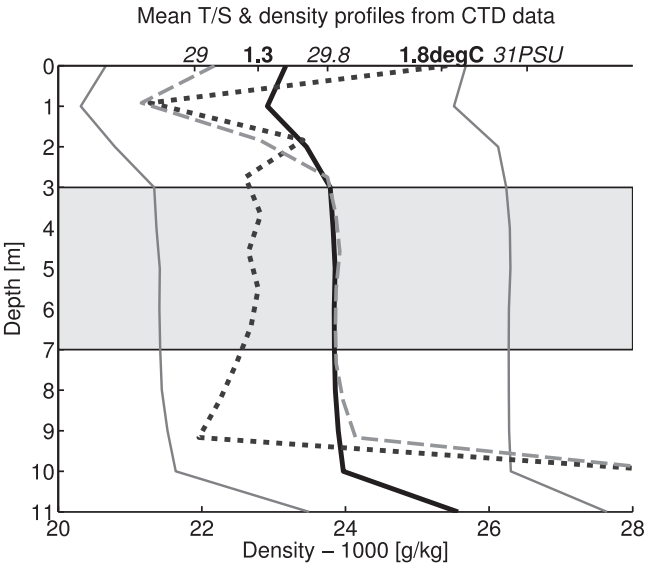

Figure 2. Mean T/S and density profiles for all ship CTD data north of 65°N excluding profiles with salinity less than 20 PSU in the uppermost 10 m. The mean density profile is shown in heavy black as the difference from 1000 g kg^{-1}, and thin grey lines at a distance of one standard deviation illustrate density variability. Mean T/S profiles are shown in dotted and broken lines, respectively, with scales given on the top axis in bold and italics, respectively. The shaded rectangle indicates the 3–7 m region where T/S and density are nearly constant, indicating a region representative of sea-surface foundational values.

between endpoints. However, the mean salinity profile has the lowest variance over the range of 1–5 m as calculated by a similar process. ITP microCATs record T/S data with a minimum depth typically between 5 m and 6 m, and ITP SST/S representatives are therefore calculated as an average of 1–3 values measured by profiler data binned at 7 m and shallower microCAT data.

Gridded SST/S product data was linearly interpolated in space and time to obtain representative product values collocated with each observation. Validation of data products requires comparison against independent data not used in their generation. The assembled IPY database contents were held privately for data contributors and only recently made publicly available, and roughly one-half of the post-IPY CTD data was obtained from closed sources. Most of the data therefore could not have been used for that purpose as it was not available for real-time production. A small portion of older data included in this study may have been used for quality control purposes by e.g., GHRSST in producing low-level background fields on which several analyses are based.

The statistical correspondence between measurements and products quantifies how well the product represents the observed ocean. In situ observations are known with accuracy and certainty, so the quality of a product or model can be roughly ascertained from the fit to the observations. Correlation, product-minus-observation residual bias, and centered root mean squared difference (RMSD) between products and observations data are able to assess how well a product tracks the background state and captures the temporal and spatial variability of the ocean. Anomaly correlation (and associated potential skill) between products and observations is not used; the necessary selection of background climatologies for this metric is likely to favor certain products over others.

4. Results

Direct comparison of representative observational data with gridded products permits estimation of product reliability with an operating assumption that 3–7 m averaging of in situ observations targets the same SST/S intended by the gridded data. Figures 3 and 4 illustrate the quality of fit between the included SST analyses and SST/S models, respectively, and in situ observations. The spatial distributions of these product-minus-observation residuals are charted geographically in Figures 5 and 6, with residual statistics appearing in Tables 1 and 2. Negative values in this table may be interpreted as products being cold relative to observations and the true ocean state. Figure 7 shows comparison of products to CTD and ITP data groups in the form of Taylor diagrams [*Taylor*, 2001], which resolve second-order statistics and should be considered in conjunction with the biases presented in Tables 1 and 2. Use of Taylor diagrams is intended to provide a convenient visual comparison and rough ordering of products in relation to how well they match the variability of observations.

4.1. Comparison Against Ship-Based CTD Data Sets

Satellite analysis sea-surface temperatures compare favorably with CTD measurements, and tend to have small negative SST biases. K10_SST and RSS_SST analyses have the smallest biases and RMSDs relative to the IPY data set. During the post-IPY period, the smallest magnitude biases of -0.02°C are found for RSS_SST and MURSST products with the SST of the latter showing slightly more precision. For IPY and post-IPY data sets respectively, GAMSSA SST has the largest overall biases of -0.27°C and -0.93°C. The GAMSSA SST cold bias in the AO, previously identified in *Dash et al*. [2012], is exaggerated against post-IPY CTD due to poor representation of SST along the coastal Chukchi and Laptev Seas during the warm and low-ice year 2012. All other included satellite-derived SST analyses have biases and errors against CTD measurements comparable to those found against different in situ measurements in previous studies [e.g., *Donlon et al*., 2012; *Dash et al*., 2012; *Høyer et al*., 2012; *Martin et al*., 2012].

Most gridded analyses correlate strongly with CTD observations as visible in the right two-thirds of Table 1 and left columns of Figure 3. With correlations (as squared Pearson coefficients) universally above 80%, SST analyses track the variability of the IPY CTD set well. Post-IPY CTD correlations are lower in general, with the GAMSSA product showing a low correlation of 59% due to large errors during 2012. OSTIA and MURSST products continue matching 83% and 87%, respectively, of the total CTD variance.

The satellite analyses are similar compared to all CTD measurements, but Taylor diagram in the left plot of Figure 7 reveals a rough ranking of all products considered on the basis of second-order statistics. The

