Ocean Dynamics

North Atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. Part 1: Dataset validation and RMS variability

--Manuscript Draft--

Manuscript Number:	ODYN-D-18-00150R1
Full Title:	North Atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. Part 1: Dataset validation and RMS variability
Article Type:	Topical Collection - Liege 2018
Keywords:	Sea surface temperature; sea surface salinity; North Atlantic; decennial variability
Corresponding Author:	Gilles Reverdin Centre National de la Recherche Scientifique FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Centre National de la Recherche Scientifique
Corresponding Author's Secondary Institution:	
First Author:	Gilles Reverdin
First Author Secondary Information:	
Order of Authors:	Gilles Reverdin
	Andrew Friedman
	Léon Chafik
	Naomy Penny Holliday
	Tanguy Szekely
	Hedinn Valdimarson
	Igor Yashayaev
Order of Authors Secondary Information:	
Funding Information:	CNRS (xxx) — Dr Gilles Reverdin
Abstract:	We present a binned annual product (BINS) of sea surface temperature (SST), salinity (SSS) and water density observations for 1896–2015 of the subpolar North Atlantic between 40°N and 70°N, mostly excluding the shelf areas. The product, which resolves space scales on the order of 200 to 500 km, reproduces most of the pluri-annual variability in different time series covering at least the last three decades or of the along-track ship monitoring. Comparisons with other SSS and SST gridded products available since 1950 suggest that BINS captures the large decadal to multidecadal variability. Comparison with the HadSST3 SST product since 1896 also indicates that the decadal and multidecadal variability is usually well reproduced, with small differences in long-term trends or in areas with marginal data coverage in either of the two products. Outside of the Labrador Sea and Greenland margins, pluri-annual variability is rather similar in different seasons. Variability at periods longer than 15 years is a large part of the total pluri-annual variability, both for SST and SSS, except possibly in the south-western part of the domain. Variability in SST and SSS increases towards the west, with the contribution of salinity variability to density dominating that of temperature in the western Atlantic, except close to the Gulf Stream and North Atlantic Current in the southwest area. Weaker variability and larger relative temperature contributions to density changes are found in the eastern part of the gyre

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	and south of Iceland.
Additional Information:	
Question	**Response**
Is this manuscript (or major parts of it) currently submitted to another journal or has it been submitted before and then either been rejected or withdrawn?	NO

Click here to access/download;Manuscript;SPG.part1.04092018.pdf

Click here to view linked References

North Atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. Part 1: Dataset validation and RMS variability

G. Reverdin[1], A. R. Friedman[2], L. Chafik[3], N. P. Holliday[4], T. Szekely[5], H. Valdimarson[6], I. Yashayaev[7]

[1] Sorbonne-Université, CNRS/IRD/MNHN (LOCEAN), Paris, France

[2] School of Geosciences, University of Edinburgh, UK. ORCID ID https://orcid.org/0000-0001-6994-2037

[3] Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden. ORCID ID https://orcid.org/0000-0002-5538-545X

[4] National Oceanography Centre, Southampton, UK

[5] IUEM, Brest, France

[6] Marine and Freshwater Research Institute, Reykjavik, Iceland

[7] Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS, Canada

Corresponding author: G. Reverdin, Laboratoire d'Océanographie et de climatologie par expérimentation et analyse numérique, Institut Pierre Simon Laplace, Sorbonne-Université, case 100, 4 pl. Jussieu, 75252 Paris Cedex 05, France (gilles.reverdin@locean-ipsl.upmc.fr ; tel : 33-1-44272342 fax : 33-1-44273805 ; orcid.org/0000-0002-5583-8236)

Abstract

We present a binned annual product (BINS) of sea surface temperature (SST), salinity (SSS) and water density observations for 1896–2015 of the subpolar North Atlantic between 40°N and 70°N, mostly excluding the shelf areas. The product, which resolves space scales on the order of 200 to 500 km, reproduces most of the pluri-annual variability in different time series covering at least the last three decades or of the along-track ship monitoring. Comparisons with other SSS and SST gridded products available since 1950 suggest that BINS captures the large decadal to multidecadal variability. Comparison with the HadSST3 SST product since 1896 also indicates that the decadal and multidecadal variability is usually well reproduced, with small differences in long-term trends or in areas with marginal data coverage in either of the two products. Outside of the Labrador Sea and Greenland margins, pluri-annual variability is rather similar in different seasons. Variability at periods longer than 15 years is a large part of the total pluri-annual variability, both for SST and SSS, except possibly in the south-western part of the domain. Variability in SST and SSS increases towards the west, with the contribution of salinity variability to density dominating that of temperature in the western Atlantic, except close to the Gulf Stream and North Atlantic Current in the southwest area. Weaker variability and larger relative temperature contributions to density changes are found in the eastern part of the gyre and south of Iceland.

Keywords: sea surface temperature, sea surface salinity, North Atlantic, decennial variability

1. Introduction

The North Atlantic subpolar gyre (SPG) receives salty surface water from the subtropical North Atlantic as well as fresh water originating from the Arctic and the Nordic seas and circulating mostly on and along the shelves. It is a formation site of deep and intermediate waters and thereby a major contributor to the lower limb of the Atlantic meridional overturning circulation (AMOC) (Rhein et al. 2011). There is wintertime densification of the surface Labrador and Irminger Seas resulting in intermittent deep convection reaching and exceeding 2400 m in the mid-1990s and 2000 m in recent years (Yashayaev and Loder 2017). Deep dense overflows from the Nordic Seas and the associated mixing with and entrainment of local waters also add important hydrographic connections between the regions and the surface and deep layers. In both instances, one expects that changes in surface density will modulate these vertical exchanges and the properties of the deep and intermediate waters formed, and thus the AMOC. Many model studies (at various spatial resolution) also suggest an influence of surface salinity changes in this region on changes in the AMOC (Frankignoul et al. 2009; Rahmstorf et al. 2015, Böning et al. 2016). Whereas the connection between these surface changes and AMOC is still subject to discussion (Lozier 2012; Williams et al. 2015; Buckley and Marshall 2016), the relation between changes of surface density and upper or intermediate water formation is rather well established (Yashayaev and Loder 2016 2017; Piron et al. 2017). Seawater density is a function of temperature and salinity, with the relative influence of salinity increasing when temperature decreases. Thus, investigating surface density variability and attempting to establish its past patterns of change requires a concurrent analysis of temperature and salinity fields.

Hydrographic data have been used to analyze salinity and temperature with a low spatial resolution over the North Atlantic and neighboring seas, but only on decadal or longer time scales since the mid-1960s and in more restricted areas since 1950 (Polyakov et al. 2005; Skliris et al. 2014; Holliday et al. 2015; Yashayaev 2007; Yashayaev and Loder 2016 2017). These data also rarely provide direct indications of winter surface density field (the local indications can be found in Yashayaev and Loder (2016 2017), but not broad-scale) that is the variable key to the connection between the surface ocean and the ocean interior. In the last twenty years, the advent of profiling floats (Lavender et al. 2000; Roemmich et al. 2009;

Riser et al. 2016) has allowed year-round three-dimensional monitoring of large-scale upper and intermediate ocean variability (Tesdal et al. 2018), but these measurements do not yet help to resolve the decadal to multidecadal frequencies. Multidecadal time series with at least annual or seasonal resolution have also been produced at a few sites: Rockall Trough (Holliday et al., 2015), the Faroe-Shetland Channel (Hughes et al., 2012), south of Iceland (Icelandic hydrographic surveys), in the center of the Labrador Sea (Yashayaev and Loder 2016 2017), or at station Mike and hydrographic sections in the Norwegian Sea (Yashayaev and Seidov 2015). Otherwise, mapped monthly fields of T and S have been constructed by different objective mapping methods based on the EN4 (Good et al. 2013), CORA (Cabanes et al. 2013) and Ishii et al. (2006) datasets that could have some coverage in a large part of this domain for the last 65 years. Monthly mapped sea surface temperature (SST) fields have also been produced with sufficient coverage into large parts of the SPG since the 1890s using mostly ship log data with a shift in the last two decades to drifter SST data. Carefully estimated bias correction has been applied to subsets of these data, such as done in Hadley Centre SST (HadSST3; Kennedy et al. 2011a and 2011b). However, these binned SST datasets present gaps in some northern regions, in particular before 1922.

Long-term box averages in the SPG with near-annual resolution were presented in Reverdin (2010) and Friedman et al. (2017). Reverdin (2010) described a record of SST and SSS in the northeastern part of the SPG from 1895–2009. There, co-variability of temperature (T) and salinity (S) was found on decadal or multidecadal time scales in different seasons, with a slight dominance of temperature on density variability (partially compensated by the salinity contribution). Friedman et al. (2017) examined pluri-annual (1-2-1 smoothed over successive years) salinity time series in larger boxes over the Atlantic from 20°S–70°N (the temperature reports associated with the salinity data were not examined). The largest salinity root mean square (RMS) variability was found in the northern tropics, from 5°–20°N. In the subpolar North Atlantic, RMS variability was shown to decrease to the northeast from the Labrador Sea to the Nordic seas (Friedman et al. 2017, Fig. 1), though the large grid box size did not permit finer examination of the spatial structure.

Building on the box averaging in these previous two studies, we present a higher-resolution binned product of temperature and salinity since 1896 in the subpolar North Atlantic and southern Nordic seas and draw first conclusions obtained with the dataset. Focusing on the region north of 40°N allows us to describe co-located SSS and SST at higher spatial scales

(less than 500km) than presented previously, without having large periods with data gaps (except during or just after WWI and WWII in some sub-regions).

In this paper we present and validate the binned product at pluri-annual time scales, and also discuss its RMS variability. A second part of this study (in preparation) examines the decadal and multidecadal variability, and relationship with North Atlantic climate variations. The dataset sources and construction are described in Section 2. In Section 3 we validate the dataset and compare it with other products. In Section 4, we examine the spatial distribution of temporal variability of SSS, SST, and surface density, followed by summary and discussion in Section 5.

2. Data and methods

2.1. Dataset construction

The main data sources are described in Reverdin et al. (1994), Reverdin (2010), and Friedman et al. (2017). A significant part of the temperature and salinity (reported as practical salinity) data was collected on merchant and other selected vessels (including the weather ships), in particular before the mid-1970s north of 50°N, and before WWII and since 1977 south of 50°N. This is also the case since 1993, mostly along two ship tracks between southern Newfoundland and Reykjavik and between the North Sea and Greenland (Reverdin et al. 2018), as well as near France and between the English Channel and eastern North America. For the first two tracks, intake temperature was measured part of the time, XBTs were dropped and water samples collected, which allows for identification / estimation of possible temperature and salinity biases in the records from thermosalinographs (TSG) (Alory et al. 2015; Reverdin et al. 2018). The identification of the temperature biases is less certain for the other vessels since 1993, but water samples were also collected to correct salinity biases in TSG records. This is complemented by near-surface hydrographic data during cruises, and since 1996, by measurements from drifters, and upper level data (often near 5-8m depth) from Argo floats (Roemmich et al. 2009; Riser et al. 2016), as well as from earlier PALACE floats (Lavender et al. 2000; Davis et al. 2001).

The depth of sampling and the methods of collection and analysis have changed in time, as well as data accuracy, and thus some subsets of the data require correction of identified biases, both for temperature and salinity, as commented in the earlier papers. Bias correction

follows Reverdin et al. (1994) for ship data before the 1990s (based on comparison with surface data from Nansen casts). There is nonetheless the possibility of seasonal stratification between the data collected below the surface (for example, from TSG data, Argo floats and CTD/Nansen casts) and the surface data (for example from drifters, or in earlier times from bucket sampling). In particular, the change in depth of sampling in time can result in artificial time variability, due to near-surface stratification. In this region and for the data considered, this effect/bias is expected to be usually small, in particular between 55°N and 65°N due to the rarity of low wind conditions, and away from shelves with surface freshwater sources (melting sea ice or icebergs, outflows from fjords, rivers).

We construct annual time series of deviations from the seasonal cycle in boxes for four different seasons, using a similar methodology as in Friedman et al. (2017) and applied from 1896–2015 for both T and S, over smaller spatial bins. This new box-product will be referred to as BINS. The boundaries of the 34 boxes were redesigned to better fit with oceanic fronts and main currents than in Friedman et al. (2017). One box in the central Labrador Sea corresponds to what is used by Yashayaev and Loder (2016) for creating a combined Labrador Sea time series, and box boundaries in the eastern subpolar North Atlantic are chosen as in Holliday et al. (2015). Except for one box off southwest Greenland and one box on the southern part of the Grand Banks, the boxes do not overlap with the shelves and with areas of large seasonal sea ice cover. In the northeastern SPG, the resulting boxes provide a slightly lower spatial resolution than what was used in Reverdin (2010), but for which the different time series were rather well correlated. To give an idea of the resolution achieved, the SSS and SST late-winter climatologies are mapped onto the BINS grid boxes (**Fig. 1**), which shows that the main characteristics of spatial variability are retained by this choice of grid.

The construction of box time series and their error estimates is described in Friedman et al (2017). Here, the main steps are summarized, with specifics and additions related to BINS further described. The SSS and SST climatologies were constructed on a 1°×1° grid with little spatial smoothing and gap-filling, with an adjustment to the mean and up to three sinusoidal harmonics of the year, when data was sufficient. Deviations from this climatology of individual data are estimated (T being in °C and S reported in the practical salinity scale 1978 (PSS-78; Fofonoff 1985; UNESCO 1981). For T, outliers were removed using a spatially variable threshold, based on the local standard deviation. [S outliers were removed similarly,

with less influence on the variability]. We compared this to a fixed 3°C error threshold, which yielded generally similar results albeit with smaller interannual excursions. Larger differences were found in the central Labrador Sea in the 1920s, a period with extremely low data coverage. In each box, these individual deviations are then combined for the individual seasons and years, by median-averaging the individual deviations. The seasonal anomalies are 1-2-1 smoothed over successive years (not applied at the start and end years) and then averaged annually from December–November, inversely weighted by error. [The annual period was incorrectly stated as March–February in Friedman et al. (2017)]. Before being combined, the seasonal anomalies were adjusted to a common baseline over the 40-year period from 1956–1995, which was chosen due to data coverage.

The fall season (September–November) is not included in the central Labrador and southwest Labrador boxes due to the noisy and poorly correlated S time series in the fall in these two regions. In the Norwegian Sea, central Labrador, southwest Labrador, and West Greenland, the winter season (December–February) data are not included before 1947 due to insufficient coverage.

The total number of years with data is shown in **Fig. 2a**. Coverage is below 100 years in the North Irminger Sea and central Labrador Sea grid boxes; all other grid boxes have more than 100 years of coverage. The periods of missing data occur mainly during WWI and WWII, and around 1900 (**Fig. 2b**). Multi-year data gaps are linearly interpolated, and attributed an error, equal to the standard deviation of the whole time series (which is probably an overestimate). A few missing end years (mostly 1896) in some boxes were filled by extending the first non-missing value. As in Friedman et al. (2017), regional average error terms are constructed by RMS averaging the error terms of their constituent individual boxes (which assumes that the errors are uncorrelated in different boxes). The BINS SSS regional average fields are compared with those of Friedman et al. (2017) in Appendix A for the common period 1896-2013, showing differences most commonly within the ±2 standard error ranges.

Figs. **2c–2d** present an estimate of the average pluri-annual grid box error as the mean error term divided by the error term during the year when there was no coverage and the error is maximum (1943, except in the West Greenland box, when it is 1917). This is thus an indication of the ratio of noise relative to signal. This shows for S, the smallest relative errors are in the central SPG. For temperature, the largest relative errors are in the western part of

the domain, where intra-annual variability and eddy variability are largest and not adequately sampled in this data set. The normalized area-averaged error for SSS and SST is shown in **Fig. 2e** and mirrors data coverage, with a larger value in the last year (2015) because error is not reduced in that year by the filtering.

Density is computed from the annual binned SST and SSS values using EOS80 (UNESCO, 1981), and referring the annual anomalies to the March average T and S values. Because SST is lower in winter and density becomes less sensitive to temperature variations when temperature decreases, this choice will slightly decrease the dependency of density on SST compared to using an annual average reference. What motivated this choice was that we wanted to have an estimate typical of winter conditions. Density errors are calculated from the errors in pluri-annual SSS and SST, assuming that they are uncorrelated.

2.2 Seasonal dependency of the pluri-annual variability

The combination of the four seasonal time series to create an annual time series in each bin implicitly assumes that the four seasons portray comparable signals/patterns of pluri-annual variability. This is particularly important if one season is missing, as is the case for the winter season in some boxes for the first part of the records. To get a sense of how well this holds for individual seasons, we combine the time series of individual bins into seven main multi-box regions (**Fig. 3a**): the central Labrador Sea / West Greenland Shelf (3 grid boxes), the southern Nordic seas (3 grid boxes), the north-east part of the SPG (5 grid boxes), the central SPG (6 grid boxes), a large central area of the intergyre gyre south of 50°N (8 boxes), an eastern Atlantic region (two boxes), and four boxes in the southwestern region around the Grand Banks. The regions are chosen for similar source data and coherent variability signals (except for the last one). For example, the two Norwegian Sea grid boxes contain mostly Norwegian and Swedish surface data before WWII, and afterwards much hydrographic data. The central Labrador Sea / West Greenland region contains hydrographic data from the Ice Patrol and other cruises in the southwest, and a mix of hydrographic data and merchant ship data until 1960 in the shelf box close to Greenland. South of 50°N, many different countries contributed to the surface data before WWII, and the French data sources are a large contribution since the mid-1970s. Three grid boxes are not included in any of these regions due to low seasonal correlations in S and more varied data sources: the southwest Labrador Sea, south Greenland/ southwest Irminger Sea, and the western area north of the Gulf Stream.

The seasonal correlations are summarized in **Table 3**. For the central SPG and north-east SG regions (**Fig. 3b-c**, rows 1 and 2), the time series are most complete in all seasons. There, the different seasons present rather similar variability, both for T and S, although for T, the summer (June-August season) deviates at times from the other seasons and presents more variance in the central SPG. Furthermore, after 1960, S anomalies in Sep-Nov in central SPG tend to be below the ones in the other seasons, and the opposite before. This might either be the sign of a real long-term seasonal change, or the effect of changes in the dataset around 1960 (transition in this region from bucket collection to other means of collecting water, for example).

In the southern Nordic seas (**Fig. 3b-c**, bottom row), the different seasons also present rather similar variability, at least since 1950. Before that, this remains largely true for SSS, albeit with more variability in summer. For SST, this also holds before WWI, however between 1920 and 1940, the different seasons present non-coherent variability. We suspect that this might be caused by inhomogeneities in the dataset, either spatially, with a southward shift of the sampling in late autumn to early spring in this period, or due to lower data quality and larger random as well as systematic errors affecting the results and lowering correlations, such as the documented use of engine room or intake temperature data from Norwegian vessels. This is for this reason and the limited winter sampling coverage that we chose not to incorporate the winter season in the annual averages before 1947 in the two Norwegian Sea boxes.

In the central Labrador Sea / West Greenland Shelf region (**Fig. 3b-c**, row 3), on the other hand, while there is some similarity between the different seasons in SST (less so for the summer season), there is much less coherence in SSS, even during the last 50 years of more regular sampling. Notice in particular the large differences between the spring and autumn seasons, which originate from the two boxes in central Labrador Sea and along west Greenland. These differences (and the much larger RMS variability in autumn) justify the choice of not including the autumn season when combining the different seasons to create the BINS time series in this region. In the Grand Banks region, however, we have less correlation with winter season but with similar variance (**Table 3**), so that we chose to retain the winter season. Elsewhere, and the statistics associated with the comparison between the seasonal time series (**Table 3**) justify the choice to have retained all seasons in order to reduce the sampling uncertainty in the BINS pluri-annual time series.

3. Dataset validation

In order to validate the BINS SST and SSS time series, we compare it first to local compilations of high-quality hydrographic data. We also compare the binned time series to different gridded time series to better characterize the properties of BINS. This includes comparison to gridded analyzed T-S products from hydrographical data from 1950-2015, and with the HADSST3 gridded SST product from 1896-2015. In addition, Appendices A and B show comparisons with Friedman et al (2017) and Reverdin et al. (2018), which are based on different binning (in time and space) of similar data sources.