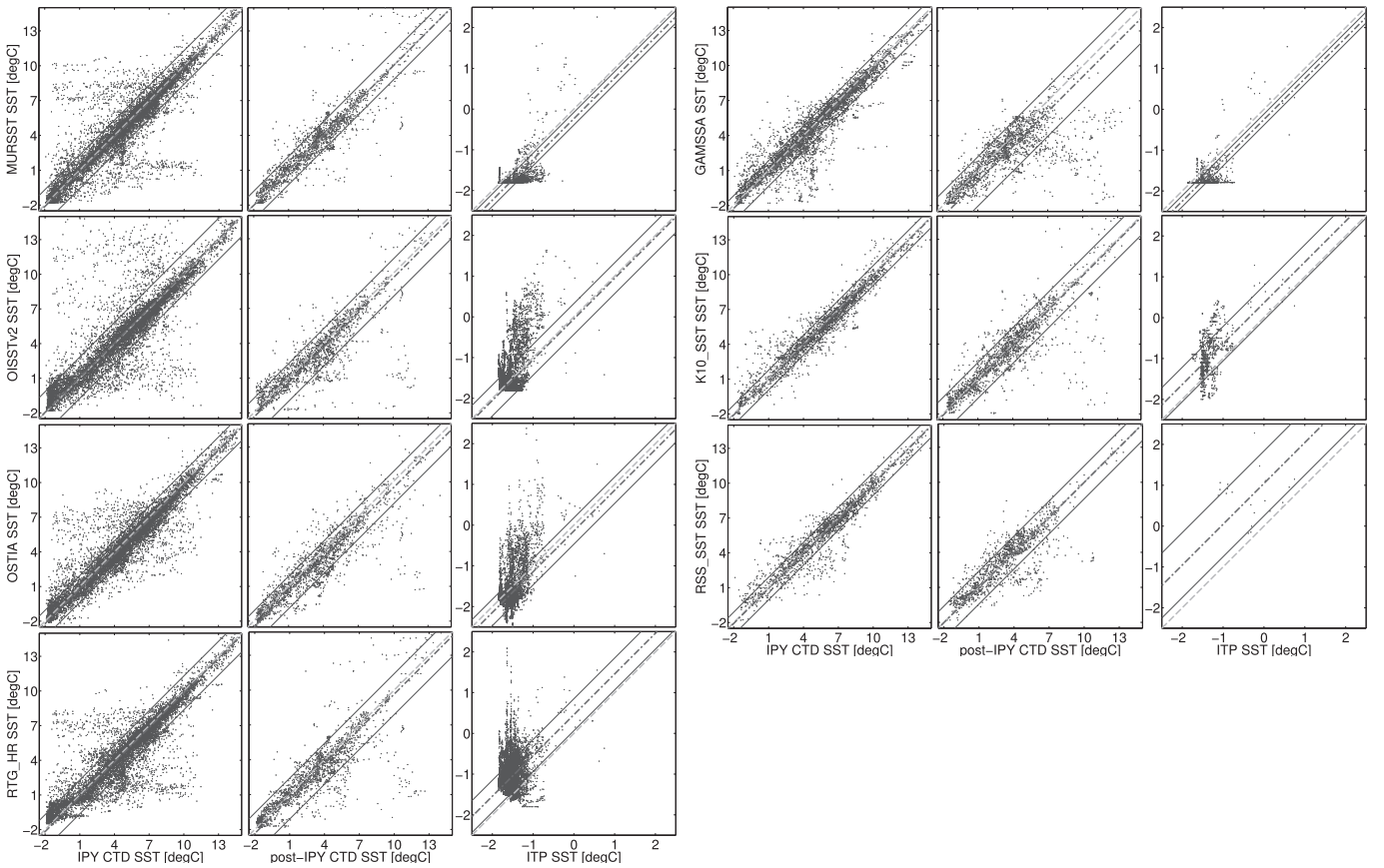

Figure 3. Correlation of SST analyses with respect to (left columns) IPY CTD, (center columns) post-IPY CTD, and (right columns) ITP observations. The descending orders of products are: MURSST, OISSTv2, OSTIA, and RTG_HR on the left side; and GAMSSA, K10_SST, and RSS_SST on the right. In each plot, the thick light-grey dashed line indicates perfect matching; the thin dark-grey line shows this ideal relation offset by product bias; and the thin solid lines bound the bias-offset fit at a distance of the product-minus-observation RMSD.

analyses demonstrate remarkable fit to the variability of CTD SST, with standard errors (as measured by RMSD) within 6% of the true CTD SST standard deviation. The OISSTv2 analysis, despite having the lowest magnitude bias of −0.02 °C, has a larger RMSD and consequently lower overall correlation than other analyses. MURSST analysis appears to be the strongest overall analysis that gives full Arctic coverage without masking sea-ice although OSTIA SST has a higher correlation and lower RMSD in total.

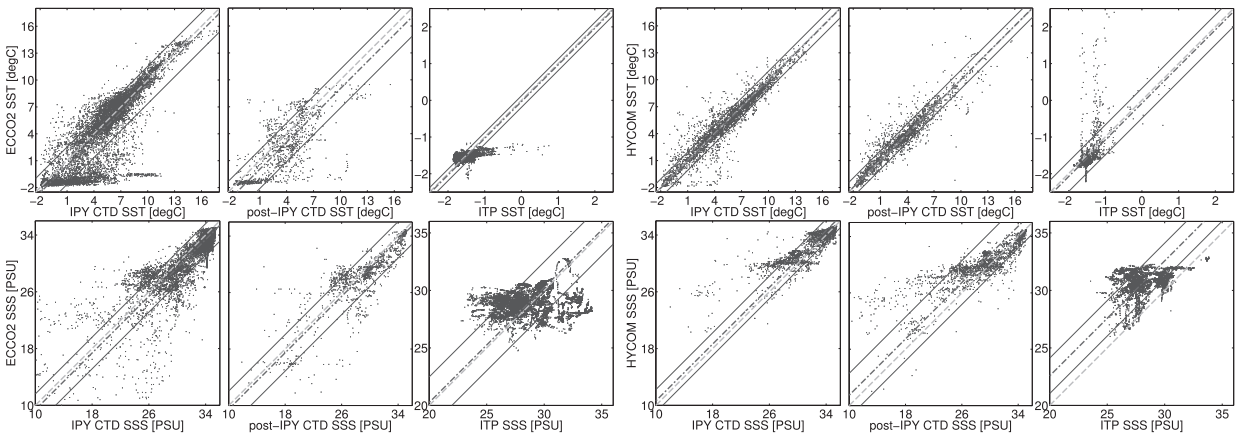

Figure 4. Correlation of SST/S model analyses with respect to (left columns) IPY CTD, (center columns) post-IPY CTD, and (right columns) ITP observations. ECCO2 is shown in the left portion and HYCOM in the right. SST comparison illustrated on the top row and SSS on the bottom. The plotted lines are equivalent to those of Figure 3.

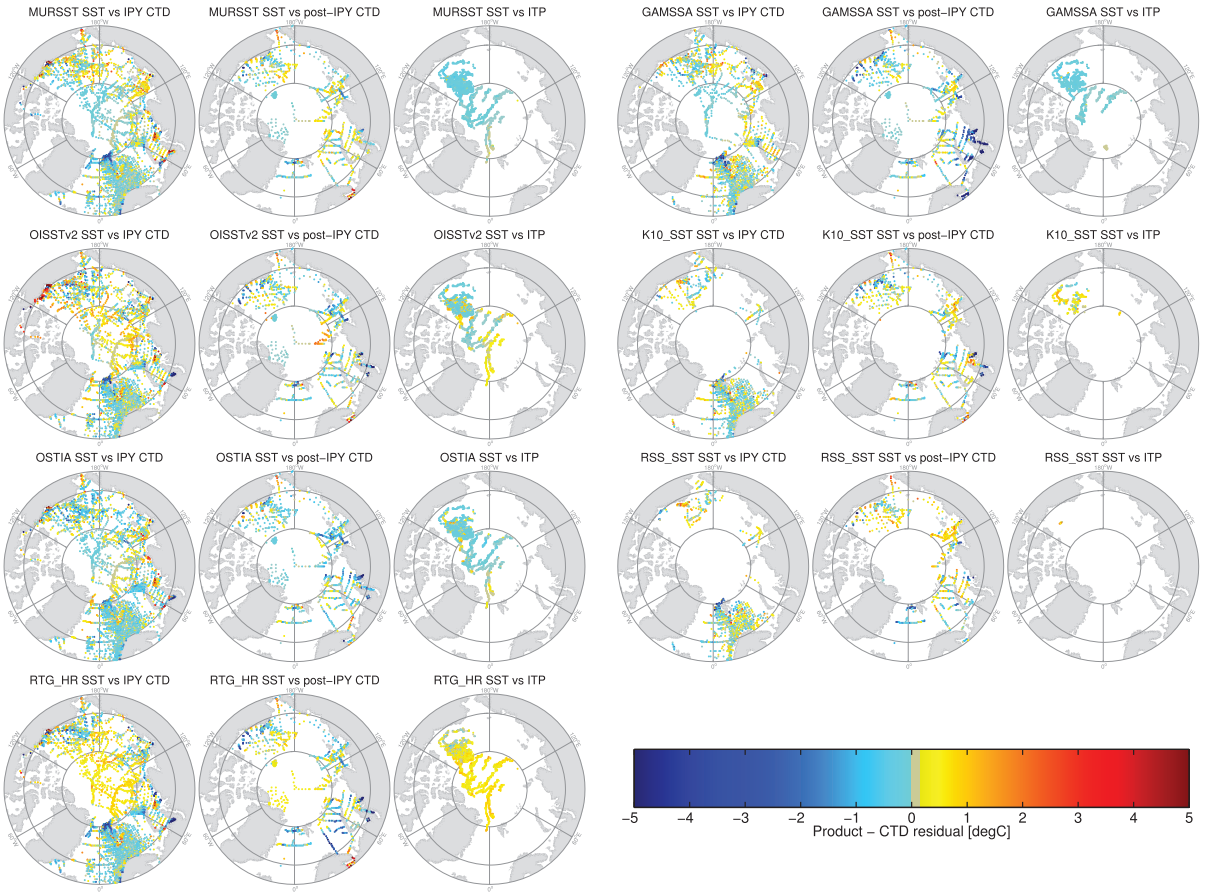

Figure 5. Analysis-minus-observations SST residuals. Plot layout is the same as in Figure 3. Plotted points represent residuals which are bin-averaged on the AVHRR northern hemisphere 25 km equal-area EASE grid [*Brodzik and Knowles*, 2002] to improve visibility and do not coincide with locations in Figure 1. The residual values are given by color. Greyish colors indicate residuals smaller than ±0.15°C.

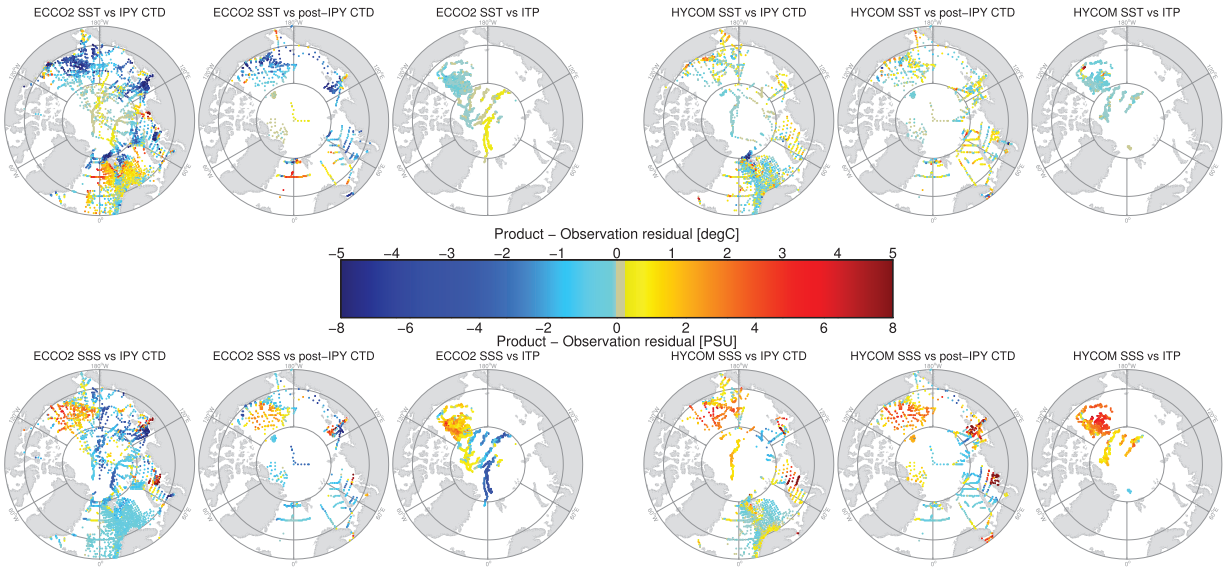

Figure 6. Model-minus-observation SST/S residuals. Plot layout is the same as in Figure 4, and points are plotted in the same manner as those in Figure 5. Greyish colors indicate residuals smaller than ±0.15°C for SST and ±0.16 PSU for SSS.

Table 1. Product Bias, RMSD, and Correlation for Ship CTD Data Groups[a]

	Bias			RMSD			Correlation		
	IPY	p-IPY	All CTD	IPY	p-IPY	All CTD	IPY	p-IPY	All CTD
OISSTv2	+0.05	−0.38	−0.02	1.78	1.62	1.77	0.80	0.76	0.80
OSTIA	−0.22	−0.54	−0.26	1.16	1.40	1.20	0.91	0.82	0.90
RTG_HR	−0.17	−0.35	−0.20	1.47	1.76	1.51	0.85	0.73	0.84
MURSST	−0.14	−0.02	−0.12	1.45	1.22	1.42	0.86	0.87	0.86
GAMSSA	−0.27	−0.93	−0.46	1.09	2.12	1.49	0.92	0.59	0.84
K10_SST	−0.08	−0.13	−0.10	0.86	1.54	1.18	0.93	0.76	0.87
RSS_SST	−0.07	−0.02	−0.05	0.97	1.41	1.18	0.90	0.70	0.86
ECCO2	−0.56	−0.97	−0.61	2.09	2.16	2.10	0.77	0.61	0.76
SSS	−0.73	−0.58	−0.71	2.12	2.51	2.17	0.73	0.69	0.74
HYCOM	−0.07	+0.03	−0.03	1.02	1.03	1.03	0.93	0.90	0.93
SSS	+0.64	+1.27	0.87	1.67	2.79	2.17	0.79	0.72	0.79

[a]The top portion of the table shows SST analyses with bias and RMSD units in °C. The bottom portion corresponds to model SST/S with units of °C and PSU, respectively, for bias and RMSD. The rightmost portion of the table gives product-observation correlations as squared Pearson coefficients.