3.1 BINS and hydrographic time series

There are a few published compiled hydrographic time series that are available in the SPG, either from repeated seasonal cruises or a compilation of data from different origins: the Rockall Trough time series since 1975 (Holliday et al. 2015); South Icelandic stations (Sevlogsbanki 5 (Sb5) and Stokksnes 5 (St5)) since the 1970s; and the central Labrador Sea since 1941 (Yashayaev and Loder, 2009 2016). Notice also the time series in the Nordic Seas summarized by (Yashayaev and Seidov 2015). The sites of the Rockall Trough and central Labrador time series fit rather well within some BINS grid boxes, and the South Icelandic time series are just to the north of one BINS grid box (**Fig. 4a**). The results based on hydrographic time series often comprise vertical averages, excluding near-surface (5 to 10 m thick) layer, which is often determined by specifics of performing oceanographic stations and technological limitations. This averaging depth range was not optimized, but in Rockall Trough and near Iceland, it fits with the expected local maximum winter mixed layer. The time series are then averaged over annual means (after removal of an annual cycle), and smoothed with a 1-2-1 filter over successful years, similar to BINS.

The comparisons of these upper ocean time series with BINS are presented in **Fig. 4b** (with statistics in **Table 4**). At each site BINS is correlated with the vertically integrated time series, although the surface time series tend to have larger amplitudes. In the Labrador Sea, we compared different ranges, with the best correlation for the comparison to 20-50 m, which is within the mixed layer at least 6 months of the year and closely follows variations in seasonal discharge of freshwater and its interannual variations. There is still a correlation when considering the averages over the 20-200 m layer, which is within the winter mixed layer most years. The thicker layer time series however present less extreme amplitudes than the

shallower ones. In the Rockall Trough, we also plot the surface time series based only on winter data, which is also correlated with the vertically integrated time series. The largest difference is found for South Iceland, with a low BINS S in the late 1980s that is not found in the hydrographic time series. However, there is also a difference in latitude between the BINS box and the site further north of the hydrographic time series.

Overall, we find that BINS surface time series are significantly correlated at 0-lag with upper ocean time series from hydrography, both for T and S (correlation coefficients are in the range 0.73 to 0.88 for the different time series of salinity or temperature), and exhibit very comparable RMS variability, as described in **Table 4a**. The analysis of lag correlation (summarized in **Table 4b**) indicates that the annually averaged BINS T and S precede by one year the vertically integrated upper ocean analysis both at the Rockall and Southern Iceland sites (for SST, the one-year lag correlation coefficients are respectively 0.80 and 0.89; and for SSS, respectively 0.77 and 0.88), as expected from winter ventilation of the upper ocean and time series analysis of the 20 to 40-year long weather-ship time series in the SPG (Reverdin et al., 1997). Interestingly, this is not the case for the Norwegian Sea time series.

3.2. BINS and gridded objective analyses

We compare the BINS time series with other gridded products of hydrographic data at their near-surface level that we average over the bins of the BINS analysis and filter in time to correspond to the same time resolution (1-2-1 running average over successive years December to November). This includes the 1°×1° EN4, version 4.2.1 (Good et al. 2013) fields from 1950-2015, as well as the 1950-2012 1°×1° product of Ishii et al. (2006), version 6.12 (referred to as ISHII). We also compared BINS to the CORA5.0 fields (Cabanes et al. 2013) that have been extended to 1950, but that exclude Nansen cast data, and thus cannot be used for the analysis of salinity before 1980, and have much coarser spatial sampling in temperature than EN4 or ISHII. On the other hand, when sampling is sufficient (mostly, in the northeastern part), results with CORA are comparable to the ones of ISHII, and will not be discussed further.

Because of the differences in sampling and mapping, the comparison in individual boxes with BINS often presents a large scatter. We will thus average the time series over the same domains as for the seasonal time series (**Fig. 3a)**. [One difference is that the central SPG and north-east SPG are combined into one SPG region]. Except in the Labrador Sea or near the

Grand Banks, the time series of the three products (**Fig. 5**) are well correlated (**Table 5**). Correlation coefficients are larger for SST than for SSS, and for SSS, are usually larger between ISHII and BINS than between EN4 and BINS. In the Labrador Sea (row 2), there are differences both for SST and for SSS. For SSS, the large positive peak in BINS originate from a data gap in one of the three bins that are averaged in the regional time series. On the other hand, the altogether weaker SST-signal in ISHII in this box, something also found, but to a lesser domain in the other regions might result from the objective mapping method used or from more strident tests on outliers (Ishii and Kimoto, 2009). We also notice in the southern Labrador Sea for the last 20 years smaller positive SSS (and SST) anomalies in BINS than in the other two products, possibly a result of the thermosalinograph (TSG) data incorporated in BINS and not in the other products (Reverdin et al. 2018). On the other hand, near the Grand Banks during the last 20 years, T is higher in BINS than in the other products. Elsewhere, the major variability is portrayed with similar amplitude and roughly at the same time in the different products. For example in the SPG or southern Nordic seas, the salinity time series portray the events sometimes described as great salinity anomalies (Belkin et al. 1998), the largest one in the central SPG happening in the early 1970s and referred to as the Great Salinity Anomaly (Dickson et al. 1988). The figures also suggest larger differences between the different products in the 1990s, in particular in the intergyre region.

The high overall agreement might result from a share of common hydrographic data incorporated in the different products. However, there is also a large portion of specific data that are not in all the products (T from mechanical or expandable bathythermographs are used in EN4 and ISHII, but not in BINS, and different sets of surface data are in BINS, but not used in EN4 and ISHII). There are also large methodological differences between the different products, both in data selection and validation and in mapping techniques and how seasonal anomalies are grouped or not. These comparisons, as well as the comparisons presented in 3.1, thus reinforce our assessment of usually small sampling errors for the BINS product during that period, and that the product captures the major events of pluri-annual variability at the grid box scale since 1950.

3.3 BINS and HadSST3

For long-term surface temperature, we make use of comparisons with the HadSST3 monthly product (Kennedy et al. 2011a and 2011b). Data in BINS and HadSST3 originate from different data streams. Although some data are certainly incorporated in both products, the

overlap is probably rather minimal, as discussed in Reverdin et al. (1994). In particular, the data before 1920 north of 58°N in BINS originate mostly from Danish and Norwegian sources which are not incorporated in HadSST3. In the last two decades, data from TSGs on merchant vessels or from profiling floats are a large contribution to BINS but are not included in HadSST3, whereas the drifter SST data used in HadSST3 are mostly not in BINS. Even in the post-WWII to mid-1970s period, when ocean weather ships contribute both to HadSST3 and BINS, temperature data in the two products often originate from different sensors. BINS and HadSST3 also account for biases and uncertainties differently. We note that HadSST3 does not apply smoothing or interpolation between neighboring grid cells, facilitating more direct regional comparisons.

We examine the monthly HadSST3 anomalies, version 3.1.1.0 (median realization). We area-average the HadSST3 grid boxes into similar sub-regions similar to those shown in **Fig. 3a**. The southwest Labrador Sea BINS grid box is also examined, which closely corresponds to one HadSST3 grid box. Generally however, the large 5°×5° HadSST3 grid boxes do not align directly with BINS grid boxes; the specific grid boxes are shown in **Fig. C1**. We average the HadSST3 anomalies into 3-month seasons (skipping over missing values). The seasonal anomalies are 1-2-1 smoothed (excluding endpoints) over successive years, and the seasonal anomalies are averaged into Dec–Nov annual means (again skipping over missing values). The winter season is not included in the southern Nordic, central Labrador, or southwest Labrador annual means before 1947 to match the BINS coverage. Likewise, the fall season is excluded from the central Labrador and southwest Labrador sub-regions. In general, we find that the BINS data are more seasonally coherent at low frequencies before WWII (not shown).

The HadSST3 and BINS regional time series are compared in **Fig. 6**, with the respective 1896-2015 and 1950-2015 correlations listed in **Table 6**. Overall, the products are very highly correlated since 1950. In the first half of the record, the largest differences between the products are generally found during periods of data gaps and linear interpolation near the end of WWI and WWII, when BINS error estimates are also large. Among regions, the agreement is less good in the southern Nordic seas (**Fig. 6b**) before the 1940s; we also find that the HadSST3 seasons are also much less correlated before the 1940s (not shown). To some extent, the early differences come from the most poorly sampled northwestern part of this region, but also from Danish and Norwegian data incorporated in this product and not in

HadSST3. Furthermore, in the 1930s, the differences might also originate from Norwegian data interpreted in BINS as recorded continuously on a water line inside the ship, and thus attributed a large positive bias. The central Labrador Sea / West Greenland shelf (**Fig. 6a**) also shows less agreement before 1930. The products correspond very well after the 1930s, although HadSST3 has less of a cooling in the 1980s. The two SPG regions are both very strongly correlated over the length of the record (**Fig. 6c,d**). The Southwest Labrador Sea grid box (**Fig. 6e**) generally corresponds well after the gap periods in the early 1900s. In the regions further south (**Fig. 6f,g,h**), the two time series usually show a good agreement, and are most of the time compatible with the BINS error estimate. The largest differences are near the Grand Banks (**Fig. 6g**) before 1920, maybe as a result of different seasonal coverage; and in the intergyre region (**Fig. 6f**) in the early 1950s. This is the only place where the error estimate does not explain the difference between the curves, maybe because data in BINS for some of these boxes were only available in the summer season for a few years in the early 1950s.

4. Characteristics of T, S and surface density RMS variability

In this section, we examine the linear trends and the RMS variability of the linearly detrended pluri-annual time series. The 1896–2015 BINS least-squares linear trends are shown in **Fig. 7a–7c**. We construct error estimates for the linear trends by block resampling using the time series error estimates (resampled every 2 years to account for the 1-2-1 filter). The SSS trend (**Fig. 7a**) shows a negative trend throughout most of the domain, except for the southwest and part of the eastern boundary; the pattern is consistent with the 1896–2013 trend shown in Friedman et al. (2017). The SST trend (**Fig. 7b**) shows strong warming in the Gulf Stream, and to a lesser extent along the eastern margin. There is cooling in West Greenland / Central Labrador, and near 52-55°N/30°W. However, error bars due to higher frequency variability or sampling are often large. Both the increase in T and the decrease in S contribute to an overall decrease in surface density (**Fig. 7c**).

The 1896–2015 HadSST3 trend is shown in (**Fig. 7d**). The trends are calculated for grid boxes with 5 or less years with missing data. [Note that specific seasons are not removed from the spatial grid boxes as for the HadSST3 indices discussed above]. Error terms for HadSST3 are computed by multiplying the white-noise standard error of the slope by a factor of sqrt(2) on account of the 1-2-1 filter. Like BINS, the HadSST3 trend shows significant warming in the Gulf Stream extension, and also along the eastern margin. HadSST3 also shows a region

of cooling in the central SPG from 50-60°N, referred to as the 'warming hole' (Drijfhout et al., 2012), which is not well reproduced in BINS. Checking this particular difference suggests that data in the pre-1917 period originated mostly from one group of data from the ICES archive, often reported at set longitudes, and for which we applied an overall correction of -0.2°C on temperature, but with little comparison data to define it (Reverdin et al. 1994). Notice however that nonetheless, for the area average of the central SPG (**Fig. 6c**), the difference between BINS and HADSST3 is not significant for that period.

The detrended RMS standard deviation of the pluri-annual T and S time series (**Figs. 8a** and **8b)** correspond to the expectation that large-scale temperature RMS standard deviation diminishes from west to east; whereas the salinity RMS values are largest in the west, then near 50°N with a decrease from west to east and from the southern to the northern SPG. This pattern is indicative on one hand of larger air-sea and hydrographical/heat forcings in the western SPG, and on the other hand of the presence during some winters of a rather shallow mixed layer in the western SPG (even in the Labrador Sea, where deep convection also often takes place (Yashayaev and Loder 2016)). In mid-ocean near 45-50°N, there might also be a contribution of displacements of hydrographical fronts.

Surface density standard deviation (**Fig. 8c**) diminishes both from west to east and south to north. The relative contributions of T and S to detrended density variability are shown in **Fig. 8d**. The contributions of SST and SSS to surface density variability are opposite in most places (as there is a positive correlation between SST and SSS variability on intra-annual to centennial time scale in a large part of this domain, so the effects of these variables on density variability tend to compensate each other). This also might explain a pattern of density rms variability which presents weaker zonal gradients in the west than what is found in SST or SSS.

Interestingly, in the westernmost part of the domain, except near the Gulf Stream, the salinity contribution to density variability dominates over the temperature contribution (**Fig. 8d**), whereas in the eastern and northern parts of the SPG, temperature dominates. Thus in the first region, density tends to be positively correlated to SSS and SST, whereas in the other region it is negatively correlated with SSS and SST. In general, the SSS contribution is slightly larger for the non-detrended data (not shown), due to its relatively larger centennial trends in this part of the ocean (discussed in part 2 of this study).

Next, we show the total RMS variability contained at multidecadal frequencies. We low-pass filter the time series using a spectral filter with half power at a period close to 15 years. **Fig. 9** shows the percentage of the detrended pluri-annual variance explained by the low-pass filtered time series. The low-pass filtered variance percentage of SSS (**Fig. 9a**) and SST (**Fig. 9b**) are less correlated spatially than overall RMS variability. For S, what striking is the lower percentages in the western part (as low as 40%) than in the eastern part of the domain (larger than 60%), whereas for T, it is mostly the northern part which has a larger percentage of variance in the multidecadal frequencies, with percentage of detrended variance reaching up to 80%. The low percentages of multidecadal variability for SSS in the northwest is consistent with the penetration in the ocean interior of short-lived freshwater pulses originating from the western shelves and the rim of the Labrador Sea (examples of recent short-term SSS variability in this region are presented in Tesdal et al. (2018)).

The low-pass filtered time series explain comparatively less density variance, particularly in the western SPG region (**Fig. 9c**, values usually less than 50%), which suggests coordinated SSS and SST variations at lower frequencies with thus partial compensation on their contribution to density changes. Indeed, the low-pass filtered correlation of SSS and SST (**Fig. 8d**) is positive in most grid boxes, with large magnitudes in the central SPG, around Greenland and in the south-west north of the Gulf Stream as well as around the Grand Banks. Notice however, less significant correlation in other areas of the western, southern, and eastern parts of the domain close to Europe. The relationship between SSS and SST variations will be examined further in part 2 of this study.

5. Discussion and Conclusions

We have constructed a spatially averaged binned product BINS of pluri-annual T, S, and surface density for 1896-2015 north of 40°N, similar to what was done for S in Friedman et al. (2017), but at a higher spatial resolution. We deliberately excluded the shelf regions (except for one box along southwestern Greenland and one box on the southern Grand Banks), which have different dynamics, variability and sampling issues (Reverdin et al. 2018). We have shown that in recent decades, the BINS pluri-annual variability of SST and SSS is coherent with published time series of upper ocean heat and salinity based on high quality hydrographic data. It provides also rather similar results to an analysis of SST and SSS

variability along two ship tracks since 1993, based on the same in situ data set for its main part (Reverdin et al. 2018).

Furthermore, BINS compares well with two gridded products of hydrographic data (EN4 from 1950-2015 and ISHII from 1950-2012), which use different data selection and mapping methods. This suggests that in this region, the processing and spatial averaging of the individual data in the spatial bins done in BINS is usually sufficient to portray the pluri-annual variability of SST on the grid box scale, at least since 1950, and that during this period, the data set used does not contain large biases compared to these other data sets. On pluri-annual time scales, BINS SST also fits largely with what is portrayed in HadSST3, except in the pre-WWII period for the Labrador Sea and the Norwegian Sea. Insufficient coverage or erroneous data, either in HadSST3 or BINS, might be causing these discrepancies. Thus, more care should be taken when interpreting variability in BINS before 1950.

The product combines time series of pluri-annual variability in different seasons. This was required to reduce sampling errors to a reasonable level (in particular pre-1950), but this might also mask or alias important differences between the seasons. For example in the northern part of the subtropical gyre (thus, south of the region we investigate) a strong seasonal modulation in the recent trends towards increasing SSS has been documented, associated to important seasonal changes of the hydrological cycle (Yu et al. 2017). In the Labrador Sea, too, on an interannual time scale, large summer-time anomalies of SSS have been documented in the recent decade (2008 and 2012) that don't have a clear counterpart in other seasons (Tesdal et al. 2018). Such events, and also the seasonality of atmospheric forcing, seasonal thermal stratification and mixing, could induce large pluri-annual differences between the anomalies in different seasons. Indeed, **Fig. 3** suggests that the differences between seasons are large in the central Labrador Sea, but for the other open-ocean regions, different seasons present more similar decadal and longer frequencies. Possibly, insufficient accuracy of the binned time series and the change in collection and measurement techniques could preclude further investigation of the seasonality of the decadal or longer variability with this dataset. For example, in boxes of the southern SPG, we noticed SSS differences in the autumn around 1960 that could be a data artifact. Differences with HadSST3 taking place around 2000 could also be data-related (either in HadSST3 with drifting buoy data becoming more numerous, or in BINS with the Argo data afterwards).

Nonetheless, a more homogeneous 25-year dataset in the central SPG resolving seasonal variability along well sampled ship lines (Reverdin et al. 2018) also pointed out a similarity in pluri-annual variability for the different seasons, although interannual RMS variability tended to be larger there too in early summer for SST.

This aside, the seasonal inhomogeneity in data coverage, with less sampling in winter in particular before 1955, should have little impact on the reconstructed pluri-annual time series. We however warn that this might not be a wise substitute for the seasonal time series, in particular for studies of winter conditions the central Labrador Sea or water mass formation, but then obviously the available binned data present more gaps, and the original seasonal time series have larger error bars.

Density is a non-linear function of temperature and salinity, with the character of non-linearity increasing at low T. We deliberately emphasized winter characteristics when estimating density from temperature and salinity time series, but use the BINS time series which combine SST and SSS anomalies in the four (or three) different seasons. We noticed that in some regions pluri-annual variability has smaller amplitude in winter (Table 3 summarizing **Fig. 3**), thus this choice might result in an overestimation of winter-time pluri-annual density variability. It would be also misleading to use these density pluri-annual time series as a proxy for water density in other seasons.

We do not have a direct comparison of SSS BINS time series prior to the 1950s. We base our evaluation that they provide reasonable estimates for this early period mostly on comparisons of the SST BINS time series with HadSST3 binned data. However, before 1950 in the Nordic Seas and Labrador Sea, the binned data have larger uncertainties and winter seasons are poorly (or not at all) sampled. Thus, clearly our estimates for SSS are also less certain in these regions prior to 1950. A possibility for further validation would be to compare with independent proxy-based estimates of surface variability. Few such proxy time series have been derived in this region resolving the multidecadal variability in SST and even less in SSS, and their uncertainties are probably too large for such a validation (Hall et al. (2010), for a site in the Iceland Basin (Gardar drift); Richter et al. (2009), for a site in the northern Rockall Trough area (Feni drift); Moffa-Sanchez et al. (2017) and references therein as well as Thornhalley et al. (2018), for a site in the southeastern Labrador Sea).

BINS has sufficient spatial resolution to resolve large-scale patterns of RMS standard deviation of the variability in SST, SSS and density (this was done after removing linear trends). RMS variability in SSS and to a lesser degree in SST increases towards the west, with the contribution to density of SSS variability dominating that of SST in the west, whereas in the eastern and northern parts of the SPG, temperature typically dominates the variability. The portion of variability in multidecadal frequencies (here with a cut-off at 15 years) could be ascertained. It is found both for SST and SSS to be a very large part of pluri-annual variability, with lesser relative contribution in the west for SSS and in the southern part for SST. For density, it is less prominent, except in the northeastern part of the SPG, where SST variability contribution dominates and is very much at multidecadal frequencies (also see Holliday et al. 2015). The higher frequencies are more influenced by the sampling uncertainties, in particular the effect of interpolation and interannual smoothing during data gaps in either specific seasons or all seasons (near WWII, for example). Thus this analysis of the relative importance of multidecadal variability is more dataset-dependent than the RMS variability

Acknowledgements

This is a contribution to the French SSS observation service, which is supported by French agencies INSU/CNRS, IRD, CNES and IPEV, as well as from SOERE CTDO2. The Rockall Trough time series were provided with support from the UK Natural Environment Research Council (Extended Ellett Line program, National Capability). The station time series south of Iceland were provided with Icelandic support. The annual oceanographic monitoring of the Labrador Sea was initiated as a Canadian contribution to the World Ocean Circulation Experiment in 1990 and is presently conducted as a core component of the Atlantic Zone Off-Shelf Monitoring Program (AZOMP) run by the Bedford Institute of Oceanography of Fisheries and Oceans Canada. The International Argo Program is part of the Global Ocean Observing System (Argo, 2000). Argo data are available from the Coriolis Global Data Center, Institut français de recherche pour l'exploitation de la mer (Ifremer). The HadSST3 and EN4 data were provided by the Met Office Hadley Center, and the ISHII data were provided by the NCAR Research Data Archive. A.R.F. was supported by SOERE CTDO2 and the ERC funded project TITAN (EC-320691). L.C. acknowledges support from the Swedish National Space Board (SNSB; Dnr 133/17).