Models show larger biases and error than analyses products, but appear to have statistics comparable with those of the satellite analyses. HYCOM SST has a +0.03 °C bias against post-IPY CTD-derived SST compared to the −0.97 °C bias of ECCO2 SST, and the corresponding RMSD for HYCOM, 1.03°C, is less than half that of ECCO2 as in Table 1. Figure 7 shows that HYCOM SST has a higher correlation, lower RMSD, and closer fit to the observed SST variability amplitude than ECCO2 for non-ITP data. In these metrics, the HYCOM+NCODA SST fits CTD SST than the included ECCO2 model does.

The bottom of Table 1 includes model-minus-observation SSS residual statistics in the rows labeled "SSS" below each model name, and Figures 4 and 6 include visualizations of SSS residuals. The ECCO2 and HYCOM SSSs differ in sign of bias; ECCO2 under-represents post-IPY CTD-derived SSS by an average of −0.58 PSU compared the + 1.27 PSU bias of HYCOM, and the RMSD of ECCO2 is roughly 10% lower than that of HYCOM. The difference in bias between these model-derived data consistently exceeds 1 PSU across all observation groups, but both model SSSs correlate well with the CTD data (Figure 4, right) with increasing error in model values as observed salinity decreases.

4.1.1. Regional Patterns in Residuals

Spatial distributions of product SST errors show similarities; see Figure 5. Product SSTs are overall slightly cool toward the north Atlantic Ocean with a tendency be warm relative to near-surface CTD toward the Pacific Ocean. In the central AO, there is less agreement: MURSST and GAMSSA SSTs remain cooler than observations while OISSTv2 and RTG_HR analyses are typically warmer. Products also share similar regions where analysis SSTs differ greatly from CTD SST such as the Kara and Beaufort Seas, and the Fram and Bering Straits. Figure 6 shows the geographic distribution of SST/S residuals for the two models. Regions of interest in the ECCO2 model are the Greenland and Barents Seas where SST is overestimated in contrast to all other products considered, and the Pacific sector of the Arctic Ocean where the model notably

Table 2. Product Bias, RMSD, and Correlation for Shallow ITP Data[a]

	Summer			Winter			Annual		
	Bias	RMSD	ρ^2	Bias	RMSD	ρ^2	Bias	RMSD	ρ^2
OISSTv2	+0.24	0.51	0.42	−0.16	0.23	0.01	−0.06	0.38	0.19
OSTIA	+0.14	0.49	0.35	−0.24	0.14	0.02	−0.14	0.32	0.23
RTG_HR	+0.48	0.43	0.04	+0.45	0.38	0.00	+0.46	0.39	0.01
MURSST	−0.24	0.23	0.30	−0.24	0.11	0.10	−0.24	0.15	0.23
GAMSSA	−0.29	0.22	0.30	−0.25	0.10	0.18	−0.25	0.13	0.26
ECCO2	−0.02	0.22	0.23	−0.04	0.12	0.01	−0.04	0.16	0.10
SSS	−0.60	2.60	0.04	+0.55	2.00	0.04	+0.25	2.23	0.03
HYCOM	+0.09	0.76	0.31	−0.12	0.26	0.01	−0.08	0.41	0.19
SSS	+1.95	1.35	0.20	+2.71	1.55	0.06	+2.57	1.54	0.07

[a]Summer, winter, and total values are given on the left, center, and right thirds of the table, respectively. Bias and RMSD of product-minus-observation residuals, and correlation presented as squared Pearson coefficient, are given in the respective left, center, and right columns of each third. Top rows correspond to SST analyses, and the bottom to SST/S models. K10_SST and RSS_SST are under-sampled and not included. Summer is taken as Julian days 130–260 when ITP SST/S depart from freezing, and winter as the remainder of the year.

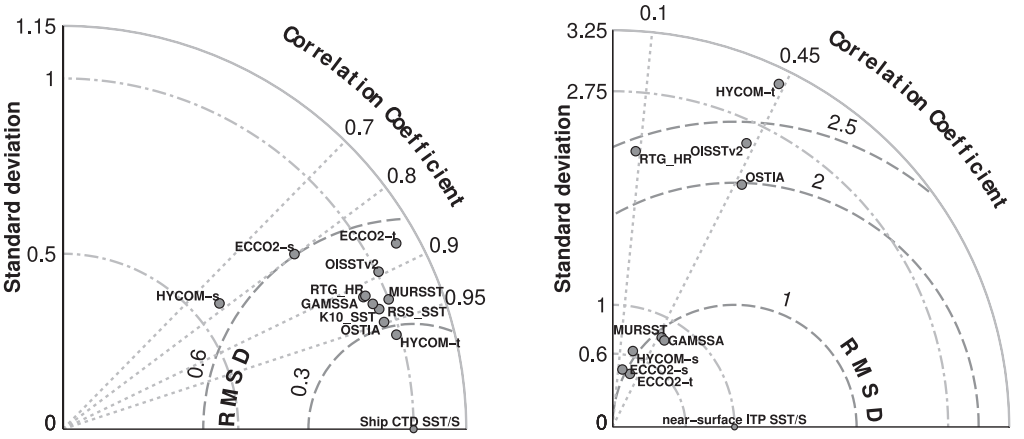

Figure 7. Comparison of nondimensional second-order statistics via Taylor diagram against all (left) ship CTD observations and (right) near-surface ITP observations. The relation shown here is independent of bias (first-order difference) and one must consider these in conjunction with Table 1.

represents the ocean water as too cold and salty both during and after IPY. For commonly resolved CTD data, ECCO2 and HYCOM products appear very salty in the Beaufort and Chukchi Seas. Figure 7 shows that against all CTD data, HYCOM SST has a lower overall RMSD and stronger correlation of 93% than even the satellite analyses, but the model SSS has a larger (and positive) bias and does not track the variability found in CTDs as well as ECCO2 does.

For SST/S models, correlations with in situ data are weaker than for SST analyses, and the accuracy of model SST/S values is not constant throughout the AO. HYCOM SST residuals are more uniform across the AO shelves with no discernible change in bias between the Eurasian and Amerasian shelves; the spatial distribution of residuals resembles those of the satellite analyses. ECCO2 SST shows a bias of -1.72°C for all CTD measurements along the the Siberian shelves between E60° and E180° while a $+0.25^{\circ}$C bias is present for CTD measurements in the Norwegian and Barents Seas region.