References

Alory G, Delcroix T, Téchiné P, Diverrès D, Varillon D, Cravatte S, … Roubaud F (2015) The French contribution to the voluntary observing ships network of sea surface salinity. Deep Sea Res I 105: 1–18. http://doi.org/10.1016/j.dsr.2015.08.005

Argo (2000) Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE. http://doi.org/10.17882/42182

Belkin I M, Levitus S, Antonov J, Malmberg S.-A (1998) "Great Salinity Anomalies" in the North Atlantic. Prog Oceanogr 41(1): 1–68. http://doi.org/10.1016/S0079-6611(98)00015-9

Böning C W, Behrens E, Biastoch A, Getzlaff K, Bamber J L (2016) Emerging Impact of Greenland Meltwater on Deepwater Formation in the North Atlantic Ocean. Nature Geoscience. https://doi.org/10.1038/ngeo2740.

Buckley M W, Marshall J (2016) Observations, Inferences, and Mechanisms of the Atlantic Meridional Overturning Circulation: A Review. Reviews of Geophysics 54 (1): 2015RG000493. https://doi.org/10.1002/2015RG000493.

Cabanes C, Grouazel A, von Schuckmann K, Hamon M, Turpin V, Coatanoan C, Paris F, et al (2013) The CORA Dataset: Validation and Diagnostics of in-Situ Ocean Temperature and Salinity Measurements. Ocean Sci 9 (1): 1–18. https://doi.org/10.5194/os-9-1-2013.

Davis R E, Sherman J T, Dufour J (2001) Profiling ALACEs and other advances in autonomous subsurface floats. J Atmos Ocean Tech 18: 982-993.

Dickson R R, Meincke J, Malmberg M S-A, Lee A J (1988) The ‘‘Great Salinity Anomaly’’ in the northern North Atlantic 1968 – 1982. Prog Oceanogr 20: 103– 151.

Drijfhout S, van Oldenborgh G J, Cimatoribus A (2012) Is a Decline of AMOC Causing the Warming Hole above the North Atlantic in observed and modeled warming patterns?" J Clim 25: 8373–79. https://doi.org/10.1175/JCLI-D-12-00490.1.

Ebisuzaki W (1997) A method to estimate the statistical significance of a correlation when the data are serially correlated. J Clim 10(9): 2147–2153.

Fofonoff N P (1985) Physical properties of seawater: A new salinity scale and equation of state for seawater. J Geophys Res 90(C2): 3332–3342. https://doi.org/10.1029/JC090iC02p03332.

Frankignoul C, Deshayes J, Curry R (2009) The Role of Salinity in the Decadal Variability of the North Atlantic Meridional Overturning Circulation. Climate Dynamics 33(6): 777–93. https://doi.org/10.1007/s00382-008-0523-2.

Friedman A R, Reverdin G, Khodri M, Gastineau G (2017) A new record of Atlantic sea surface salinity from 1896 to 2013 reveals the signatures of climate variability and long-term trends. Geophys Res Lett 44: 1866–1876. DOI:10.1002/2017GL072582.

Good S-A, Martin M J, Rayner N A (2013) EN4: Quality Controlled Ocean Temperature and Salinity Profiles and Monthly Objective Analyses with Uncertainty Estimates. J Geophys Res 118(12): 6704–16. https://doi.org/10.1002/2013JC009067.

Hall I R, Boessenkool K P, Barker S, McCave I N, Elderfield H (2010) Surface and deep ocean coupling in the subpolar North Atlantic during the last 230 years. Paleoceanogr 25. doi:10.1029/2009PA001886.

Holliday N P, Cunningham S A, Johnson C, Gary S F, Griffiths C, Read J F, Sherwin T (2015) Multidecadal variability of potential temperature, salinity, and transport in the eastern subpolar North Atlantic. J Geophys Res 120. doi:10.1002/2015JC010762.

Hughes, S. L., Holliday, N. P., Gaillard, F., and the ICES Working Group on Oceanic Hydrography, 2012. Variability in the ICES/NAFO region between 1950 and 2009: observations from the ICES Report on Ocean Climate. – ICES Journal of Marine Science, doi:10.1093/icesjms/fss044.Icelandic hydrographic surveys …

Ishii M, Kimoto M, Sakamoto K, Iwasaki S I (2006) Steric sea level changes estimated from historical ocean subsurface temperature and salinity analyses. J Oceanography 62(2): 155-170.

Kennedy J J, Rayner N A, Smith R O, Parker D E, Saunby M (2011) Reassessing Biases and Other Uncertainties in Sea Surface Temperature Observations Measured in Situ since 1850: 1. Measurement and Sampling Uncertainties. J Geophys Res 116. doi:201110.1029/2010JD015218.

Kennedy J J, Rayner N A, Smith R O, Saunby M, Parker D E (2011) Reassessing biases and other uncertainties in sea-surface temperature observations since 1850: 2. Biases and homogenisation. J Geophys Res 116 D14104. doi:10.1029/2010JD015220

Lavender K L, Davis R E, Owens W B (2000) Mid-depth recirculation observed on the interior Labrador and Irminger seas by direct velocity measurements. Nature 607: 66- 69.

Lozier M S (2012) Overturning in the North Atlantic. Ann Rev of Marine Sci 4(1): 291–315. http://doi.org/10.1146/annurev-marine-120710-100740

Moffa-Sanchez P, Hall I R (2017) North Atlantic variability and its links to European climate over the last 3000 years. Nature comm. 8:1726, 555. doi:10.1038/s41467-017-01884-8.

Piron A, Thierry V, Mercier H, Caniaux G (2017) Gyre-scale deep convection in the subpolar North Atlantic Ocean during winter 2014–2015. Geophys Res Lett 44: 1439–1447. doi:10.1002/2016GL071895.

Polyakov I V, Bhatt U S, Simmons H L, Walsh D, Walsh J E, Zhang X (2005) Multidecadal Variability of North Atlantic Temperature and Salinity during the Twentieth Century. J Clim 18(21): 4562–4581. http://doi.org/10.1175/JCLI3548.1

Rahmstorf S, Box J E, Feulner G, Mann M E, Robinson A, Rutherford S, Schaffernicht E J (2015) Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. Nature Climate Change 5(5): 475–480. http://doi.org/10.1038/nclimate2554

Reverdin G (2010) North Atlantic Subpolar Gyre Surface Variability (1895–2009). J Clim 23(17): 4571–4584. http://doi.org/10.1175/2010JCLI3493.1

Reverdin G, Cayan D, Dooley H, Ellett D, Levitus S, du Penhoat Y, Dessier A (1994) Surface Salinity of the North-Atlantic - Can We Reconstruct Its Fluctuations Over the Last 100 Years. Prog Oceanogr 33(4): 303–346. http://doi.org/10.1016/0079-6611(94)90021-3

Reverdin G, Cayan D, Kushnir Y (1997) Decadal variability of hydrography in the upper northern North Atlantic 1948-1990. J Geophys Res 102: 8505-8531.

Reverdin G, Alory G, Diverres D, Bringas F, Goni G, Heilmann L, Chafik L, Szekely T, Friedman A R (2018) North Atlantic subpolar gyre along predetermined ship tracks since 1993: a monthly data set of surface temperature, salinity, and density. Earth Syst Sci Data 10: 1403-1415. https://doi.org/10.5194/essd-10-1403-2018.

Rhein M, Kieke D, Hüttl-Kabus S, Roessler A, Mertens C, Meissner R, Klein B, Böning C W, Yashayaev I (2011) Deep water formation, the subpolar gyre, and the meridional overturning circulation in the subpolar North Atlantic. Deep Sea Res II 58: 1819-1832. doi:10.1016/j.dsr2.2010.10.061

Richter T O., Peeters F J C, Weering T C E (2009) Late Holocene (0-2.4 ka BP) surface water temperature and salinity variability, Feni Drift, NE Atlantic Ocean. Quat Sci Rev 28: 1941-1955.

Riser S C, Freeland H J, Roemmich D, Wijffels S, Troisi A, Belbéoch M, Gilbert D, et al. (2016) Fifteen Years of Ocean Observations with the Global Argo Array. Nature Climate Change 6 (2): 145–53. https://doi.org/10.1038/nclimate2872.

Roemmich, D, Johnson G C, Riser S, Davis R E, Gilson J, Owens W B, Garzoli S L, Schmid C, Ignaszewski M (2009) The Argo Program: Observing the global ocean with profiling floats. Oceanography. 22:34-43. 10.5670/oceanog.2009.36

Skliris N, Marsh R, Josey S A, Good S A, Liu C, Allan R P (2014) Salinity changes in the World Ocean since 1950 in relation to changing surface freshwater fluxes. Climate Dynamics 43(3-4): 709–736. http://doi.org/10.1007/s00382-014-2131-7

Tesdal J, Abernathey R P, Goes J I, Gordon A L, Haine T W (2018) Salinity Trends within the Upper Layers of the Subpolar North Atlantic. J Clim 31: 2675–2698. https://doi.org/10.1175/JCLI-D-17-0532.1

Thornhalley D J R, Oppo D W, Ortega P, Robson J I, Brierley C M, Davis R, Hall I R, Moffa-Sanchez P, Rose N L, Spooner P T, Yashayaev I, Keigwin L D (2018) Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. Nature 556. https://doi.org/10.1038/s41586-018-0007-4.

UNESCO (1981) The Practical Salinity Scale 1978 and the International Equation of State of Seawater 1980. UNESCO technical papers in marine science 36: 25pp.

Williams R G, Roussenov V, Lozier M S, Smith D (2015) Mechanisms of heat content and thermocline change in the subtropical and subpolar North Atlantic. J Clim 28: 9803-9813. doi:10.1175/JCLI-D-15-0097.1.

Yashayaev I (2007) Hydrographic changes in the Labrador Sea, 1960-2005. Prog Oceanogr 73: 242-276. doi:10.1016/j.pocean.2007.04.015.

Yashayaev I, Loder J W (2009) Enhanced production of Labrador Sea Water in 2008, Geophys Res Lett 36 L01606. doi:10.1029/2008GL036162.

Yashayaev I, Loder J W (2016) Recurrent replenishment of Labrador Sea water and associated decadal-scale variability. J Geophys Res 121:8095-8114. doi:10.1002/2016JC012046.

Yashayaev I, Loder J W (2017) Further intensification of deep convection in the Labrador Sea in 2016. Geophys Res Lett 44:1429-1438, doi:/10.1002/2016GL071668.

Yashayaev I, Seidov D (2015) The role of the Atlantic water in multidecadal ocean variability in the Nordic and Barents Seas. Prog Oceanogr 132: 68-127.

Yu L, Jin X Liu H (2017) Poleward Shift in Ventilation of the North Atlantic Subtropical Underwater. Geophysical Research Letters 44. doi: https://doi.org/10.1002/2017GL075772

Figures.

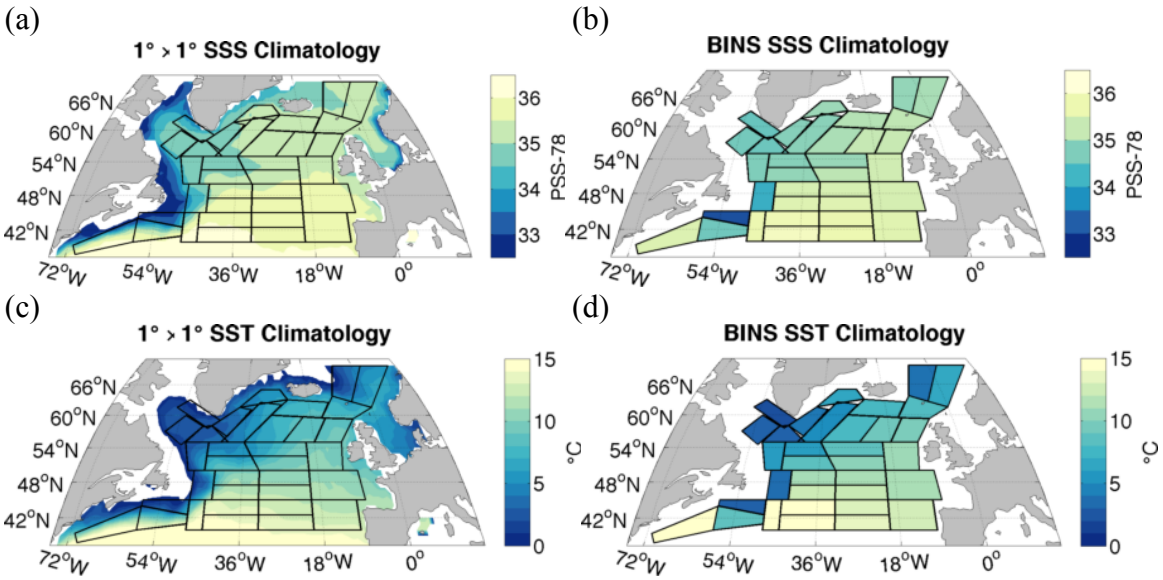


Figure 1. Gridded climatologies (approximately 1896–2000). (a) and (c): 1°×1° March climatologies for (a) SSS and (c) SST. (b) and (d): BINS March–May climatologies for (b) SSS and (d) SST.

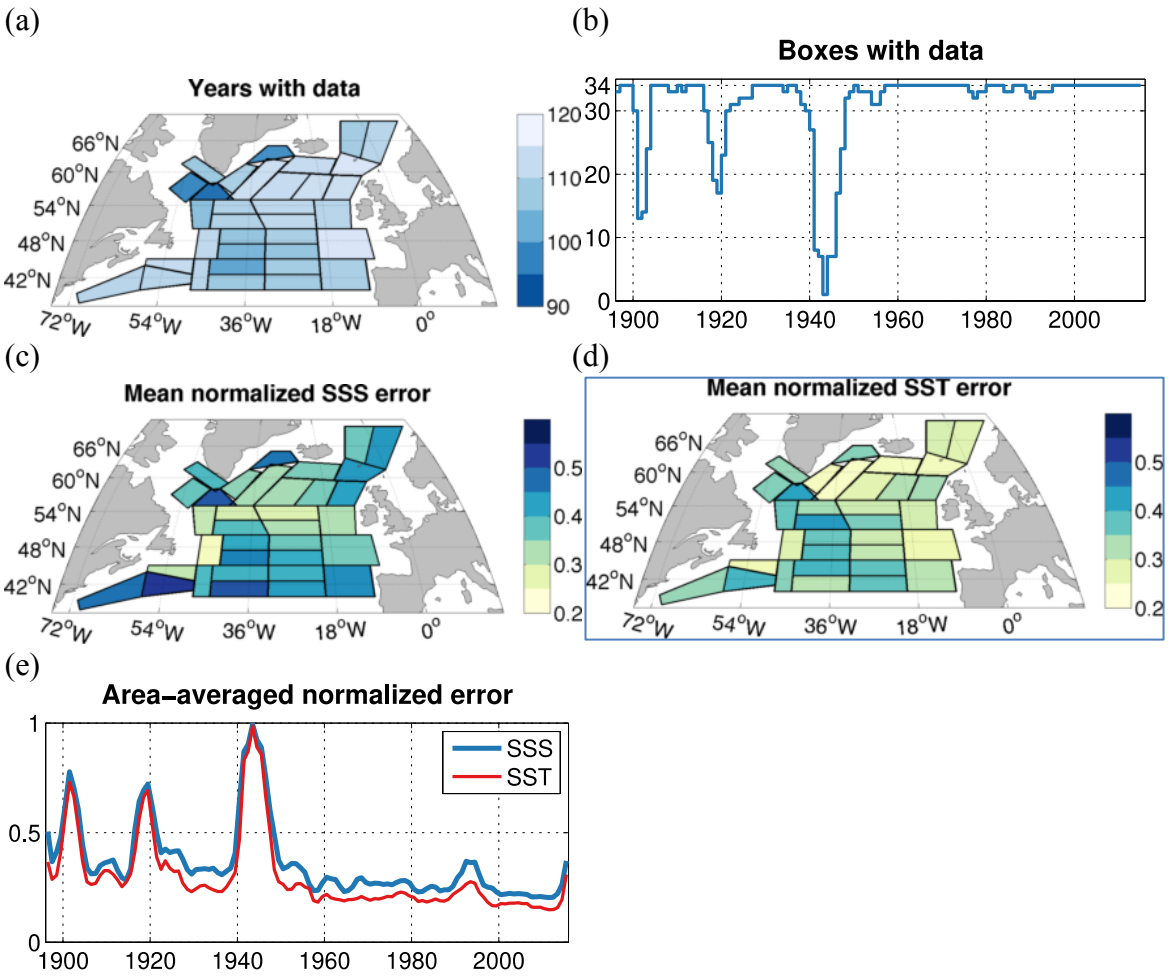
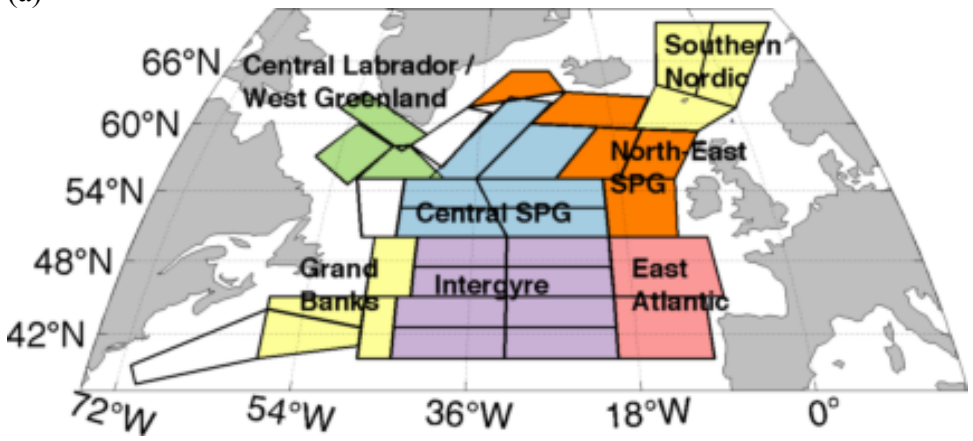
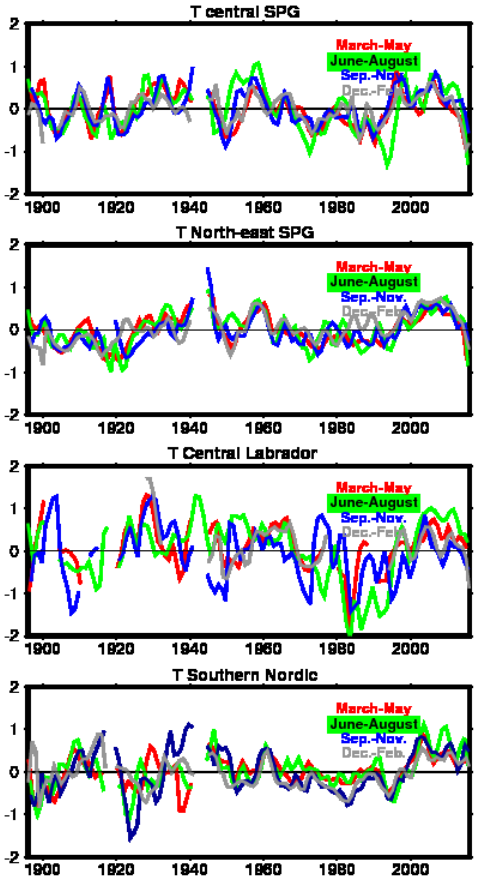


Figure 2. (a) Outline of BINS boxes, with number of years with data from 1896–2015 (out of a maximum of 120). (b) Time series of total BINS boxes with coverage (out of a maximum 34). The data counts in (a) and (b) are from after applying the 1-2-1 filter. (c) Mean annual SSS grid box error, defined as the mean grid box error divided by its maximum error (usually in 1943). (d) Same as (c), for SST. (e) Area-weighted normalized annual grid box error for SSS and SST.