SSSs in these regions are more accurate in the ECCO2 model, which also lacks the overly salty Norwegian coastal waters found in HYCOM. The warm, salty surface errors found in the HYCOM representation of the Kara sea contrasts the cool, fresh bias found in ECCO2 throughout the Siberian Shelf. In the central AO, the HYCOM product is cool and salty relative to IPY CTD data while ECCO2 is slightly warm and fresh. Similar patterns are found compared to the longer time series of ITP measurements. Both model data sets appear cool and strongly salty in the Beaufort Sea region despite large differences in AO-wide SSS biases. MURSST SST also has biases of $+0.35^{\circ}$C and -0.30°C for these regions, respectively. Similar but smaller differences between errors in these regions are found in GAMSSA and OISSTv2 products. No other large-area differences in regional biases are detected in residuals relative to CTD data but clusters of large errors in several smaller data-abundant regions such as the Kara Sea and northern Svalbard coast beg further investigation not pursued here. In comparison to SSS observations which both models resolve, HYCOM and ECCO2 represent the Beaufort and Chukchi regions with strong positive salinity biases, but this region is more localized in ECCO2 whereas the overly saline region extends into Bering Sea and the central AO in HYCOM.

Model SST/S and analysis SST are less accurate in regions of runoff influence, such as in the Kara Sea and near outflows of the Lena and Mackenzie Rivers. In these areas, both models over-estimate salinity and many products under-estimate SST. The regional presence of these biases and increased inaccuracy may be due to the choice of averaging CTD data over 3–7 m; the mean profile over all CTD measurements with salinity less than 20 PSU in 0–3 m depths has a monotonic decrease in temperature between 3 m and 10 m. This suggests that calculated SST cold residuals in freshwater regions may be reduced if deeper CTD measurements were averaged to represent observations.

4.2. Comparison Against ITP Near-Surface Data

ITPs move with the ice and their records form a time-series, so associated residuals are correlated in time and are substantially less independent than the CTD observations. Further, the errors are distributed nonnormally

due to the constraint on water temperature minimum. Checking products against ITP data may not be as meaningful as it is for CTD data, but this comparison is diagnostic in its own right. Against ITP data in the 3–7 m range, the SST analyses display more pronounced differences. The K10_SST and RSS_SST analyses sample only about 1200 and 20, respectively, of the approximately 14,000 ITP profiles due to sea-ice masking in those products, and neither analysis is further included in this discussion of fit to ITP-derived SST.

RTG_HR SST shows a distinctive strong positive bias of 0.45°C and less than 1% correlation with ITP data, and this bias is evident against CTD data sets as well in Figure 5. OISSTv2 SST has the lowest bias for ITP data at −0.05°C, but, along with OSTIA and RTG_HR, over-represents the inherent variance of ITP SST by more than 2.5 times (Figure 7, right). MURSST and GAMSSA statistics are nearly identical with biases near −0.25°C and RMSDs roughly 0.14°C.

Among the analyses considered, MURSST and OSTIA SST appear equal in providing both strongest correlation and lowest biases to ITP-derived SST for products which include a separate sea-ice analysis. The salinities in the ECCO2 model have smaller RMSD and bias than those in HYCOM for all data sets including ITP, although they still under-represent the variability of observations. Both models are competitive with satellite analyses in terms bias and RMSD for ITP data. Compared to the ECCO2 product, HYCOM SST has larger RMSD and a more pronounced over-estimation of ITP temperature variability, although the SSS residual second-order statistics are similar for both models. ECCO2 SST exists in a temperature range much smaller than the data while HYCOM SST is opposite (Figure 4, bottom).

4.3. Improving Fit to ITP Near-Surface Data

Hydrographical data beneath sea-ice for sub surface layers are scarce and none of it is remotely sensed. Instead, MW satellite bands are able to resolve approximate sea-ice concentration (SIC), and product SST can be relaxed to a freezing-point or other climatological state as SIC approaches 100%. This parametric approach is used by many of satellite analyses with a target freezing-point near −1.8°C implicitly assuming salinities near 35 PSU [*Rayner et al*., 2003; *Donlon et al*., 2012]. Roughly 75% of ITP temperature data in the 3–7 m range used for comparison is within 0.01°C of the freezing-point calculated from the observed salinity.

The difference between ITP temperature and freezing temperature calculated from ITP salinity shows a change during Julian days 130–260, when the mean of this difference increases in order from 1/1000°C to 1/10°C with a corresponding change in standard deviation from 0.027°C in winter to 0.138°C in summer. Salinity distribution of ITP measurements is multimodal and shows both seasonal and inter-annual variability, reflecting the changes in freshwater distribution and rapid freshening. The upper-most decile of near-surface averaged ITP salinities never exceed 30.5 PSU for any season in the years 2008–2012.

Different biases of several products are detected in seasonally partitioned residual statistics presented in the left and center columns of Table 2. No seasonal changes in bias appear to be present in RTG_HR, MURSST, GAMSSA, or ECCO2 SSTs. However, biases with seasonal variation are apparent for OISSTv2, OSTIA, and HYCOM residuals against ITP, and differ from the total biases presented in Table 2. These products show warm biases of +0.24°C, +0.14°C, and +0.09°C, respectively, in summer contrasting cold biases of −0.16°C, −0.24°C, and −0.12°C, respectively, in winter. The winter biases against ITP data are roughly in the same range as those against the summer-focused CTD data. This suggests that rather than a cold bias in the winter, products are simply warmly biased in the central AO during the summer, and this bias is undetected by CTD measurements concentrated near and on continental shelves. Large errors for ITP data are localized in the Beaufort Gyre where surface freshwater content depend on a complex interaction of terrestrial hydrological cycles, atmospheric modes, and sea-ice melt [*Timmermans et al*., 2011; *Morison et al*., 2012]. Inaccuracies in that region may reflect the rapid salinity divergence from previous climatologies in the Beaufort Gyre and Canada Basin where ITPs provide the most data.