(a)


66°N
60°N
54°N
48°N
42°N
72°W
54°W
36°W
18°W
0°
Central Labrador / West Greenland
Southern Nordic
North-East SPG
Central SPG
Grand Banks
Intergyre
East Atlantic

(b)



(c)

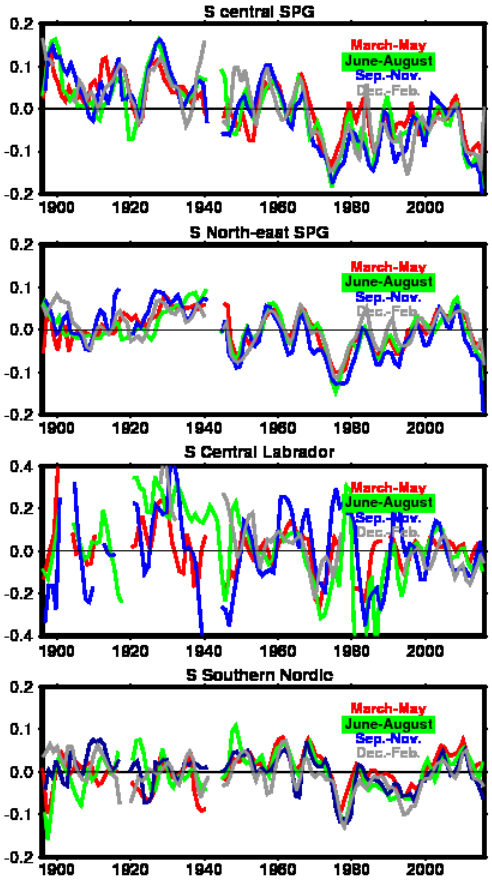


Figure 3. (a) Regional domains described in the text. (b-c) Seasonal time series for (b) T (°C) and (c) S (PSS-78) averaged over four domains portrayed in (a): central SPG, north-east SPG, central Labrador Sea / West Greenland, and the southern Nordic Seas.

Table 3. Seasonal RMS and Pearson correlation coefficient. RMS values are in °C for T, and PSS-78 for S (×1000). Corresponding regions are shown on Fig. 3a.

Region	**Length**		**12-2**	**3-5**	**6-8**	**9-11**		**12-2**	**3-5**	**6-8**	**9-11**
Central SPG	112-116 yrs	**T RMS**	0.304	0.407	0.505	0.404	**S RMS**	73	60	72	79
		T corr.					**S corr.**				
		12-2		0.58	0.52	0.57	**12-2**		0.72	0.77	0.78
		3-5			0.55	0.71	**3-5**			0.79	0.82
		6-8				0.62	**6-8**				0.86

Region	**Length**		**12-2**	**3-5**	**6-8**	**9-11**		**12-2**	**3-5**	**6-8**	**9-11**
N-E SPG	113-115 yrs	**T RMS**	0.339	0.338	0.375	0.336	**S RMS**	42	41	44	55
		T corr.					**S corr.**				
		12-2		0.65	0.54	0.60	**12-2**		0.72	0.75	0.73
		3-5			0.76	0.71	**3-5**			0.81	0.78
		6-8				0.74	**6-8**				0.80

		12-2	**3-5**	**6-8**	**9-11**		**12-2**	**3-5**	**6-8**	**9-11**
Southern Nordic	**T RMS**	0.358	0.351	0.425	0.524	**S RMS**	41	40	47	41
	T corr.					**S corr.**				
	12-2		0.31	0.40	0.61	**12-2**		0.50	0.43	0.66
	3-5			0.69	0.41	**3-5**			0.66	0.52
	6-8				0.57	**6-8**				0.61

		12-2	**3-5**	**6-8**	**9-11**		**12-2**	**3-5**	**6-8**	**9-11**
Central Labrador/	**T RMS**	0.47	0.385	0.482	0.602	**S RMS**	130	95	131	162
West Greenland*	**T corr.**					**S corr.**				
	12-2		0.81	0.20	0.60	**12-2**		0.37	0.50	0.27
	3-5			0.60	0.78	**3-5**			0.56	0.50
	6-8				0.52	**6-8**				0.42

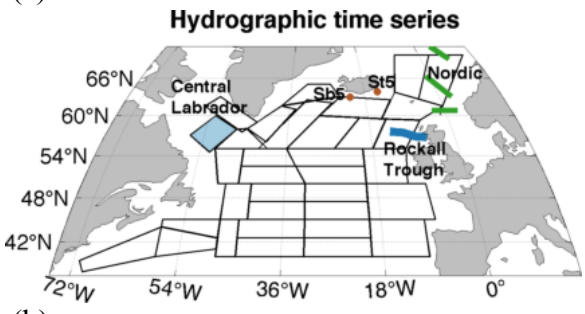
		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Grand Banks	**T RMS**	0.950	0.792	0.890	0.990	**S RMS**	148	156	132	150
	T corr.					**S corr.**				
	12-2		0.26	0.33	0.29	**12-2**		0.09	0.24	0.51
	3-5			0.56	0.21	**3-5**			0.49	0.33
	6-8				0.44	**6-8**				0.50

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Intergyre	**T RMS**	0.49	0.42	0.61	0.49	**S RMS**	87	76	82	100
	T corr.					**S corr.**				
	12-2		0.56	0.19	0.42	**12-2**		0.64	0.65	0.73
	3-5			0.44	0.53	**3-5**			0.51	0.55
	6-8				0.42	**6-8**				0.59

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
East Atlantic	**T RMS**	0.37	0.37	0.41	0.51	**S RMS**	53	61	55	54
	T corr.					**S corr.**				
	12-2		0.59	0.41	0.57	**12-2**		0.40	0.53	0.56
	3-5			0.57	0.52	**3-5**			0.59	0.34
	6-8				0.38	**6-8**				0.51

* small number of common points, and varying number of boxes; thus not reliable, in particular for S.
Both for Southern Nordic and Central Labrador, very different number of years in winter and other seasons (only on the order of 60 years with winter data in some of the boxes included in the average)

(a)



(b)

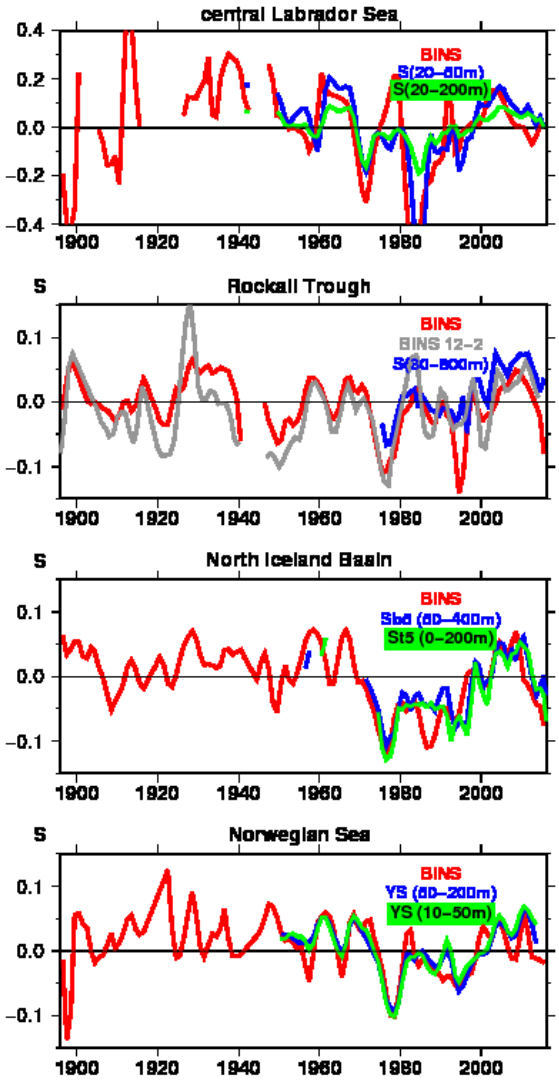


Figure 4. (a) Map showing hydrographic time series locations with BINS grid boxes. (b) Comparison for S of BINS (in red) with hydrographic time series low passed 1-2-1 over successive years; Central Labrador Sea (top; green and blue for different layer averages), Rockall Trough (top middle; the green curve is from BINS but for the winter (Dec-Feb.) season; blue is the vertically averaged (30-800m) S time series), Nor Iceland Basin (lower middle; two vertically averaged station time series are presented); and eastern Norwegian Sea (bottom; the blue and green time series is a compilation of the offshore data (median averaged) in 3 Norwegian Sea sections presented in Yashayaev and Seidov (2015), for two different depth ranges).

Table 4a. Hydrographic sections and BINS: RMS variability for S (PSS-78) and T (°C), and lag-0 correlation coefficient.

		RMS (S)		corr (S)		RMS (T)		corr (T)
Region	**Years S**	**section**	**BINS**		**Years T**	**section**	**BINS**	
S Rockall (30–800m)	37	.040	0.048	0.73	37	0.284	0.350	0.80
S Iceland (50–400m)	47	.050	0.043	0.81	47	0.413	0.391	0.87
Central Labrador (20–50m)	69	.155	0.169	0.76	67	0.715	0.716	0.74
Norwegian Sea (10-50m)	64	.039	0.038	0.82	64	0.31	0.31	0.87

Table 4b. Hydrographic section lag correlation with BINS.

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
South Rockall	**–2**	35	0.57	35	0.56
(30–800m)	**–1**	36	0.70	36	0.65
	0	37	0.73	37	0.80
	+1	37	0.77	37	0.80
(BINS leads)	**+2**	37	0.68	37	0.72

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
South Iceland	**–2**	45	0.54	44	0.59
(50–400m)	**–1**	46	0.68	45	0.74
	0	47	0.81	45	0.87
	+1	47	0.88	45	0.89
(BINS leads)	**+2**	47	0.83	45	0.81

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
Eastern Norwegian Sea	**–2**	64	0.48	64	0.70
(10-50m)	**–1**	64	0.72	64	0.81
	0	64	0.82	64	0.87
	+1	64	0.74	64	0.83
(BINS leads)	**+2**	64	0.54	64	0.72

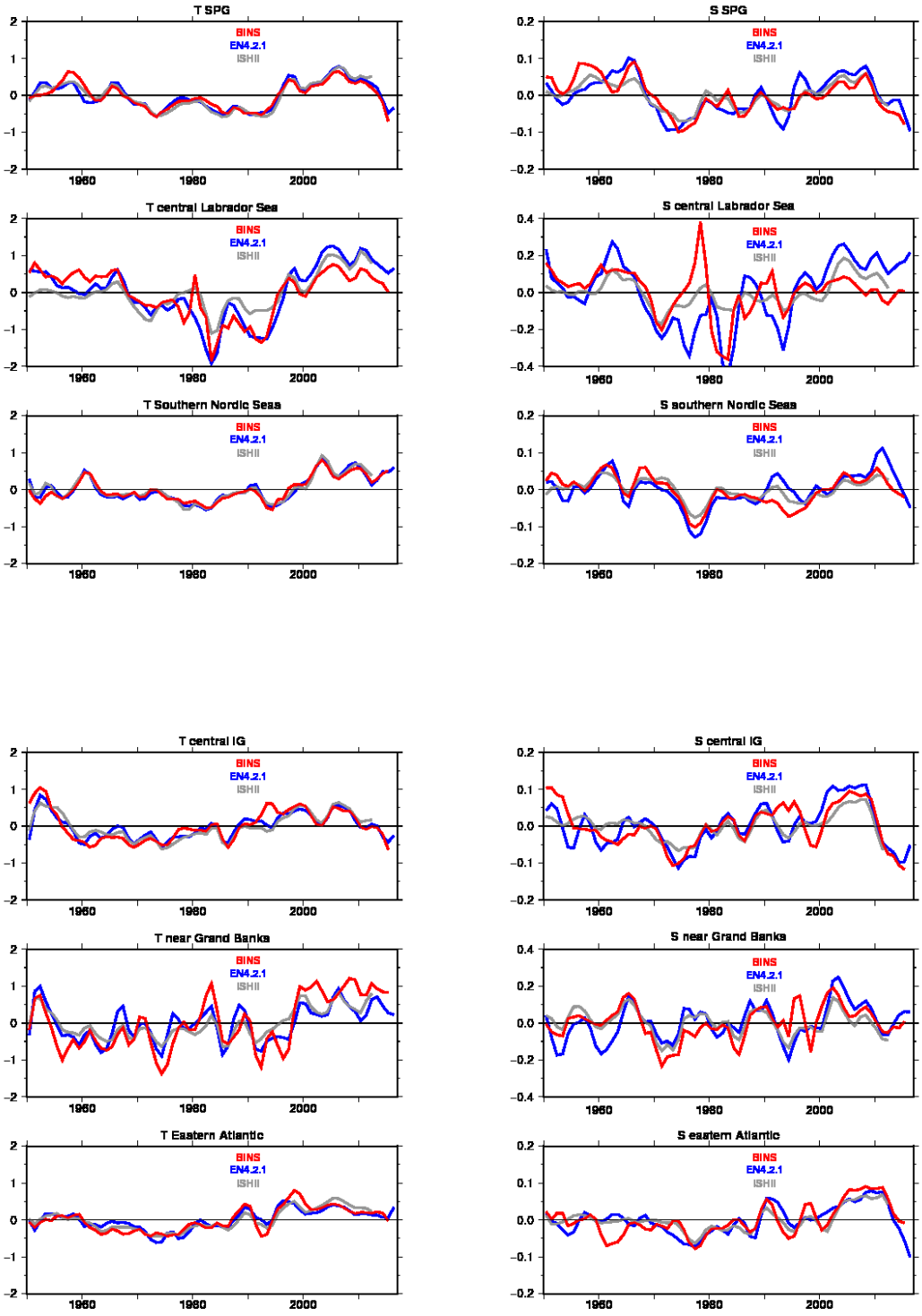


Figure 5. Comparison of regional T (left, °C) and S (right, PSS-78) time series from BINS with EN4 (1950-2015) and ISHII (1950-2012) surface products. The regional domains are the ones presented in Fig. 3. Here, the SPG region includes both the central SPG and the north-east SPG (and the south Greenland/southwest Irminger Sea box).

Table 5. RMS variability and correlation with BINS: (a) EN4, (b) ISHII.

(a) EN4 / BINS		**RMS (S) (PSS-78)**		**corr (S)**	**RMS (T) °C**		**corr (T)**
Region	**Time period**	**EN4**	**BINS**		**EN4**	**BINS**	
SPG	1950–2015	0.050	0.046	0.78	0.36	0.34	0.93
Central Labrador/ West Greenland	1950–2015	0.173	0.121	0.49	0.76	0.63	0.87
Central Labrador/ West Greenland	excluding 1976–1979	0.169	0.112	0.67	0.78	0.64	0.87
Southern Nordic S	1950–2015	0.045	0.039	0.75	0.34	0.32	0.96
Intergyre	1950–2012	0.057	0.057	0.78	0.33	0.43	0.85
Grand Banks	1950–2012	0.098	0.114	0.40	0.47	0.69	0.42
East Atlantic	1950–2012	0.037	0.042	0.73	0.26	0.32	0.84

(b) ISHII / BINS		**RMS (S) (PSS-78)**		**corr (S)**	**RMS (T) °C**		**corr (T)**
Region	**Time period**	**ISHII**	**BINS**		**ISHII**	**BINS**	
SPG	1950–2012	0.034	0.045	0.90	0.38	0.34	0.95
Central Labrador/ West Greenland	1950–2012	0.075	0.124	0.55	0.50	0.64	0.74
Central Labrador/ West Greenland	excluding 1976-79	0.077	0.115	0.61	0.52	0.65	0.75
Southern Nordic	1950–2012	0.028	0.040	0.91	0.35	0.32	0.9
Intergyre	1950–2012	0.034	0.045	0.71	0.34	0.34	0.62
Grand Banks	1950–2012	0.068	0.124	0.45	0.41	0.65	0.42
East Atlantic	1950–2012	0.030	0.040	0.73	0.28	0.32	0.65

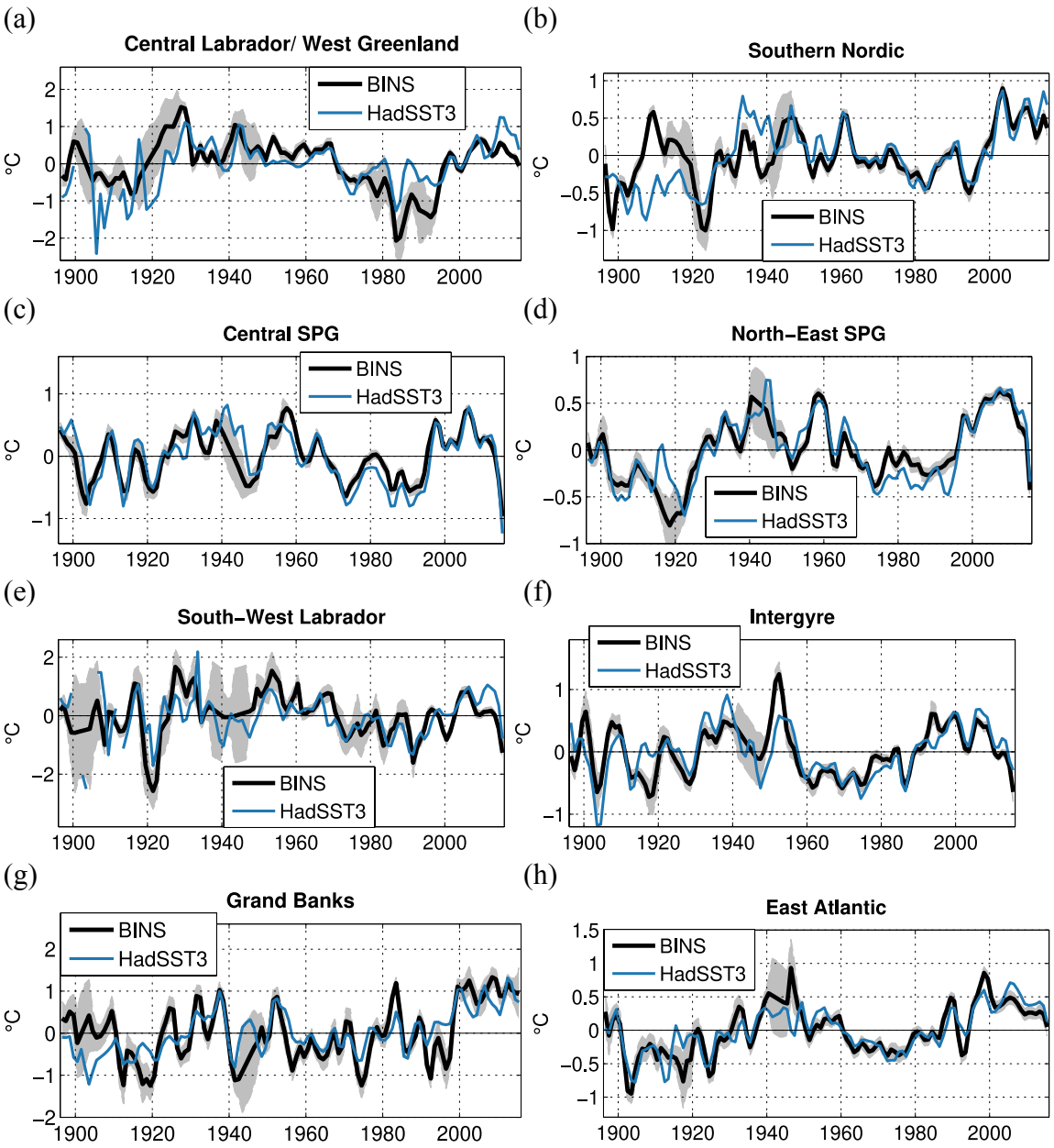


Figure 6. Area-averaged BINS and HadSST3 SST anomalies: (a) central Labrador / West Greenland, (b) southern Nordic, (c) central SPG, (d) north-east SPG, (e) south-west Labrador, (f) intergyre, (g) Grand Banks / Labrador Current, (h) east Atlantic. Shading indicates ±2 BINS error terms. The corresponding regions are shown in **Fig. C1**.