An increased warm bias during summer for certain products may be due to under-relaxation of temperatures toward freezing when the presence of surface melt-ponds causes SIC under-estimation. Cold bias in winter for these products is likely the result of prescribed climatological freezing temperatures: freezing-point prescribed near −1.8°C is significantly cooler than temperatures observed by ITPs. The under-estimation of sea-ice in summer and over-estimation of background salinity is consistent with the seasonally partitioned correlation diagrams shown in Figure 8. In the top row of this figure, summer-time measurements cluster near −1.6°C while analysis SSTs are frequently warmer by 1/2°C. The bottom row illustrates strongly constrained

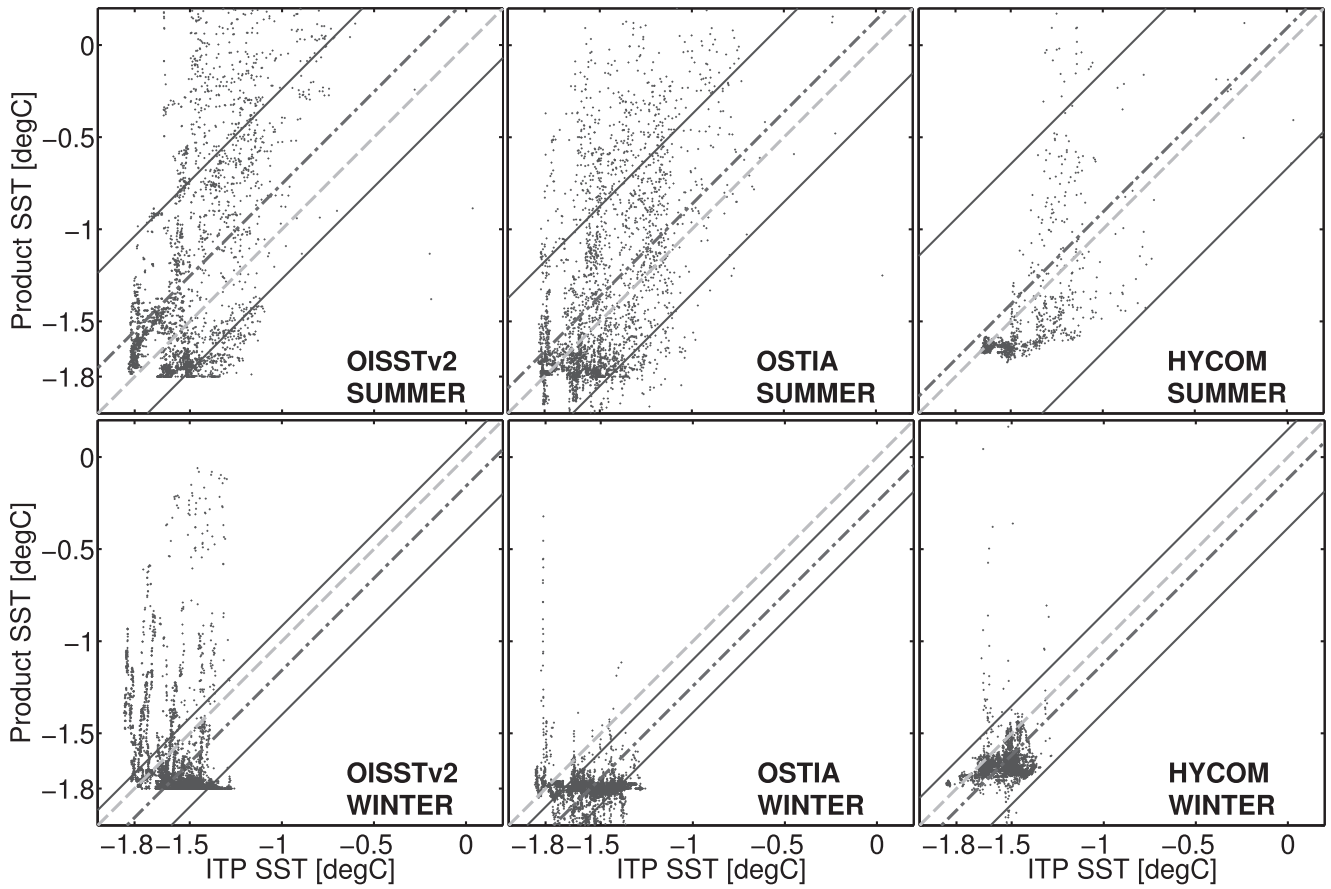

Figure 8. Seasonal comparison of select products to ITP measurements. The (left) OISSTv2, (middle) OSTIA, and (right) HYCOM products are biased (top) positively in summer and (bottom) negatively in winter. Summer-time residual distribution shows a less constrained temperature range compared to observed values. Winter residuals show strongly constrained product SST targeting −1.8°C while actual measurements are centered near −1.6°C.

temperatures near −1.8°C while the observed near-freezing temperatures are warmer by roughly 1/6°C. Another possible source of summer warm bias in these products is the under-estimation of wind stress used in transforming the satellite skin temperatures to product SST, but this is not explored here.

Product SST P may be corrected to better fit in situ ITP data by using product-minus-observation residuals ΔP and SIC I. In this discussion, I^* is a critical SIC above which ice influences SST. Linear regression of residuals against SIC greater than I^* parametrizes residuals in terms of SIC: $\Delta P = \alpha \cdot I + \beta$ where α, β are statistically determined coefficients. The corrected product values $P_1^* = P + (\alpha \cdot I + \beta) \cdot \{I^* \leq I \leq 1\}$ target the mean full-ice SST observed by ITP and have an improved mean linear relation to ITP near-surface temperature.

Corrected SST P_1^* have biases smaller than 10^{-2} °C against ITP data, but most ITP data are associated with SIC greater than 95%. As a result, the residuals of P_1^* show little improvement in higher-order statistics and the values of (α, β) fluctuate for random sub-samples of ITP data. However, the full-ice SST corrections $\gamma = \alpha + \beta$ are robust and correspond to freezing temperature corrections. Table 3 shows (α, γ) -values for several products using critical SIC of 50%. The corrected freezing temperatures $P + \gamma$ are approximately −1.62 °C for all products presented.

As a result of inadequate variation of SIC among ITP measurements, the method presented above tends to correct only product bias. The frequency of ITP data with respect to SIC resembles log-normal distribution so SST correction can be improved using a normalized logarithm of SIC. Transforming SIC into $\tilde{I} = \log(2-I)/\log(2-I^*)$ redistributes the dense cluster of ITP data associated with high SIC more appropriately onto the interval $0 \leq \tilde{I} \leq 1$. This decreases the weight of ITP associated with high SIC and yields a better conditioned least-squares problem for linear regression.

Table 3. Least-Squares Slope (α) and Temperature Correction at 100% SIC (γ) for Linear Correction to Various Products for SIC Above 50%[a]

	Summer		Winter		Annual	
	α (°C)	γ (°C)	α (°C/%)	γ (°C)	α (°C/%)	γ (°C)
OISSTv2	1.61	−0.15	0.71	0.20	1.11	0.11
OSTIA	0.29	−0.12	0.16	0.24	0.29	0.15
MURSST	0.11	0.24	0.15	0.25	0.14	0.24
ECCO2	0.12	0.03	0.16	0.05	0.15	0.04
HYCOM	−1.26	0.12	0.26	0.10	0.26	0.10

[a]Original product values, P, better reflect ITP observations by computing $P_1^* = P + \gamma + \alpha \cdot (I - 1)$ for $I > I^* = 0.5$.