Table 6. HadSST3 correlation with BINS SST

Region	Time period	corr	Time period	corr
Central SPG	1896–2015	0.84	1950–2015	0.95
N-E SPG	1896–2015	0.87	1950–2015	0.93
Central Labrador/ West Greenland	except 1900–01, 1910	0.57	1950–2015	0.72
SW Labrador	except 1900–01, 1904–1905, 1912	0.56	1950–2015	0.71
Southern Nordic	1896–2015	0.65	1950–2015	0.93
Intergyre	1896–2015	0.74	1950–2015	0.84
Grand Banks	1896–2015	0.67	1950–2015	0.82
East Atlantic	1896–2015	0.82	1950–2015	0.89

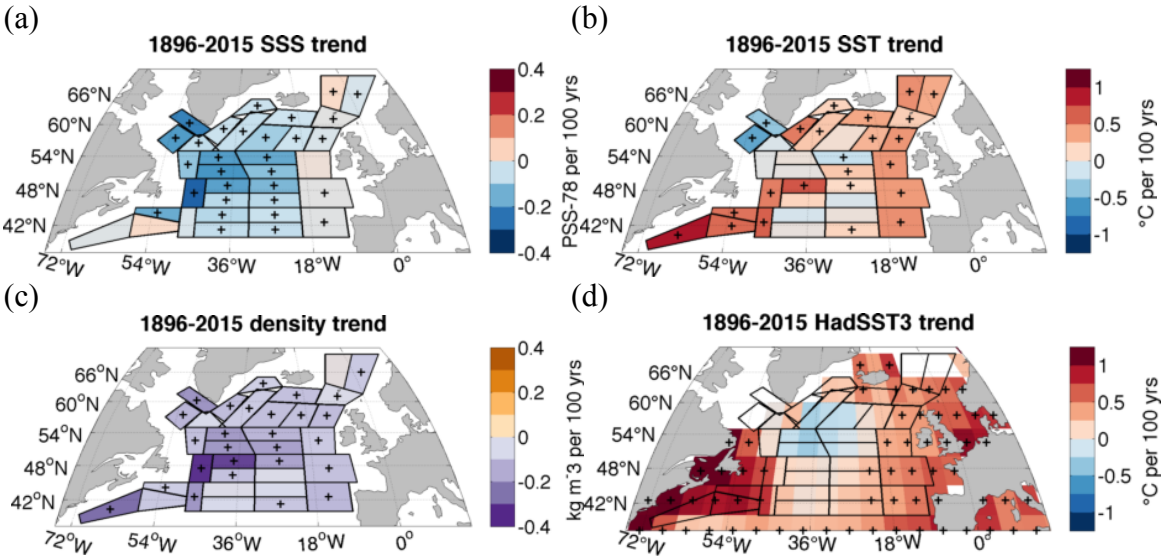


Figure 7. Trend over 1896–2015 (a) SSS; (b), SST; and (c) density; per 100 years. Pluses indicate where the slope magnitude is larger than twice the estimated error. (d) HadSST3 trend, per 100 years.

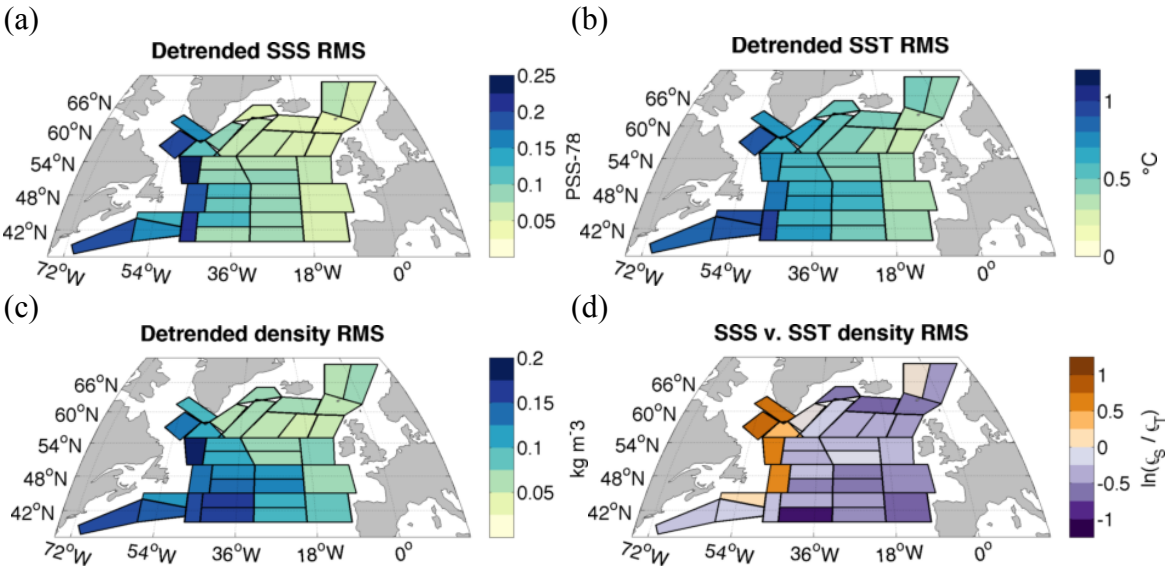


Figure 8. Detrended pluri-annual (1-2-1 filtered) RMS variability for (a) SSS, (b) SST, and (c) density (ρ). (d) Ratio of the contributions of detrended SSS and SST to density RMS variability (plotted as a logarithm).

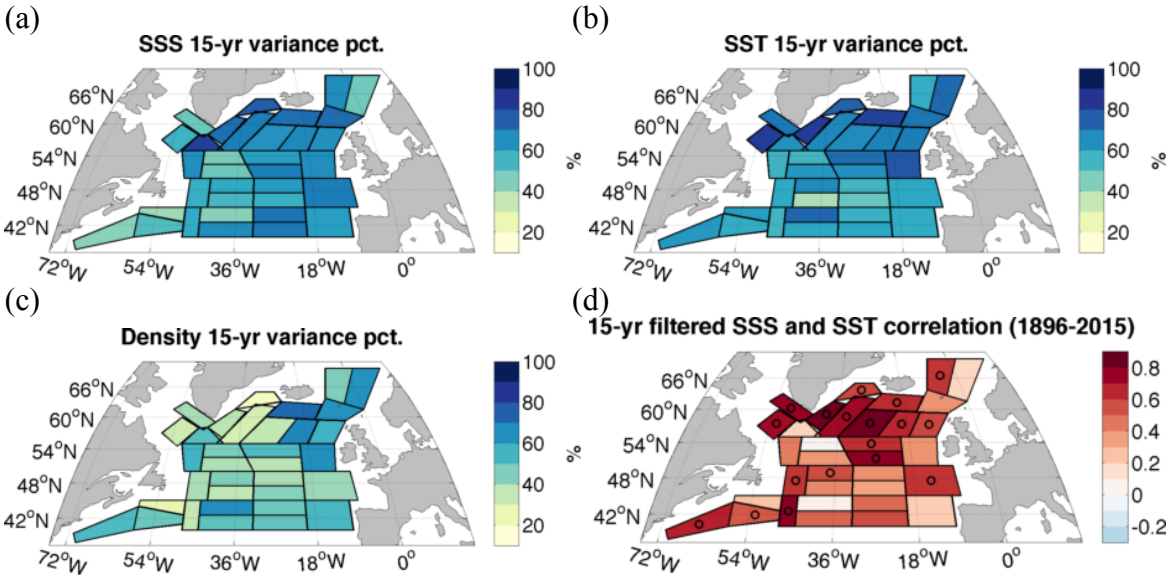


Figure 9. Percentage of detrended pluri-annual (1-2-1) variance explained by the 15-year low-pass filtered time series for (a) SSS, (b) SST, and (c) density. (d) Detrended correlation coefficient between the low-pass filtered SSS and SST. Circles indicate where the correlation is significant at $p<0.05$, estimated with a random-phase bootstrap test to account for serial autocorrelation (Ebisuzaki, 1997).

Appendix A. Comparison with Friedman et al. (2017)

The BINS boxes mainly use the same underlying SSS data as the large boxes north of 40°N in Friedman et al. (2017). The datasets cover a similar area (**Fig. A1a**), though as mentioned previously, the BINS boxes more carefully avoid shelf regions (except for southwest Greenland and the southern part of the Grand Banks). Most of the source data used in the two analyses are the same. Additionally, BINS also incorporates two small datasets from the 1900s and 1910s, plus a few recent transects (and 2014–2015). There are more gaps in the 19 BINS time series than in Friedman et al. (2017): in particular in 1918-1921, and during and just after WWII; these gaps are linearly interpolated in the smaller boxes. Also, when sampling is poor, but varying geographically within the larger boxes, it is possible that some spatial variability is aliased in the temporal variability in the larger boxes.

Fig. A1b compares the NATL index from Friedman et al. (2017), area-averaged SSS from 45°–62°N, with SSS averaged over a similar area in BINS. [NATL from Friedman et al. (2017) is only plotted through 2012, as 2013 was subject to endpoint smoothing]. The two products are very highly correlated (r=0.94, 1896–2012), and compatible considering the differences in area and error estimates. Greater differences are found for smaller regions in the first half of the record, particularly during the gap years mentioned above (not shown)

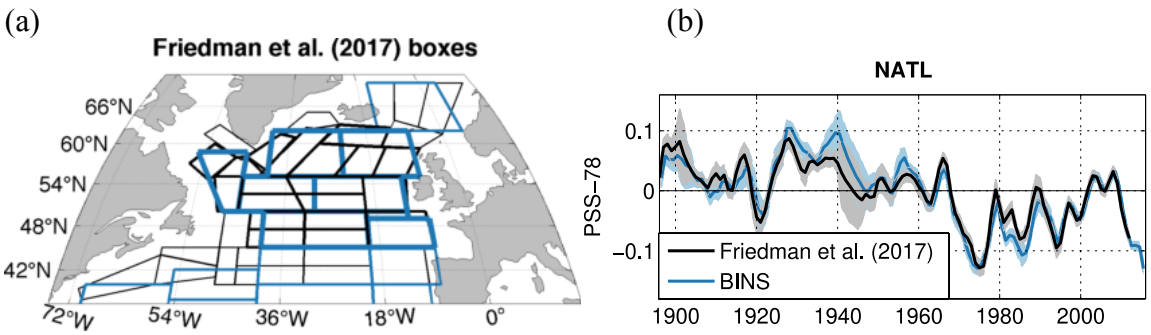


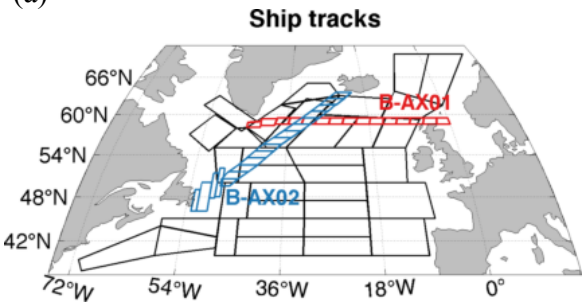
Figure A1. (a) Grid boxes from Friedman et al. (2017) (blue) and BINS (black). Thick lines indicate the NATL region. (c) Comparison of NATL index from Friedman et al. (2017) and BINS grid boxes. Anomalies are from 1896–2012; shading indicates ±2 error terms.

Appendix B. Comparison with Reverdin et al. (2018)

As a check on the BINS time series, we compare the boxes with time series constructed with mostly similar data, monthly binned along two ship routes since mid-1993, intersecting near 59.5°N/32°W: AX02 between Iceland and southern Newfoundland and AX01 between the North Sea and southern Greenland, mostly along 59.5°N (Reverdin et al. 2018), shown in **Fig. B1a**. Time series along AX02 start in July 1993 with few gaps, whereas for AX01 some large data gaps were filled until late 1997. These time series, referred to as B-AX01 and B-AX02, provide increased spatial resolution at seasonal time scales, and portray very coherent variability where they intersect.

We illustrate the comparison of pluri-annual variability of T and S between B-AX01 and BINS for overlapping boxes in the common period 1993-2015 (**Fig. B1b**). For S, the corresponding filtered RMS variability is larger in B-AX01 than in BINS by up to 20%, but with very high correlation coefficients (all larger than 0.95). The smaller RMS amplitudes of salinity in BINS probably result from the larger box sizes and the resulting spatial averaging. RMS variability is more similar for T, but with a slightly smaller correlation in the Iceland Basin (0.80) where gaps in 1993-1997 were the most common in B-AX01. Altogether, the comparisons for the two ship tracks suggest that the method used for BINS in the box averaging to produce pluri-annual variability yields correct results when data coverage is sufficient.

(a)



(b)

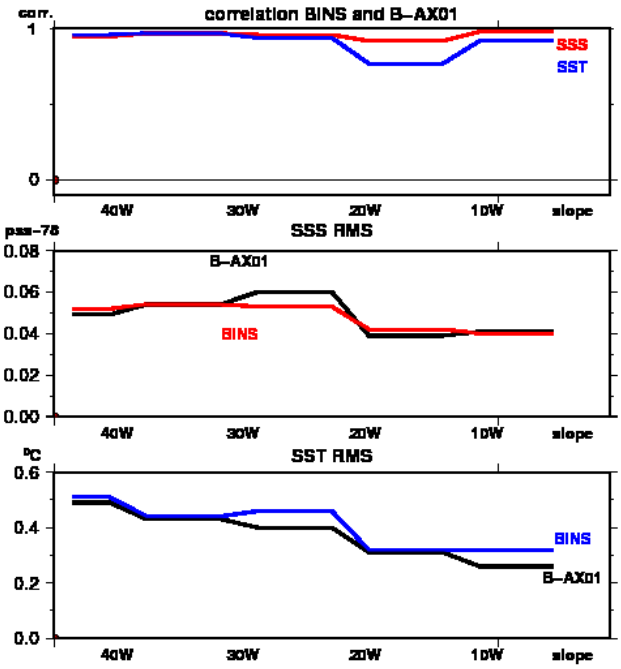


Figure B1. (a) AX01 (red) and AX02 (blue), from (Reverdin et al. 2018). (b) Comparison of B-AX01 with BINS, 1993–2015. The monthly time series of B-AX01 have been yearly-averaged (Dec-Nov), low-pass (1-2-1) filtered over successive years, and then averaged over the bins in BINS: correlation coefficient (top) for T (blue) and S (red), and RMS (middle for S and bottom for T) (B-AX01 in black, BINS in blue (T) and red (S)).

Appendix C: Locations of BINS and HadSST3 grid boxes.

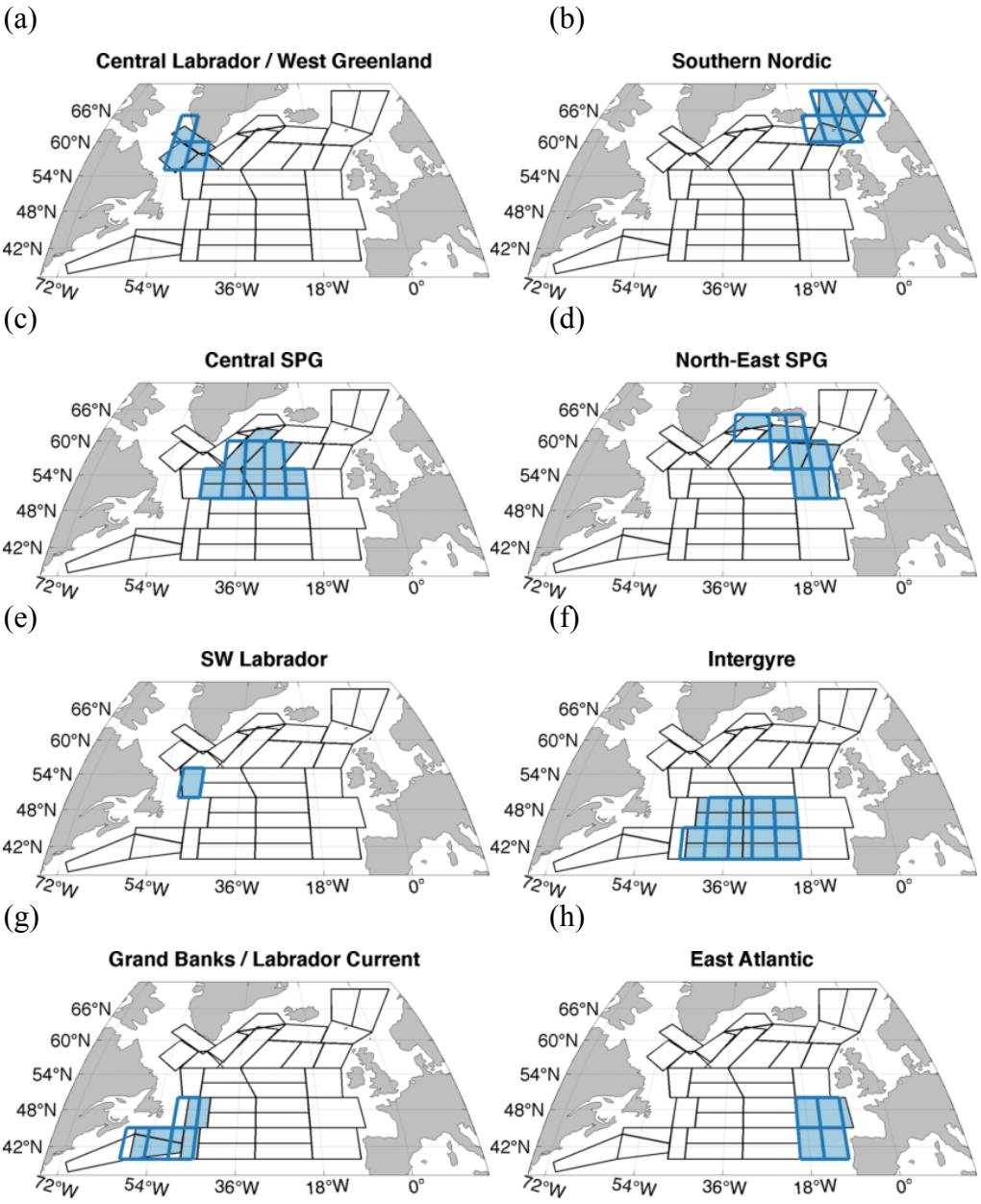


Figure C1. Locations of BINS and HadSST3 grid box regions used in the Section 3.3. Thick blue lines show the 5°×5° HadSST3 grid boxes; shading shows the corresponding BINS grid boxes: (a) central Labrador / West Greenland, (b) southern Nordic, (c) central SPG, (d) north-east SPG, (e) south-west Labrador, (f) intergyre, (g) Grand Banks / Labrador Current, (h) east Atlantic.

Click here to view linked References

Click here to access/download;Manuscript;SPG.part1.16112018.pdf

North Atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. Part 1: Dataset validation and RMS variability

G. Reverdin[1], A. R. Friedman[2], L. Chafik[3], N. P. Holliday[4], T. Szekely[5], H. Valdimarsson[6], I. Yashayaev[7]

[1] Sorbonne-Université, CNRS/IRD/MNHN (LOCEAN), Paris, France. ORCID ID https://orcid.org/0000-0002-5583-8236

[2] School of Geosciences, University of Edinburgh, UK. ORCID ID https://orcid.org/0000-0001-6994-2037

[3] Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden. ORCID ID https://orcid.org/0000-0002-5538-545X

[4] National Oceanography Centre, Southampton, UK. ORCID ID https://orcid.org/0000-0002-9733-8002

[5] IUEM, Brest, France

[6] Marine and Freshwater Research Institute, Reykjavik, Iceland

[7] Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS, Canada

Corresponding author: G. Reverdin, Laboratoire d'Océanographie et de climatologie par expérimentation et analyse numérique, Institut Pierre Simon Laplace, Sorbonne-Université, case 100, 4 pl. Jussieu, 75252 Paris Cedex 05, France (gilles.reverdin@locean-ipsl.upmc.fr ; tel : 33-1-44272342 fax : 33-1-44273805 ; orcid.org/0000-0002-5583-8236)

Abstract

We present a binned annual product (BINS) of sea surface temperature (SST), sea surface salinity (SSS) and sea surface density (SSD) observations for 1896–2015 of the subpolar North Atlantic between 40°N and 70°N, mostly excluding the shelf areas. The product of bin averages over spatial scales on the order of 200 to 500 km, reproduces most of the interannual variability in different time series covering at least the last three decades or of the along-track ship monitoring. Comparisons with other SSS and SST gridded products available since 1950 suggest that BINS captures the large decadal to multidecadal variability. Comparison with the HadSST3 SST product since 1896 also indicates that the decadal and multidecadal variability is usually well reproduced, with small differences in long-term trends or in areas with marginal data coverage in either of the two products. Outside of the Labrador Sea and Greenland margins, interannual variability is rather similar in different seasons. Variability at periods longer than 15 years is a large part of the total interannual variability, both for SST and SSS, except possibly in the south-western part of the domain. Variability in SST and SSS increases towards the west, with the contribution of salinity variability to density dominating that of temperature in the western Atlantic, except close to the Gulf Stream and North Atlantic Current in the southwest area. Weaker variability and larger relative temperature contributions to density changes are found in the eastern part of the gyre and south of Iceland.

Keywords: sea surface temperature, sea surface salinity, surface density, North Atlantic, decadal variability

1. Introduction

The North Atlantic subpolar gyre (SPG) receives salty surface water from the subtropical North Atlantic as well as fresh water originating from the Arctic and the Nordic seas and circulating mostly on and along the shelves. It is a formation site of deep and intermediate waters and thereby a major contributor to the lower limb of the Atlantic meridional overturning circulation (AMOC) (Rhein et al. 2011). There is wintertime densification of the surface Labrador and Irminger Seas resulting in intermittent deep convection reaching and exceeding 2400 m in the mid-1990s and 2000 m in recent years (Yashayaev and Loder 2017). Deep dense overflows from the Nordic Seas and the associated mixing with and entrainment of local waters also add important hydrographic connections between the regions and the surface and deep layers. In both instances, one expects that changes in surface density will modulate these vertical exchanges and the properties of the deep and intermediate waters formed, and thus the AMOC. Many model studies (at various spatial resolution) also suggest an influence of surface salinity changes in this region on changes in the AMOC (Frankignoul et al. 2009; Rahmstorf et al. 2015, Böning et al. 2016). Whereas the connection between these surface changes and AMOC is still subject to discussion (Lozier 2012; Williams et al. 2015; Buckley and Marshall 2016), the relation between changes of surface density and upper or intermediate water formation is rather well established (Yashayaev and Loder 2016 2017; Piron et al. 2017). Seawater density is a function of temperature and salinity, with the relative influence of salinity increasing when temperature decreases. Thus, investigating surface density variability and attempting to establish its past patterns of change requires a concurrent analysis of temperature and salinity fields.

Hydrographic data in the North Atlantic and neighboring seas have been used to analyze salinity and temperature on decadal time scales with a low spatial resolution , but only since the mid-1960s and in more restricted areas since 1950 (Polyakov et al. 2005; Skliris et al. 2014; Holliday et al. 2015; Yashayaev 2007; Yashayaev and Loder 2016 2017). These data also rarely provide direct indications of winter surface density field (regional estimates can be found in Yashayaev and Loder (2016 2017)) that is the variable key to the connection between the surface ocean and the ocean interior. In the last twenty years, the advent of profiling floats (Lavender et al. 2000; Roemmich et al. 2009; Riser et al. 2016) has allowed

year-round three-dimensional monitoring of large-scale upper and intermediate ocean variability (Tesdal et al. 2018), but these measurements do not yet help to resolve the decadal to multidecadal frequencies. Multidecadal time series with at least annual or seasonal resolution have also been produced at a few sites: Rockall Trough (Holliday et al., 2015), the Faroe-Shetland Channel (Hughes et al., 2012), south of Iceland (Icelandic hydrographic surveys), in the center of the Labrador Sea (Yashayaev and Loder 2016 2017), or at station Mike and hydrographic sections in the Norwegian Sea (Yashayaev and Seidov 2015). Otherwise, mapped monthly fields of T and S have been constructed by different objective mapping methods based on the EN4 (Good et al. 2013), CORA (Cabanes et al. 2013) and Ishii et al. (2006) datasets that could have some coverage in a large part of this domain for the last 65 years. Monthly mapped sea surface temperature (SST) fields have also been produced with sufficient coverage into large parts of the SPG since the 1890s using mostly ship log data with a shift in the last two decades to drifter SST data. Carefully estimated bias correction has been applied to subsets of these data, such as done in Hadley Centre SST (HadSST3; Kennedy et al. 2011a and 2011b). However, these binned SST datasets present gaps in some northern regions, in particular before 1922.

Long-term box averages in the SPG with near-annual resolution were presented in Reverdin (2010) and Friedman et al. (2017). Reverdin (2010) described a record of SST and SSS in the northeastern part of the SPG from 1895–2009. There, co-variability of temperature (T) and salinity (S) was found on decadal or multidecadal time scales in different seasons, with a slight dominance of temperature on density variability (partially compensated by the salinity contribution). Friedman et al. (2017) examined interannual (1-2-1 smoothed over successive years) salinity time series in larger boxes over the Atlantic from 20°S–70°N (the temperature reports associated with the salinity data were not examined). The largest salinity root mean square (RMS) variability was found in the northern tropics, from 5°–20°N. In the subpolar North Atlantic, RMS variability was shown to decrease to the northeast from the Labrador Sea to the Nordic seas (Friedman et al. 2017, Fig. 1), though the large grid box size did not permit finer examination of the spatial structure.

Using similar box averaging as in these previous two studies, we present a higher-resolution binned product of temperature and salinity since 1896 in the subpolar North Atlantic and southern Nordic seas and draw first conclusions obtained with the dataset. Focusing on the region north of 40°N allows us to describe co-located SSS and SST at higher spatial scales

(less than 500km) than presented previously, without having large periods with data gaps (except during or just after WWI and WWII in some sub-regions).

In this paper we present and validate the binned product at interannual time scales, and also discuss its RMS variability. A second part of this study (in preparation) examines the decadal and multidecadal variability, and relationship with North Atlantic climate variations. The dataset sources and construction are described in Section 2. In Section 3 we validate the dataset and compare it with other products. In Section 4, we examine the spatial distribution of temporal variability of SSS, SST, and surface density, followed by summary and discussion in Section 5.

2. Data and methods

2.1. Dataset construction

The main data sources are described in Reverdin et al. (1994), Reverdin (2010), and Friedman et al. (2017). A significant part of the temperature and salinity (reported as practical salinity) data was collected on merchant and other selected vessels (including the weather ships), in particular before the mid-1970s north of 50°N, and before WWII and since 1977 south of 50°N. This is also the case since 1993, mostly along two ship tracks between southern Newfoundland and Reykjavik and between the North Sea and Greenland (Reverdin et al. 2018), as well as near France and between the English Channel and eastern North America. For the first two tracks, intake temperature was measured part of the time, XBTs were dropped and water samples collected, which allows for identification / estimation of possible temperature and salinity biases in the records from thermosalinographs (TSG) (Alory et al. 2015; Reverdin et al. 2018). The identification of the temperature biases is less certain for the other vessels since 1993, but water samples were also collected to correct salinity biases in TSG records. This is complemented by near-surface hydrographic data during cruises, and since 1996, by measurements from drifters, and upper level data (often near 5-8m depth) from Argo floats (Roemmich et al. 2009; Riser et al. 2016), as well as from earlier PALACE floats (Lavender et al. 2000; Davis et al. 2001).

The depth of sampling and the methods of collection and analysis have changed in time, as well as data accuracy, and thus some subsets of the data require correction of identified biases, both for temperature and salinity, as commented in the earlier papers. Bias correction

follows Reverdin et al. (1994) for ship data before the 1990s (based on comparison with surface data from Nansen casts). There is nonetheless the possibility of seasonal stratification between the data collected below the surface (for example, from TSG data, Argo floats and CTD/Nansen casts) and the surface data (for example from drifters, or in earlier times from bucket sampling). In particular, the change in depth of sampling in time can result in artificial time variability, due to near-surface stratification. In other parts of the world, this has been documented due to surface heating and evaporation or precipitation (Boutin et al., 2016). In regions of sustained wind conditions, and away from shelves with surface freshwater sources (melting sea ice or icebergs, outflows from fjords, river plumes), we expect stratification on the order of or less than 0.01 in practical salinity and 0.1°C in temperature. Sustained wind conditions are commonly melt in all seasons near 55-65°N, and probably a little less so in particular in spring/summer further south and north.

We construct annual time series of deviations from the seasonal cycle in boxes for four different seasons, using a similar methodology as in Friedman et al. (2017) and applied from 1896–2015 for both T and S, over smaller spatial bins. This new box-product will be referred to as BINS. The boundaries of the 34 boxes were redesigned to better fit with oceanic fronts and main currents than in Friedman et al. (2017). One box in the central Labrador Sea corresponds to what is used by Yashayaev and Loder (2016) for creating a combined Labrador Sea time series, and box boundaries in the eastern subpolar North Atlantic are chosen as in Holliday et al. (2015). Except for one box off southwest Greenland and one box on the southern part of the Grand Banks, the boxes do not overlap with the shelves and with areas of large seasonal sea ice cover. In the northeastern SPG, the resulting boxes provide a slightly lower spatial resolution than what was used in Reverdin (2010), but for which the different time series were rather well correlated. To give an idea of the resolution achieved, the SSS and SST late-winter climatologies are mapped onto the BINS grid boxes (**Fig. 1**), which shows that the main characteristics of spatial variability are retained by this choice of grid.

The construction of box time series and their error estimates is described in Friedman et al (2017). Here, the main steps are summarized, with specifics and additions related to BINS further described. The SSS and SST climatologies were constructed on a 1°×1° grid with little spatial smoothing and gap-filling, with an adjustment to the mean and up to three sinusoidal harmonics of the year, when data was sufficient. Deviations from this climatology of

individual data are estimated (T being in °C and S reported in the practical salinity scale 1978 (PSS-78; Fofonoff 1985; UNESCO 1981). For T, outliers were removed using a spatially variable threshold (between 3°C and 6°C), inspired by the local standard deviation. We compared this to a fixed 3°C error threshold, which yielded generally similar results albeit with smaller interannual excursions. Largest differences were found in the central Labrador Sea in the 1920s, a period with extremely low data coverage. S outliers were removed similarly, with less influence on the variability. In each box, these individual deviations are then combined for the individual seasons and years, by median-averaging the individual deviations. The seasonal anomalies are 1-2-1 smoothed over successive years (not applied at the start and end years) and then averaged annually from December–November, inversely weighted by error. [The annual period was incorrectly stated as March–February in Friedman et al. (2017)]. Before being combined, the seasonal anomalies were adjusted to a common baseline over the 40-year period from 1956–1995, which was chosen due to data coverage.

The fall season (September–November) is not included in most areas north of 45°N in the western Atlantic, including in the central Labrador and southwest Labrador boxes due to poorly correlated T or S time series in the fall with the other seasons in these regions. In the Norwegian Sea, central Labrador, southwest Labrador, and West Greenland, the winter season (December–February) data are not included before 1947 due to insufficient coverage.

The total number of years with data is shown in **Fig. 2a**. Coverage is below 100 years in the North Irminger Sea and central Labrador Sea grid boxes; all other grid boxes have more than 100 years of coverage. The periods of missing data occur mainly during WWI and WWII, and around 1900 (**Fig. 2b**). Multi-year data gaps are linearly interpolated, and attributed an error, equal to the standard deviation of the whole time series (which is probably an overestimate). A few missing end years (mostly 1896) in some boxes were filled by extending the first non-missing value. As in Friedman et al. (2017), regional average error terms are constructed by RMS averaging the error terms of their constituent individual boxes (which assumes that the errors are uncorrelated in different boxes). The BINS SSS regional average fields are compared with those of Friedman et al. (2017) in Appendix A for the common period 1896-2013, showing differences most commonly within the ±2 standard error ranges.

Figs. **2c**–**2d** present an estimate of the average interannual grid box error as the mean error term divided by the error term during the year when there was no coverage and the error

is maximum (1943, except in the West Greenland box, when it is 1917). This is thus an indication of the ratio of noise relative to signal. This shows for S, the smallest relative errors are in the central SPG. For temperature, the largest relative errors are in the western part of the domain, where intra-annual variability and eddy variability are largest and not adequately sampled in this data set. The normalized area-averaged error for SSS and SST is shown in **Fig. 2e** and mirrors data coverage, with a larger value in the last year (2015) because error is not reduced in that year by the filtering.

Density is computed from the annual binned SST and SSS values using EOS80 (UNESCO, 1981), and referring the annual anomalies to the March average T and S values. Because SST is lower in winter and density becomes less sensitive to temperature variations when temperature decreases, this choice slightly decreases the dependency of density on SST compared to using an annual average reference. What motivated this choice was that we wanted to have an estimate typical of winter conditions. Density errors are calculated from the errors in interannual SSS and SST, assuming that they are uncorrelated.

2.2 Seasonal dependency of the interannual variability

The combination of the four seasonal time series to create an annual time series in each bin implicitly assumes that the four seasons portray comparable signals/patterns of interannual variability. This is particularly important if one season is missing, as is the case for the winter season in some boxes for the first part of the records. To get a sense of how well this holds for individual seasons, we combine the time series of individual bins into seven main multi-box regions (**Fig. 3a**): the central Labrador Sea / West Greenland Shelf (3 grid boxes), the southern Nordic seas (3 grid boxes), the north-east part of the SPG (5 grid boxes), the central SPG (6 grid boxes), a large central area of the intergyre gyre south of 50°N (8 boxes), an eastern Atlantic region (two boxes), and four boxes in the southwestern region around the Grand Banks. The regions are chosen for similar source data and coherent variability signals (except for the last one). For example, the two Norwegian Sea grid boxes contain mostly Norwegian and Swedish surface data before WWII, and afterwards much hydrographic data. The central Labrador Sea / West Greenland region contains hydrographic data from the Ice Patrol and other cruises in the southwest, and a mix of hydrographic data and merchant ship data until 1960 in the shelf box close to Greenland. South of 50°N, many different countries contributed to the surface data before WWII, and the French data sources are a large contribution since the mid-1970s. Three grid boxes are not included in any of these regions

due to low seasonal correlations in S and more varied data sources: the southwest Labrador Sea, south Greenland/ southwest Irminger Sea, and the western area north of the Gulf Stream.

The seasonal correlations are summarized in **Table 3**. For the central SPG and north-east SG regions (**Fig. 3b-c**, rows 1 and 2), the time series are most complete in all seasons. There, the different seasons present rather similar variability, both for T and S, although for T, the summer (June-August season) deviates at times from the other seasons and presents more variance in the central SPG. Furthermore, after 1960, S anomalies in Sep-Nov in central SPG tend to be below the ones in the other seasons, and the opposite before. This might either be the sign of a real long-term seasonal change, or the effect of changes in the dataset around 1960 (transition in this region from bucket collection to other means of collecting water, for example).

In the southern Nordic seas (**Fig. 3b-c**, bottom row), the different seasons also present rather similar variability, at least since 1950. Before that, this remains largely true for SSS, albeit with more variability in summer. For SST, this also holds before WWI, however between 1920 and 1940, the different seasons present non-coherent variability. We suspect that this might be caused by inhomogeneities in the dataset, either spatially, with a southward shift of the sampling in late autumn to early spring in this period, or due to lower data quality and larger random as well as systematic errors affecting the results and lowering correlations, such as the documented use of engine room or intake temperature data from Norwegian vessels. For this reason and the limited winter sampling coverage, we chose not to incorporate the winter season in the annual averages before 1947 in the two Norwegian Sea boxes.

In the central Labrador Sea / West Greenland Shelf region (**Fig. 3b-c**, row 3), on the other hand, while there is some similarity between the different seasons in SST (less so for the summer season), there is much less coherence in SSS, even during the last 50 years of more regular sampling. Notice in particular the large differences between the spring and autumn seasons, which originate from the two boxes in central Labrador Sea and along west Greenland. These differences (and the much larger RMS variability in autumn) justify the choice of not including the autumn season when combining the different seasons to create the BINS time series in this region. In the Grand Banks region, however, we have less correlation with winter season but with similar variance (**Table 3**), so that we chose to retain the winter season. Elsewhere, the statistics associated with the comparison between the seasonal time

series (**Table 3**) justify the choice to have retained all seasons in order to reduce the sampling uncertainty in the BINS interannual time series.

2.3 Other products

In the next section, we validate BINS with some of the few published compiled hydrographic time series that are available in the SPG, either from repeated seasonal cruises or a compilation of data from different origins: the Rockall Trough time series since 1975 (Holliday et al. 2015); South Icelandic stations (Sevlogsbanki 5 (Sb5) and Stokksnes 5 (St5)) since the 1970s; and the central Labrador Sea since 1941 (Yashayaev and Loder, 2009 2016). Notice also the time series in the Norwegian Sea summarized by (Yashayaev and Seidov 2015). We will also compare BINS to gridded datasets. These include the 1°×1° EN4, version 4.2.1 (Good et al. 2013) fields from 1950-2015, as well as the 1950-2012 1°×1° product of Ishii et al. (2006), version 6.12 (referred to as ISHII). We also compared BINS to the CORA5.0 fields (Cabanes et al. 2013) that have been extended to 1950, but that exclude Nansen cast data, and thus cannot be used for the analysis of salinity before 1980, and have much coarser spatial sampling in temperature than EN4 or ISHII. On the other hand, when sampling is sufficient (mostly, in the northeastern part), results with CORA are comparable to those of ISHII, and will not be discussed further. For long-term surface temperature, we also compare BINS with the HadSST3 monthly product (Kennedy et al. 2011a and 2011b).

Mis en forme : Police :Non Gras

3. Dataset validation

In order to validate the BINS SST and SSS time series, we compare it first to local compilations of high-quality hydrographic data. We also compare the binned time series to different gridded time series to better characterize the properties of BINS. In addition, Appendices A and B show comparisons with Friedman et al (2017) and Reverdin et al. (2018), which are based on different binning (in time and space) of similar data sources.

3.1 BINS and hydrographic time series

The sites of the different time series are presented in Fig. 4a. The Rockall Trough and central Labrador time series fit rather well within some BINS grid boxes, and the South Icelandic time series are just to the north of one BINS grid box. The results based on hydrographic time series often comprise vertical averages, excluding the near-surface (5 to 10 m thick) layer, which is often determined by specifics of performing oceanographic stations and

technological limitations. This averaging depth range was not optimized, but in Rockall Trough and near Iceland, it fits with the expected local maximum winter mixed layer. The time series are then averaged over annual means (after removal of an annual cycle), and smoothed with a 1-2-1 filter over successful years, similar to BINS.

The comparisons of these upper ocean time series with BINS are presented in **Fig. 4b** (with statistics in **Table 4**). At each site BINS is correlated with the vertically integrated time series, although the surface time series tend to have larger amplitudes. In the Labrador Sea, we compared different ranges, with the best correlation for the comparison to 20-50 m, which is within the mixed layer at least 6 months of the year and closely follows variations in seasonal discharge of freshwater and its interannual variations. There is still a correlation when considering the averages over the 20-200 m layer, which is within the winter mixed layer most years. The thicker layer time series however present less extreme amplitudes than the shallower ones. In the Rockall Trough, we also plot the surface time series based only on winter data, which is also correlated with the vertically integrated time series. The largest difference is found for South Iceland, with a low BINS S in the late 1980s that is not found in the hydrographic time series. However, there is also a difference in latitude between the BINS box and the site further north of the hydrographic time series.

Overall, we find that BINS surface time series are significantly correlated at 0-lag with upper ocean time series from hydrography, both for T and S (correlation coefficients are in the range 0.73 to 0.88 for the different time series of salinity or temperature), and exhibit very comparable RMS variability, as described in **Table 4a**. The analysis of lag correlation (summarized in **Table 4b**) indicates that the annually averaged BINS T and S precede by one year the vertically integrated upper ocean analysis both at the Rockall and Southern Iceland sites (for SST, the one-year lag correlation coefficients are respectively 0.80 and 0.89; and for SSS, respectively 0.77 and 0.88), as expected from winter ventilation of the upper ocean and time series analysis of the 20 to 40-year long weather-ship time series in the SPG (Reverdin et al., 1997). Interestingly, this is not the case for the Norwegian Sea time series.

3.2. BINS and gridded objective analyses

We compare the BINS time series with other gridded products of hydrographic data at their near-surface level that we average over the bins of the BINS analysis and filter in time to

correspond to the same time resolution (1-2-1 running average over successive years December to November).

Because of the differences in sampling and mapping, the comparison in individual boxes with BINS often presents a large scatter. We will thus average the time series over the same domains as for the seasonal time series (**Fig. 3a)**. [One difference is that the central SPG and north-east SPG are combined into one SPG region]. Except in the Labrador Sea or near the Grand Banks, the time series of the three products (**Fig. 5**) are well correlated (**Table 5**). Correlation coefficients are larger for SST than for SSS, and for SSS, are usually larger between ISHII and BINS than between EN4 and BINS. In the Labrador Sea (row 2), there are differences both for SST and for SSS. For SSS, the large positive peak in BINS originates from a data gap in one of the three bins that are averaged in the regional time series. On the other hand, the altogether weaker SST-signal in ISHII in this box, something also found, but to a lesser domain in the other regions might result from the objective mapping method used or from more strident tests on outliers (Ishii and Kimoto, 2009). We also notice in the southern Labrador Sea for the last 20 years smaller positive SSS (and SST) anomalies in BINS than in the other two products, possibly a result of the thermosalinograph (TSG) data incorporated in BINS and not in the other products (Reverdin et al. 2018). On the other hand, near the Grand Banks during the last 20 years, T is higher in BINS than in the other products. Elsewhere, the major variability is portrayed with similar amplitude and roughly at the same time in the different products. For example in the SPG or southern Nordic seas, the salinity time series portray the events sometimes described as great salinity anomalies (Belkin et al. 1998), the largest one in the central SPG happening in the early 1970s and referred to as the Great Salinity Anomaly (Dickson et al. 1988). The figures also suggest larger differences between the different products in the 1990s, in particular in the intergyre region.