Modified product SST values $P_2^* = P + (\delta \cdot \tilde{I} + \eta) \cdot \{0 \leq \tilde{I} \leq 1\}$ fit the ITP data in the mean more uniformly over the region $I \geq I^*$. Table 4 shows coefficients (δ, η) for several products whose modified SST absolute biases are less than $5 \cdot 10^{-3}$ °C. This method produces values of (δ, η) which are considerably more robust for random sub-samples of ITP data, and may also improve higher-order statistical correspondence between product and ITP values. For example, applying independently the summer and winter corrections reduces the RMSD of OISSTv2 and OSTIA residuals by roughly 16% and 14.5%, respectively, and results in corrected SST P_2^* with biases smaller than those of P_1^*.

5. Summary

We extracted near-surface AO temperature and salinity profiles from ship-based CTD rosette casts, xCTD, and ITP data sources during the years 2006–2013. Each profile containing valid T/S in the range 3–7 m was reduced to the mean over that range to represent observed foundational SST and SSS. These in situ SST/S observations were compared to corresponding values in OISSTv2, OSTIA, MURSST, RTG_HR, K10_SST, and RSS_SST gridded satellite analyses, as well as particular NASA/JPL ECCO2 Cube92 and HYCOM+NCODA data-assimilated model outputs. Against CTD temperature measurements, the satellite products had similar error statistics; correlation ranged between approximately 80% (OISSTv2) and 90% (OSTIA) with relative errors in variability ranging from −6% (GAMSSA and RTG_HR) to + 1% (OISSTv2). All analyses showed overall cold biases against averaged ship-CTD profiles, and bias magnitudes were less than 0.5 °C. The warm post-IPY year 2012 generated large errors for many products; exclusion of 2012 CTD profiles reduces bias magnitude to less than 0.3 °C for all products.

Bias statistics against CTD SST data for models were generally worse than those for conditioned satellite products, but models estimate SSS as well as other key quantities (e.g., heat fluxes) not considered. The reader is reminded that satellite MW salinity measurement is a more recent technological advancement [*Lagerloef and Font*, 2010], and is not listed in either model documentation as a data source. The HYCOM product SST showed smaller bias (−0.04 °C) and higher correlation (93%) with ship-based CTD data than any of the satellite analyses but showed a large salinity bias of + 1.23 PSU. ECCO2 product SST/S biases were −0.61°C and −0.42 PSU, respectively, for all ship-based CTD data. ECCO2 matched the variance of the observations with + 9% relative temperature error and −17% salinity error while corresponding values for HYCOM were −1% and −43%, respectively. Both models showed overall negative temperature biases similar to satellite analyses. Larger errors in T/S were detected in regions of direct freshwater and sea-ice

Table 4. Least-Squares Slope (δ) and Intercept (η) for Residuals Regressed Over SIC Transformed by $\tilde{I} = \log(2 - I)/\log(3/2)$ Corresponding to a Critical SIC Value of 50%[a]

	Summer		Winter		Annual	
	δ (°C)	η (°C)	δ (°C)	η (°C)	δ (°C)	η (°C)
OISSTv2	−0.76	−0.14	−0.34	+0.20	−0.52	+0.12
OSTIA	−0.13	−0.12	−0.08	+0.24	−0.14	+0.15
MURSST	−0.05	+0.24	−0.07	+0.25	−0.07	+0.25
ECCO2	−0.06	+0.03	−0.08	+0.05	−0.07	+0.05
HYCOM	+0.54	+0.12	−0.13	+0.10	−0.13	+0.11

[a]Original product values, P, better reflect ITP observations by computing $P_2^* = P + (\delta \cdot \tilde{I} + \eta)$ for $I > I^* = 0.5$.

influence. Generally, ECCO2 Cube92 appears to have a cold and fresh bias in the AO and adjunct shelves while HYCOM+NCODA output is cold and salty in those regions.

Near-surface ITP data reflect year-round T/S below multiyear sea-ice typically found over deep ocean basins in the Central Arctic. Away from the more often surveyed coastal shelves, products showed much lower correlation with ITP temperature measurements despite having small magnitude biases; the strongest nonseasonal correlation (26%) was found for GAMSSA, a product with negative bias of $-0.25\,^{\circ}$C. Product RMSDs were correspondingly very small due to the narrow range temperature found in ITP-derived data. Considering ITP data during Julian summer days 130–260 separately from winter days, seasonal changes in temperature bias appear in all products except MURSST and ECCO2, with summer temperature biases higher than those found during winter. In OISSTv2, OSTIA, and HYCOM SSTs, these biases are positive in summer and negative in winter. In summer, this is likely a consequence of SIC under-estimation due to melt ponds.

A likely diagnosis for winter-time cold bias is over-estimation of surface salinities in background climatologies. Many products appear to use freezing-point temperatures lower than observed ITP temperatures; during each year of 2009–2012, less than 10% of ITP near-surface salinities in the Canada Basin and Beaufort Sea are above 30.5 PSU while products target freezing temperatures closer to 34 or 35 PSU.

Linear correction of SST on the basis of SIC and the product-minus-ITP temperature allows for reducing product bias, but has difficulty improving second-order statistics due to the relatively low variability in SIC associated with ITP near-surface samples. A suggested linear correction of product SST against logarithmically transformed SIC may be used to de-bias products as well as reduce RMSD in this case.

Ship-collected CTD profiles and near-surface ITP data as collected are not intended to represent sea surface quantities. ITPs sample waters below sea-ice where the satellite observations are not present. Satellite-derived products and models which rely heavily on satellite SST present SST/S quantities where there are no remotely sensed data, necessitating the use of available in situ data for validation and quality diagnosis. Improvements to product quality may be made, not only through direct use of these in situ data sources for analysis and state estimation, but additionally by the incorporation of these data to update background fields and to improve the representativeness of remotely sensible proxy data such as SIC.

Acknowledgments

This study was funded by NOAA grant NA11OAR4310153 for the project "Arctic Ocean Structure during the Period of the International Polar Year" and by NSF grants 1107925 and 1203740. We further thank the support of the Russia government (p220/mega-grant 2013-220-04-157) and the Russian Scientific Foundation (grant no. 15-17-20032). Organizations contributing data to the IPY AO database are owed a debt of thanks. Other data made available by and collected from: the Beaufort Gyre Exploration Program based at the Woods Hole Oceanographic Institution (http://www.whoi.edu/beaufortgyre) in collaboration with researchers from Fisheries and Oceans Canada at the Institute of Ocean Sciences; the latter three coauthors; and the Nansen and Amundsen Basins Observational System (NABOS) program.

References

Brodzik, M. J., and K. W. Knowles (2002), EASE-Grid: A versatile set of equal-area projections and grids, *Discrete Global Grids*, *5*, 110–125.