The high overall agreement might result from a share of common hydrographic data incorporated in the different products. However, there is also a large portion of specific data that are not in all the products (T from mechanical or expandable bathythermographs are used in EN4 and ISHII, but not in BINS, and different sets of surface data are in BINS, but not used in EN4 and ISHII). There are also large methodological differences between the different products, both in data selection and validation and in mapping techniques and how seasonal anomalies are grouped or not. These comparisons, as well as the comparisons presented in 3.1, thus reinforce our assessment of usually small sampling errors for the BINS

product during that period, and that the product captures the major events of interannual variability at the grid box scale since 1950.

3.3 BINS and HadSST3

SST data in BINS and HadSST3 originate from different data streams. Although some data are certainly incorporated in both products, the overlap is probably rather minimal, as discussed in Reverdin et al. (1994). In particular, the data before 1920 north of 58°N in BINS originate mostly from Danish and Norwegian sources which are not incorporated in HadSST3. In the last two decades, data from TSGs on merchant vessels or from profiling floats are a large contribution to BINS but are not included in HadSST3, whereas the drifter SST data used in HadSST3 are mostly not in BINS. Even in the post-WWII to mid-1970s period, when ocean weather ships contribute both to HadSST3 and BINS, temperature data in the two products often originate from different sensors. BINS and HadSST3 also account for biases and uncertainties differently. HadSST3 does not apply smoothing or interpolation between neighboring grid cells, facilitating more direct regional comparisons.

We examine the monthly HadSST3 anomalies, version 3.1.1.0 (median realization). We area-average the HadSST3 grid boxes into similar sub-regions similar to those shown in **Fig. 3a**. The southwest Labrador Sea BINS grid box is also examined, which closely corresponds to one HadSST3 grid box. Generally however, the large 5°×5° HadSST3 grid boxes do not align directly with BINS grid boxes; the specific grid boxes are shown in **Fig. C1**. We average the HadSST3 anomalies into 3-month seasons (skipping over missing values). The seasonal anomalies are 1-2-1 smoothed (excluding endpoints) over successive years, and the seasonal anomalies are averaged into Dec–Nov annual means (again skipping over missing values). The winter season is not included in the southern Nordic, central Labrador, or southwest Labrador annual means before 1947 to match the BINS coverage. Likewise, the fall season is excluded from the central Labrador and southwest Labrador sub-regions. In general, we find that the BINS data are more seasonally coherent at low frequencies before WWII (not shown).

The HadSST3 and BINS regional time series are compared in **Fig. 6**, with the respective 1896-2015 and 1950-2015 correlations listed in **Table 6**. Overall, the products are very highly correlated since 1950. In the first half of the record, the largest differences between the products are generally found during periods of data gaps and linear interpolation near the end

of WWI and WWII, when BINS error estimates are also large. Among regions, the agreement is less good in the southern Nordic seas (**Fig. 6b**) before the 1940s; we also find that the HadSST3 SST in different seasons are also much less correlated there before the 1940s (not shown). To some extent, the early differences come from the most poorly sampled northwestern part of this region, but also from Danish and Norwegian data incorporated in this product and not in HadSST3. Furthermore, in the 1930s, the differences might also originate from Norwegian data interpreted in BINS as recorded continuously on a water line inside the ship, and thus attributed a large positive bias. The central Labrador Sea / West Greenland shelf (**Fig. 6a**) also shows less agreement before 1930. The products correspond very well after the 1930s, although HadSST3 has less of a cooling in the 1980s. The two SPG regions are both very strongly correlated over the length of the record (**Fig. 6c,d**). The Southwest Labrador Sea grid box (**Fig. 6e**) generally corresponds well after the gap periods in the early 1900s. In the regions further south (**Fig. 6f,g,h**), the two time series usually show a good agreement, and are most of the time compatible with the BINS error estimate. The largest differences are near the Grand Banks (**Fig. 6g**) before 1920, maybe as a result of different seasonal coverage, but also in the 1990s, when BINS tends to present lower values. Large differences are also found in the intergyre region (**Fig. 6f**) in the early 1950s. This is the only place where the error estimate does not explain the difference between the curves, maybe because, for a few years in the early 1950s, data in BINS for some of these boxes were only available in the summer season, and thus are less representative of annual averages.

4. Characteristics of T, S and surface density RMS variability

In this section, we examine the linear trends and the RMS variability of the linearly detrended interannual time series. The 1896–2015 BINS least-squares linear trends are shown in **Fig. 7a–7c**. We construct error estimates for the linear trends by block resampling using the time series error estimates (resampled every 2 years to account for the 1-2-1 filter). The SSS trend (**Fig. 7a**) shows a negative trend throughout most of the domain, except for the southwest and part of the eastern boundary; the pattern is consistent with the 1896–2013 trend shown in Friedman et al. (2017). The SST trend (**Fig. 7b**) shows strong warming in the Gulf Stream, and to a lesser extent along the eastern margin. There is cooling in West Greenland / Central Labrador, and near 52-55°N/30°W. However, error bars due to higher frequency variability or sampling are often large. Both the increase in T and the decrease in S contribute to an overall decrease in surface density (**Fig. 7c**).

The 1896–2015 HadSST3 trend is shown in (**Fig. 7d**). The trends are calculated for grid boxes with 5 or fewer years with missing data. [Note that specific seasons are not removed from the spatial grid boxes as for the HadSST3 indices discussed above]. Error terms for HadSST3 are computed by multiplying the white-noise standard error of the slope by a factor of sqrt(2) on account of the 1-2-1 filter. Like BINS, the HadSST3 trend shows significant warming in the Gulf Stream extension, and also along the eastern margin. HadSST3 also shows a region of cooling in the central SPG from 50-60°N, referred to as the 'warming hole' (Drijfhout et al., 2012), which is not well reproduced in BINS. Checking this particular difference suggests that data in the pre-1917 period originated mostly from one group of data from the ICES archive, often reported at set longitudes, and for which we applied an overall correction of -0.2°C on temperature, but with little comparison data to define it (Reverdin et al. 1994). Notice however that nonetheless, for the area average of the central SPG (**Fig. 6c**), the difference between BINS and HADSST3 is not significant for that period.

The detrended RMS standard deviation of the interannual T and S time series (**Figs. 8a** and **8b)** correspond to the expectation that large-scale temperature RMS standard deviation diminishes from west to east; whereas the salinity RMS values are largest in the west, then near 50°N with a decrease from west to east and from the southern to the northern SPG. This pattern is indicative on one hand of larger air-sea and hydrographical/heat forcings in the western SPG, and on the other hand of the presence during some winters of a rather shallow mixed layer in the western SPG (even in the Labrador Sea, where deep convection also often takes place (Yashayaev and Loder 2016)). In mid-ocean near 45-50°N, there might also be a contribution of displacements of hydrographical fronts.

Surface density standard deviation (**Fig. 8c**) diminishes both from west to east and south to north. The relative contributions of T and S to detrended density variability are shown in **Fig. 8d**. The contributions of SST and SSS to surface density variability are opposite in most places (the effects of these variables on density variability tend to compensate each other, as there is a positive correlation between SST and SSS variability on intra-annual to centennial time scale in a large part of this domain). This also might explain a pattern of density rms variability which presents weaker zonal gradients in the west than what is found in SST or SSS.

Interestingly, in the westernmost part of the domain, except near the Gulf Stream, the salinity contribution to density variability dominates over the temperature contribution (**Fig. 8d**), whereas in the eastern and northern parts of the SPG, temperature dominates. Thus in the first region, density tends to be positively correlated with SSS and SST, whereas in the other region it is negatively correlated with SSS and SST. In the northern part, this was not expected, as temperature is lower than further south, which would contribute to diminish the contribution of temperature variations to density variations. However, this is counteracted by a lower ratio of SSS over SST variability in the northern regions. In general, the SSS contribution is slightly larger for the non-detrended data (not shown), due to its relatively larger centennial trends in this part of the ocean (discussed in part 2 of this study).

Next, we show the total RMS variability contained at multidecadal frequencies. We low-pass filter the time series using a spectral filter with half power at a period close to 15 years. **Fig. 9** shows the percentage of the detrended interannual variance explained by the low-pass filtered time series. The low-pass filtered variance percentage of SSS (**Fig. 9a**) and SST (**Fig. 9b**) are less correlated spatially than overall RMS variability. For S, what striking is the lower percentages in the western part (as low as 40%) than in the eastern part of the domain (larger than 60%), whereas for T, it is mostly the northern part which has a larger percentage of variance in the multidecadal frequencies, with percentage of detrended variance reaching up to 80%. The low percentages of multidecadal variability for SSS in the northwest is consistent with the penetration in the ocean interior of short-lived freshwater pulses originating from the western shelves and the rim of the Labrador Sea (examples of recent short-term SSS variability in this region are presented in Tesdal et al. (2018)).

The low-pass filtered time series explain comparatively less density variance, particularly in the western SPG region (**Fig. 9c**, values usually less than 50%), which suggests coordinated SSS and SST variations at lower frequencies with thus partial compensation on their contribution to density changes. Indeed, the low-pass filtered correlation of SSS and SST (**Fig. 9d**) is positive in most grid boxes, with large magnitudes in the central SPG, around Greenland and in the south-west north of the Gulf Stream as well as around the Grand Banks. Notice however, less significant correlation in other areas of the western, southern, and eastern parts of the domain close to Europe. The relationship between SSS and SST variations will be examined further in part 2 of this study.

5. Discussion and Conclusions

We have constructed a spatially averaged binned product BINS of interannual T, S, and surface density for 1896-2015 north of 40°N, similar to what was done for S in Friedman et al. (2017), but at a higher spatial resolution. We deliberately excluded the shelf regions (except for one box along southwestern Greenland and one box on the southern Grand Banks), which have different dynamics, variability and sampling issues (Reverdin et al. 2018). We have shown that in recent decades, the BINS interannual variability of SST and SSS is coherent with published time series of upper ocean heat and salinity based on high quality hydrographic data. It provides also rather similar results to an analysis of SST and SSS variability along two ship tracks since 1993, based on the same in situ data set for its main part (Reverdin et al. 2018).

Furthermore, BINS compares well with two gridded products of hydrographic data (EN4 from 1950-2015 and ISHII from 1950-2012), which use different data selection and mapping methods. This suggests that in this region, the processing and spatial averaging of the individual data in the spatial bins done in BINS is usually sufficient to portray the interannual variability of SST on the grid box scale, at least since 1950, and that during this period, the data set used does not contain large biases compared to these other data sets. On interannual time scales, BINS SST also fits largely with what is portrayed in HadSST3, except in the pre-WWII period for the Labrador Sea and the Norwegian Sea. Insufficient coverage or erroneous data, either in HadSST3 or BINS, might be causing these discrepancies. Thus, more care should be taken when interpreting variability in BINS before 1950.

The product combines time series of interannual variability in different seasons. This was required to reduce sampling errors to a reasonable level (in particular pre-1950), but this might also mask or alias important differences between the seasons. For example in the northern part of the subtropical gyre (thus, south of the region we investigate) a strong seasonal modulation in the recent trends towards increasing SSS has been documented, associated to important seasonal changes of the hydrological cycle (Yu et al. 2017). In the Labrador Sea, too, on an interannual time scale, large summer-time anomalies of SSS have been documented in the recent decade (2008 and 2012) that don't have a clear counterpart in other seasons (Tesdal et al. 2018). Such events, and also the seasonality of atmospheric forcing, seasonal thermal stratification and mixing, could induce large interannual differences between the anomalies in different seasons. Indeed, **Fig. 3** suggests that the differences

between seasons are large in the central Labrador Sea, but for the other open-ocean regions, different seasons present more similar decadal and longer frequencies. Possibly, the insufficient accuracy of binned time series and the change in collection and measurement techniques could preclude further investigation of the seasonality of the decadal or longer variability with this dataset. For example, in boxes of the southern SPG, we noticed SSS differences in the autumn around 1960 that could be a data artifact. Differences with HadSST3 taking place around 2000 could also be data-related (either in HadSST3 with drifting buoy data becoming more numerous, or in BINS with the Argo data afterwards). Nonetheless, a more homogeneous 25-year dataset in the central SPG resolving seasonal variability along well sampled ship lines (Reverdin et al. 2018) also pointed out a similarity in interannual variability for the different seasons, although interannual RMS variability tended to be larger there too in early summer for SST.

This aside, the seasonal inhomogeneity in data coverage, with less sampling in winter in particular before 1955, should have little impact on the reconstructed interannual time series. We however warn that this might not be a wise substitute for the seasonal time series, in particular for studies of winter conditions the central Labrador Sea or water mass formation, but then obviously the available binned data present more gaps, and the original seasonal time series have larger error bars.

Density is a non-linear function of temperature and salinity, with the character of non-linearity increasing at low T. We deliberately emphasized winter characteristics when estimating density from temperature and salinity time series, but use the BINS time series which combine SST and SSS anomalies in the four (or three) different seasons. We noticed that in some regions interannual variability has smaller amplitude in winter (Table 3 summarizing **Fig. 3**), thus this choice might result in an overestimation of winter-time interannual density variability. It would be also misleading to use these density interannual time series as a proxy for water density in other seasons.

We do not have a direct comparison of SSS BINS time series prior to the 1950s. We estimate that they provide reasonable estimates for this early period mostly based on comparisons of the SST BINS time series with HadSST3 binned data. However, before 1950 in the Nordic Seas and Labrador Sea, the binned data have larger uncertainties and winter seasons are poorly (or not at all) sampled. Thus, clearly our estimates for SSS are also less certain in these

regions prior to 1950. A possibility for further validation would be to compare with independent proxy-based estimates of surface variability. Few such proxy time series have been derived in this region resolving the multidecadal variability in SST and even less in SSS, and their uncertainties are probably too large for such a validation (Hall et al. (2010), for a site in the Iceland Basin (Gardar drift); Richter et al. (2009), for a site in the northern Rockall Trough area (Feni drift); Moffa-Sanchez et al. (2017) and references therein as well as Thornhalley et al. (2018), for a site in the southeastern Labrador Sea).

BINS has sufficient spatial resolution to resolve large-scale patterns of RMS standard deviation of the variability in SST, SSS and density (this was done after removing linear trends). RMS variability in SSS and to a lesser degree in SST increases towards the west, with the contribution to density of SSS variability dominating that of SST in the west, whereas in the eastern and northern parts of the SPG, temperature typically dominates the variability. The portion of variability in multidecadal frequencies (here with a cut-off at 15 years) could be ascertained. It is found both for SST and SSS to be a very large part of interannual variability, with lesser relative contribution in the west for SSS and in the southern part for SST. For density, it is less prominent, except in the northeastern part of the SPG, where SST variability contribution dominates and is very much at multidecadal frequencies (also see Holliday et al. 2015). The higher frequencies are more influenced by the sampling uncertainties, in particular the effect of interpolation and interannual smoothing during data gaps in either specific seasons or all seasons (near WWII, for example). Thus this analysis of the relative importance of multidecadal variability is more dataset-dependent than the RMS variability

Acknowledgements

This is a contribution to the French SSS observation service, which is supported by French agencies INSU/CNRS, IRD, CNES and IPEV, as well as from SOERE CTDO2. The Rockall Trough time series were provided with support from the UK Natural Environment Research Council (Extended Ellett Line program, National Capability). The station time series south of Iceland were provided with Icelandic support. The annual oceanographic monitoring of the Labrador Sea was initiated as a Canadian contribution to the World Ocean Circulation Experiment in 1990 and is presently conducted as a core component of the Atlantic Zone Off-Shelf Monitoring Program (AZOMP) run by the Bedford Institute of Oceanography of Fisheries and Oceans Canada. The International Argo Program is part of the Global Ocean

Observing System (Argo, 2000). Argo data are available from the Coriolis Global Data Center, Institut français de recherche pour l'exploitation de la mer (Ifremer). The HadSST3 and EN4 data were provided by the Met Office Hadley Center, and the ISHII data were provided by the NCAR Research Data Archive. A.R.F. was supported by SOERE CTDO2 and the ERC funded project TITAN (EC-320691). L.C. acknowledges support from the Swedish National Space Board (SNSB; Dnr 133/17). Comments by two reviewers and by Semyon Grodsky were appreciated.

References

Alory G, Delcroix T, Téchiné P, Diverrès D, Varillon D, Cravatte S, … Roubaud F (2015) The French contribution to the voluntary observing ships network of sea surface salinity. Deep Sea Res I 105: 1–18. http://doi.org/10.1016/j.dsr.2015.08.005

Argo (2000) Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE. http://doi.org/10.17882/42182

Belkin I M, Levitus S, Antonov J, Malmberg S.-A (1998) "Great Salinity Anomalies" in the North Atlantic. Prog Oceanogr 41(1): 1–68. http://doi.org/10.1016/S0079-6611(98)00015-9

Böning C W, Behrens E, Biastoch A, Getzlaff K, Bamber J L (2016) Emerging Impact of Greenland Meltwater on Deepwater Formation in the North Atlantic Ocean. Nature Geoscience. https://doi.org/10.1038/ngeo2740.

Boutin J, Chao T, Asher W, Delcroix T, Drucker R, Drushka K, Kolodziejczyk N, Lee T, Reul N, Reverdin G, Schanze J, Soloviev A, Yu L, Anderson J, Brucker L, Dinnat E, Santos-Garcia A, Jones W, Maes C, Meissner T, Tang W, Vinogradova N, Ward B (2016) Satellite and In Situ Salinity : Understanding Near-Surface Stratification and Sub-footprint Variability. Bull. Amer. Meteor. Soc., 97, 1391–1407, doi: 10.1175/BAMS-D-15-00032.1.

Buckley M W, Marshall J (2016) Observations, Inferences, and Mechanisms of the Atlantic Meridional Overturning Circulation: A Review. Reviews of Geophysics 54 (1): 2015RG000493. https://doi.org/10.1002/2015RG000493.

Cabanes C, Grouazel A, von Schuckmann K, Hamon M, Turpin V, Coatanoan C, Paris F, et al (2013) The CORA Dataset: Validation and Diagnostics of in-Situ Ocean Temperature and Salinity Measurements. Ocean Sci 9 (1): 1–18. https://doi.org/10.5194/os-9-1-2013.

Davis R E, Sherman J T, Dufour J (2001) Profiling ALACEs and other advances in autonomous subsurface floats. J Atmos Ocean Tech 18: 982-993.

Dickson R R, Meincke J, Malmberg M S-A, Lee A J (1988) The ''Great Salinity Anomaly'' in the northern North Atlantic 1968 – 1982. Prog Oceanogr 20: 103– 151.

Drijfhout S, van Oldenborgh G J, Cimatoribus A (2012) Is a Decline of AMOC Causing the Warming Hole above the North Atlantic in observed and modeled warming patterns?" J Clim 25: 8373–79. https://doi.org/10.1175/JCLI-D-12-00490.1.

Ebisuzaki W (1997) A method to estimate the statistical significance of a correlation when the data are serially correlated. J Clim 10(9): 2147–2153.

Fofonoff N P (1985) Physical properties of seawater: A new salinity scale and equation of state for seawater. J Geophys Res 90(C2): 3332–3342. https://doi.org/10.1029/JC090iC02p03332.

Frankignoul C, Deshayes J, Curry R (2009) The Role of Salinity in the Decadal Variability of the North Atlantic Meridional Overturning Circulation. Climate Dynamics 33(6): 777–93. https://doi.org/10.1007/s00382-008-0523-2.

Friedman A R, Reverdin G, Khodri M, Gastineau G (2017) A new record of Atlantic sea surface salinity from 1896 to 2013 reveals the signatures of climate variability and long-term trends. Geophys Res Lett 44: 1866–1876. DOI:10.1002/2017GL072582.

Good S-A, Martin M J, Rayner N A (2013) EN4: Quality Controlled Ocean Temperature and Salinity Profiles and Monthly Objective Analyses with Uncertainty Estimates. J Geophys Res 118(12): 6704–16. https://doi.org/10.1002/2013JC009067.

Hall I R, Boessenkool K P, Barker S, McCave I N, Elderfield H (2010) Surface and deep ocean coupling in the subpolar North Atlantic during the last 230 years. Paleoceanogr 25. doi:10.1029/2009PA001886.

Holliday N P, Cunningham S A, Johnson C, Gary S F, Griffiths C, Read J F, Sherwin T (2015) Multidecadal variability of potential temperature, salinity, and transport in the eastern subpolar North Atlantic. J Geophys Res 120. doi:10.1002/2015JC010762.