Chen, Y., J. A. Francis, and J. R. Miller (2002), Surface temperature of the Arctic: Comparison of TOVS satellite retrievals with surface observations, *J. Clim.*, *15*(24), 3698–3708.

Comiso, J. (2010), *Polar Oceans From Space*, vol. 41, Springer, N. Y.

Corlett, G. K., et al. (2006), The accuracy of SST retrievals from AATSR: An initial assessment through geophysical validation against in situ radiometers, buoys and other SST data sets, *Adv. Space Res.*,*37*(4), 764–769.

Cummings, J. A. (2005), Operational multivariate ocean data assimilation, *Q. J. R. Meteorol. Soc.*, *131*(613), 3583–3604.

Dash, P., et al. (2012), Group for High Resolution Sea Surface Temperature (GHRSST) analysis fields inter-comparisons. Part 2: Near real time web-based level 4 SST Quality Monitor (L4-SQUAM), *Deep Sea Res., Part II*, *77*, 31–43.

Donlon, C. J., M. Martin, J. D. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer (2012), The operational sea surface temperature and sea ice analysis (OSTIA) system, *Remote Sens. Environ.*, *116*, 140–158.

Emery, W., S. Castro, G. Wick, P. Schluessel, and C. Donlon (2001), Estimating sea surface temperature from infrared satellite and in situ temperature data, *Bull. Am. Meteorol. Soc.*, *82*, 2773–2785.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young (1996), Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment, *J. Geophys. Res.*, *101*(C2), 3747–3764.

Gammelsrød, T., Ø. Leikvin, V. Lien, W. P. Budgell, H. Loeng, and W. Maslowski (2009), Mass and heat transports in the {NE} Barents Sea: Observations and models, *J. Mar. Sci.*, *75*(1-2), 56–69.

GCOS (2011), Systematic observation requirements for satellite-based products for climate, 2011 update, *WMO GCOS Rep.*, *154*, 1–127.

Gemmill, W., B. Katz, and X. Li (2007), Daily real-time global sea surface temperature-high resolution analysis at NOAA/NCEP, *NCEP Off. Note*, *260*, 39 pp.

Høyer, J. L., I. Karagali, G. Dybkjær, and R. Tonboe (2012), Multi sensor validation and error characteristics of arctic satellite sea surface temperature observations, *Remote Sens. Environ.*, *121*, 335–346.

Kawai, Y., and A. Wada (2007), Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: A review, *J. Oceanogr.*, *63*(5), 721–744.

Krishfield, R., J. Toole, A. Proshutinsky, and M.-L. Timmermans (2008), Automated ice-tethered profilers for seawater observations under pack ice in all seasons, *J. Atmos. Oceanic Technol.*, *25*(11), 2094–2105.

Lagerloef, G., and J. Font (2010), SMOS and Aquarius/SAC-D missions: The era of spaceborne salinity measurements is about to begin, in edited by V. Barale, J. F. R. Gower, and L. Alberotanza, *Oceanography From Space*: Revisited, pp. 35–58, Springer, N. Y.

Lee, M.-A., M.-T. Tzeng, K. Hosoda, F. Sakaida, H. Kawamura, W.-J. Shieh, Y. Yang, and Y. Chan (2010), Validation of JAXA/MODIS sea surface temperature in water around Taiwan using the Terra and Aqua satellites, *Terr. Atmos. Oceanic Sci.*, *21*(4), 727–736.

Martin, M., et al. (2012), Group for High Resolution Sea Surface temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE), *Deep Sea Res., Part II*, *77*, 21–30.

Menemenlis, D., I. Fukumori, and T. Lee (2005), Using Green's functions to calibrate an ocean general circulation model, *Mon. Weather Rev.*, *133*(5), 1244–1240.

Morison, J., R. Kwok, C. Peralta-Ferriz, M. Alkire, I. Rigor, R. Andersen, and M. Steele (2012), Changing arctic ocean freshwater pathways, *Nature*,*481*(7379), 66–70.

Nguyen, A. T, D. Menemenlis, and R. Kwok (2009), Improved modeling of the Arctic halocline with a subgrid-scale brine rejection parameterization, *J. Geophys. Res.*, *114*, C11014, doi:10.1029/2008JC005121.

Nguyen, A. T., D. Menemenlis, and R. Kwok (2011), Arctic ice-ocean simulation with optimized model parameters: Approach and assessment, *J. Geophys. Res.*, *116*, C04025, doi:10.1029/2010JC006573.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi: 10.1029/2002JD002670,

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax (2007), Daily high-resolution-blended analyses for sea surface temperature, *J. Clim.*, *20*, 5473–5496.

Robinson, I. S. (2004), *Measuring the Oceans From Space: The Principles and Methods of Satellite Oceanography*, 1 ed., Springer, N. Y.

Sakov, A., F. Counillon, L. Bertino, K. A. Lisæter, P. R. Oke, and A. Korablev (2012), TOPAZ4: An ocean-sea ice data assimilation system for the North Atlantic and Arctic, *Ocean Sci.*, *8*(4), 633–656.

Stark, J. D., C. J. Donlon, M. J. Martin, and M. E. McCulloch (2007), OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system, paper presented at Oceans 2007: Marine Challenges: Coastline to Deep Sea, IEEE/Oceanic Eng. Soc., Aberdeen, Scotland, 18–21 June.

Sy, A., and D. Wright (2000), XBT/XCTD standard test procedures for reliability and performance tests of expendable probes at sea, paper presented at 3rd Session of Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology Ship-of-Opportunity Implementation Panel (SOOPIP-III), March 28–31, La Jolla, Calif.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.

Thiébaux, J., E. Rogers, W. Wang, and B. Katz (2003), A new high-resolution blended real-time global sea surface temperature analysis, *Bull. Am. Meteorol. Soc.*, *85*(5), 645–656.

Timmermans, M.-L., A. Proshutinsky, R. A. Krishfield, D. K. Perovich, J. A. Richter-Menge, T. P. Stanton, and J. M. Toole (2011), Surface freshening in the Arctic Ocean's Eurasian Basin: An apparent consequence of recent change in the wind-driven circulation, *J. Geophys. Res.*, *116*, C00D03, doi:10.1029/2011JC006975.

Toole, J. M., R. A. Krishfield, M.-L. Timmermans, and A. Proshutinsky (2011), The ice-tethered profiler: Argo of the arctic, *Oceanography*, *24*(3), 126–135.

Zhong, A. and H. Beggs (2008), Operational implementation of global Australian multi-sensor sea surface temperature analysis, *Analysis and Prediction Operations Bulletin*. No. 77, Bureau of Meteorology, Australia, 2 October. [Available at: http://cawcr.gov.au/projects/SST/GAMSSA_BoM_Operational_Bulletin_77.pdf.]