Hughes, S. L., Holliday, N. P., Gaillard, F., and the ICES Working Group on Oceanic Hydrography, 2012. Variability in the ICES/NAFO region between 1950 and 2009: observations from the ICES Report on Ocean Climate. – ICES Journal of Marine Science, doi:10.1093/icesjms/fss044.Icelandic hydrographic surveys …

Ishii M, Kimoto M, Sakamoto K, Iwasaki S I (2006) Steric sea level changes estimated from historical ocean subsurface temperature and salinity analyses. J Oceanography 62(2): 155-170.

Kennedy J J, Rayner N A, Smith R O, Parker D E, Saunby M (2011) Reassessing Biases and Other Uncertainties in Sea Surface Temperature Observations Measured in Situ since 1850: 1. Measurement and Sampling Uncertainties. J Geophys Res 116. doi:201110.1029/2010JD015218.

Kennedy J J, Rayner N A, Smith R O, Saunby M, Parker D E (2011) Reassessing biases and other uncertainties in sea-surface temperature observations since 1850: 2. Biases and homogenisation. J Geophys Res 116 D14104. doi:10.1029/2010JD015220

Lavender K L, Davis R E, Owens W B (2000) Mid-depth recirculation observed on the interior Labrador and Irminger seas by direct velocity measurements. Nature 607: 66- 69.

Lozier M S (2012) Overturning in the North Atlantic. Ann Rev of Marine Sci 4(1): 291–315. http://doi.org/10.1146/annurev-marine-120710-100740

Moffa-Sanchez P, Hall I R (2017) North Atlantic variability and its links to European climate over the last 3000 years. Nature comm. 8:1726, 555. doi:10.1038/s41467-017-01884-8.

Piron A, Thierry V, Mercier H, Caniaux G (2017) Gyre-scale deep convection in the subpolar North Atlantic Ocean during winter 2014–2015. Geophys Res Lett 44: 1439–1447. doi:10.1002/2016GL071895.

Polyakov I V, Bhatt U S, Simmons H L, Walsh D, Walsh J E, Zhang X (2005) Multidecadal Variability of North Atlantic Temperature and Salinity during the Twentieth Century. J Clim 18(21): 4562–4581. http://doi.org/10.1175/JCLI3548.1

Rahmstorf S, Box J E, Feulner G, Mann M E, Robinson A, Rutherford S, Schaffernicht E J (2015) Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. Nature Climate Change 5(5): 475–480. http://doi.org/10.1038/nclimate2554

Reverdin G (2010) North Atlantic Subpolar Gyre Surface Variability (1895–2009). J Clim 23(17): 4571–4584. http://doi.org/10.1175/2010JCLI3493.1

Reverdin G, Cayan D, Dooley H, Ellett D, Levitus S, du Penhoat Y, Dessier A (1994) Surface Salinity of the North-Atlantic - Can We Reconstruct Its Fluctuations Over the Last 100 Years. Prog Oceanogr 33(4): 303–346. http://doi.org/10.1016/0079-6611(94)90021-3

Reverdin G, Cayan D, Kushnir Y (1997) Decadal variability of hydrography in the upper northern North Atlantic 1948-1990. J Geophys Res 102: 8505-8531.

Reverdin G, Alory G, Diverres D, Bringas F, Goni G, Heilmann L, Chafik L, Szekely T, Friedman A R (2018) North Atlantic subpolar gyre along predetermined ship tracks since 1993: a monthly data set of surface temperature, salinity, and density. Earth Syst Sci Data 10: 1403-1415. https://doi.org/10.5194/essd-10-1403-2018.

Rhein M, Kieke D, Hüttl-Kabus S, Roessler A, Mertens C, Meissner R, Klein B, Böning C W, Yashayaev I (2011) Deep water formation, the subpolar gyre, and the meridional overturning circulation in the subpolar North Atlantic. Deep Sea Res II 58: 1819-1832. doi:10.1016/j.dsr2.2010.10.061

Richter T O., Peeters F J C, Weering T C E (2009) Late Holocene (0-2.4 ka BP) surface water temperature and salinity variability, Feni Drift, NE Atlantic Ocean. Quat Sci Rev 28: 1941-1955.

Riser S C, Freeland H J, Roemmich D, Wijffels S, Troisi A, Belbéoch M, Gilbert D, et al. (2016) Fifteen Years of Ocean Observations with the Global Argo Array. Nature Climate Change 6 (2): 145–53. https://doi.org/10.1038/nclimate2872.

Roemmich, D, Johnson G C, Riser S, Davis R E, Gilson J, Owens W B, Garzoli S L, Schmid C, Ignaszewski M (2009) The Argo Program: Observing the global ocean with profiling floats. Oceanography. 22:34-43. 10.5670/oceanog.2009.36

Skliris N, Marsh R, Josey S A, Good S A, Liu C, Allan R P (2014) Salinity changes in the World Ocean since 1950 in relation to changing surface freshwater fluxes. Climate Dynamics 43(3-4): 709–736. http://doi.org/10.1007/s00382-014-2131-7

Tesdal J, Abernathey R P, Goes J I, Gordon A L, Haine T W (2018) Salinity Trends within the Upper Layers of the Subpolar North Atlantic. J Clim 31: 2675–2698. https://doi.org/10.1175/JCLI-D-17-0532.1

Thornhalley D J R, Oppo D W, Ortega P, Robson J I, Brierley C M, Davis R, Hall I R, Moffa-Sanchez P, Rose N L, Spooner P T, Yashayaev I, Keigwin L D (2018) Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. Nature 556. https://doi.org/10.1038/s41586-018-0007-4.

UNESCO (1981) The Practical Salinity Scale 1978 and the International Equation of State of Seawater 1980. UNESCO technical papers in marine science 36: 25pp.

Williams R G, Roussenov V, Lozier M S, Smith D (2015) Mechanisms of heat content and thermocline change in the subtropical and subpolar North Atlantic. J Clim 28: 9803-9813. doi:10.1175/JCLI-D-15-0097.1.

Yashayaev I (2007) Hydrographic changes in the Labrador Sea, 1960-2005. Prog Oceanogr 73: 242-276. doi:10.1016/j.pocean.2007.04.015.

Yashayaev I, Loder J W (2009) Enhanced production of Labrador Sea Water in 2008, Geophys Res Lett 36 L01606. doi:10.1029/2008GL036162.

Yashayaev I, Loder J W (2016) Recurrent replenishment of Labrador Sea water and associated decadal-scale variability. J Geophys Res 121:8095-8114. doi:10.1002/2016JC012046.

Yashayaev I, Loder J W (2017) Further intensification of deep convection in the Labrador Sea in 2016. Geophys Res Lett 44:1429-1438, doi:/10.1002/2016GL071668.

Yashayaev I, Seidov D (2015) The role of the Atlantic water in multidecadal ocean variability in the Nordic and Barents Seas. Prog Oceanogr 132: 68-127.

Yu L, Jin X Liu H (2017) Poleward Shift in Ventilation of the North Atlantic Subtropical Underwater. Geophysical Research Letters 44. doi: https://doi.org/10.1002/2017GL075772

Figures.

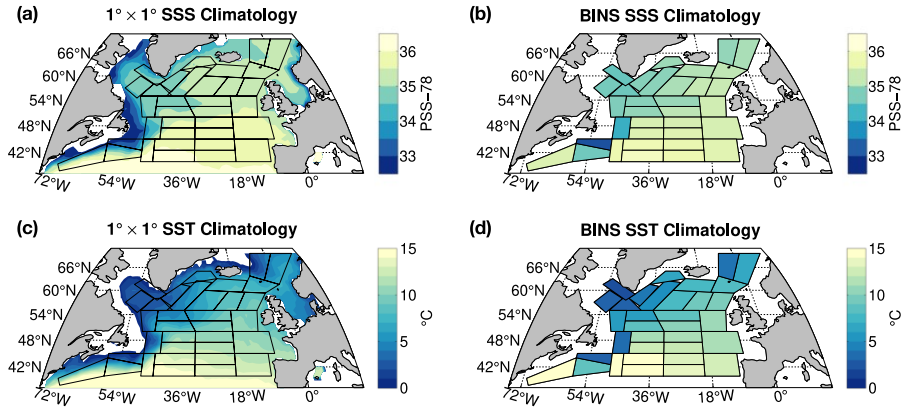


Figure 1. Gridded climatologies () (for 1896-2000). (a) and (c): 1°×1° March climatologies for (a) SSS and (c) SST. (b) and (d): BINS March–May climatologies for (b) SSS and (d) SST.

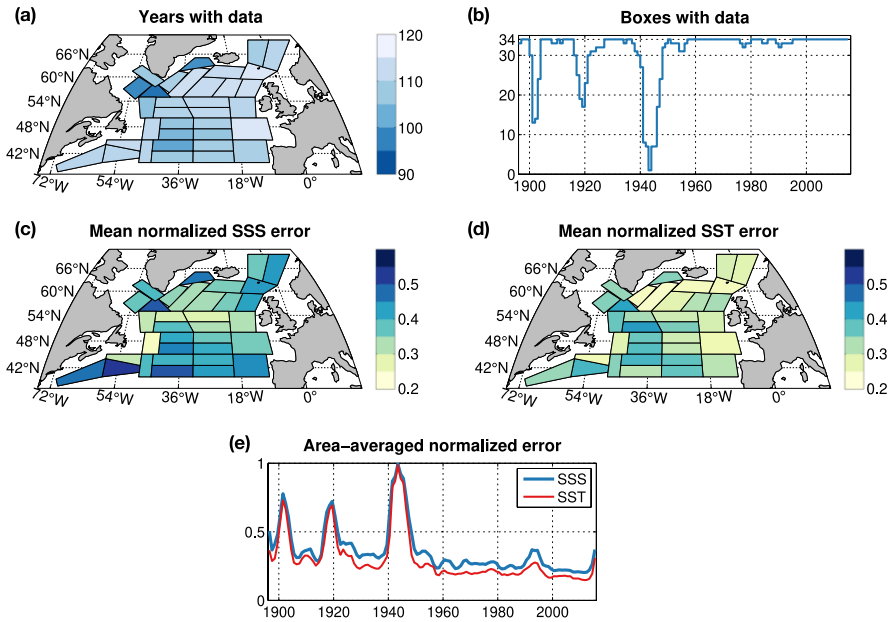


Figure 2. (a) Outline of BINS boxes, with number of years with data from 1896–2015 (out of a maximum of 120). (b) Time series of total BINS boxes with coverage (out of a maximum 34). The data counts in (a) and (b) are from after applying the 1-2-1 filter. (c) Mean annual SSS grid box error, defined as the mean grid box error divided by its maximum error (usually in 1943). (d) Same as (c), for SST. (e) Area-weighted normalized annual grid box error for SSS and SST.

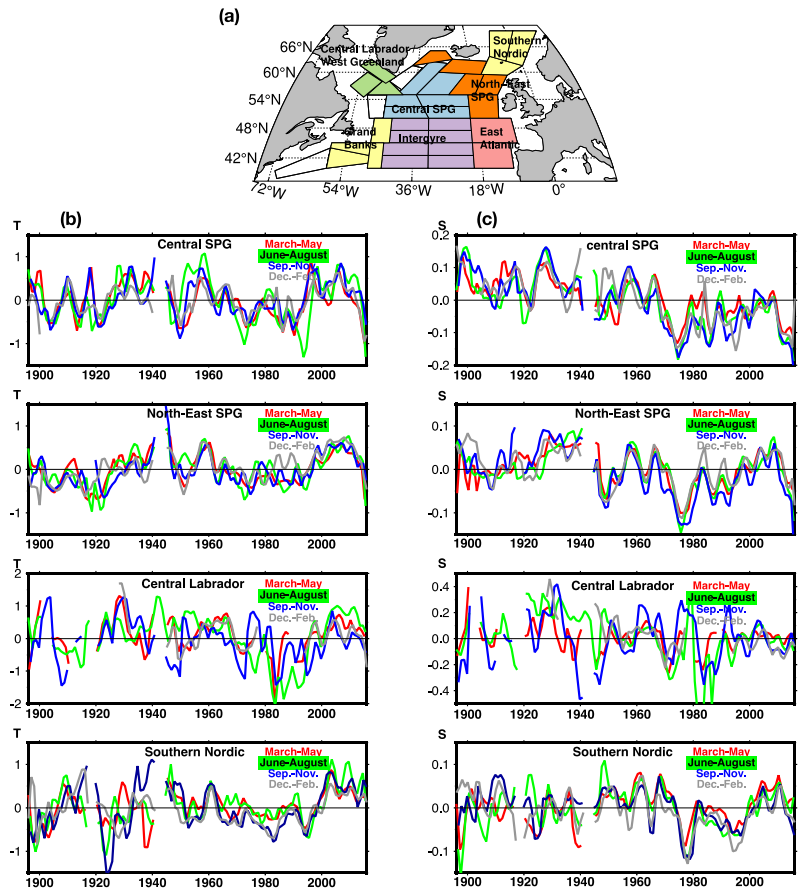


Figure 3. (a) Regional domains described in the text. (b-c) Seasonal time series for (b) T (°C) and (c) S (PSS-78) averaged over four domains portrayed in (a): central SPG, north-east SPG, central Labrador Sea / West Greenland, and the southern Nordic Seas.

Table 3. Seasonal RMS and Pearson correlation coefficient. RMS values are in °C for T, and PSS-78 for S (×1000). Corresponding regions are shown on Fig. 3a.

Region	Length		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Central SPG	112-116 yrs	**T RMS**	0.304	0.407	0.505	0.404	**S RMS**	73	60	72	79
		T corr.					**S corr.**				
		12-2		0.58	0.52	0.57	**12-2**		0.72	0.77	0.78
		3-5			0.55	0.71	**3-5**			0.79	0.82
		6-8				0.62	**6-8**				0.86

Region	Length		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
N-E SPG	113-115 yrs	**T RMS**	0.339	0.338	0.375	0.336	**S RMS**	42	41	44	55
		T corr.					**S corr.**				
		12-2		0.65	0.54	0.60	**12-2**		0.72	0.75	0.73
		3-5			0.76	0.71	**3-5**			0.81	0.78
		6-8				0.74	**6-8**				0.80

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Southern Nordic	**T RMS**	0.358	0.351	0.425	0.524	**S RMS**	41	40	47	41
	T corr.					**S corr.**				
	12-2		0.31	0.40	0.61	**12-2**		0.50	0.43	0.66
	3-5			0.69	0.41	**3-5**			0.66	0.52
	6-8				0.57	**6-8**				0.61

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Central Labrador/	**T RMS**	0.47	0.385	0.482	0.602	**S RMS**	130	95	131	162
West Greenland*	**T corr.**					**S corr.**				
	12-2		0.81	0.20	0.60	**12-2**		0.37	0.50	0.27
	3-5			0.60	0.78	**3-5**			0.56	0.50
	6-8				0.52	**6-8**				0.42

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Grand Banks	**T RMS**	0.950	0.792	0.890	0.990	**S RMS**	148	156	132	150
	T corr.					**S corr.**				
	12-2		0.26	0.33	0.29	**12-2**		0.09	0.24	0.51
	3-5			0.56	0.21	**3-5**			0.49	0.33
	6-8				0.44	**6-8**				0.50

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
Intergyre	**T RMS**	0.49	0.42	0.61	0.49	**S RMS**	87	76	82	100
	T corr.					**S corr.**				
	12-2		0.56	0.19	0.42	**12-2**		0.64	0.65	0.73
	3-5			0.44	0.53	**3-5**			0.51	0.55
	6-8				0.42	**6-8**				0.59

		12-2	3-5	6-8	9-11		12-2	3-5	6-8	9-11
East Atlantic	**T RMS**	0.37	0.37	0.41	0.51	**S RMS**	53	61	55	54
	T corr.					**S corr.**				
	12-2		0.59	0.41	0.57	**12-2**		0.40	0.53	0.56
	3-5			0.57	0.52	**3-5**			0.59	0.34
	6-8				0.38	**6-8**				0.51

* small number of common points, and varying number of boxes; thus not reliable, in particular for S.
Both for Southern Nordic and Central Labrador, very different number of years in winter and other seasons (only on the order of 60 years with winter data in some of the boxes included in the average)

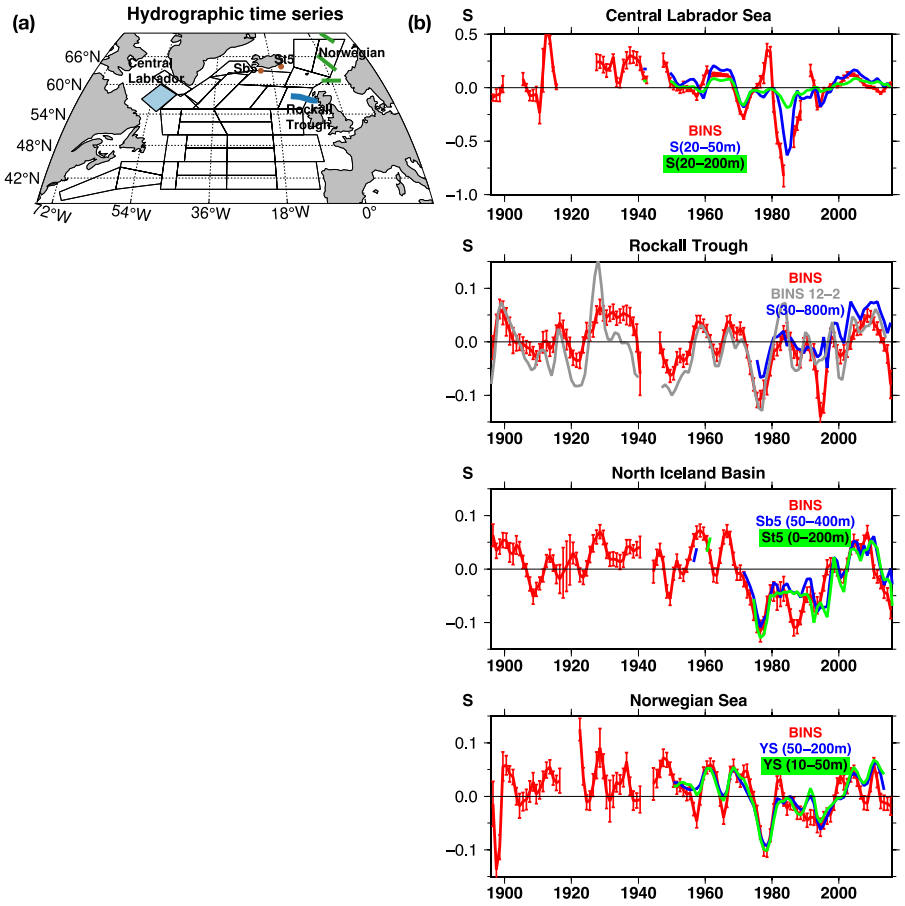


Figure 4. (a) Map showing hydrographic time series locations with BINS grid boxes. (b) Comparison for S of BINS (in red) with hydrographic time series low passed 1-2-1 over successive years; Central Labrador Sea (top; green and blue for different layer averages), Rockall Trough (top middle; the green curve is from BINS but for the winter (Dec-Feb.) season; blue is the vertically averaged (30-800m) S time series), North Iceland Basin (lower middle; two vertically averaged station time series are presented); and eastern Norwegian Sea (bottom; the blue and green time series is a compilation of the offshore data (median averaged) in 3 Norwegian Sea sections presented in Yashayaev and Seidov (2015), for two different depth ranges). One standard deviation rms error bars are added on the BINS series.

Table 4a. Hydrographic sections and BINS: RMS variability for S (PSS-78) and T (°C), and lag-0 correlation coefficient.

			RMS (S)		corr (S)		RMS (T)		corr (T)
Region	**Years S**		**section**	**BINS**		**Years T**	**section**	**BINS**	
S Rockall (30–800m)	37		.040	0.048	0.73	37	0.284	0.350	0.80
S Iceland (50–400m)	47		.050	0.043	0.81	47	0.413	0.391	0.87
Central Labrador (20–50m)	69		.155	0.169	0.76	67	0.715	0.716	0.74
Norwegian Sea (10-50m)	64		.039	0.038	0.82	64	0.31	0.31	0.87

Table 4b. Hydrographic section lag correlation with BINS.

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
South Rockall	**–2**	35	0.57	35	0.56
(30–800m)	**–1**	36	0.70	36	0.65
	0	37	0.73	37	0.80
	+1	37	0.77	37	0.80
(BINS leads)	**+2**	37	0.68	37	0.72

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
South Iceland	**–2**	45	0.54	44	0.59
(50–400m)	**–1**	46	0.68	45	0.74
	0	47	0.81	45	0.87
	+1	47	0.88	45	0.89
(BINS leads)	**+2**	47	0.83	45	0.81

Region	**Lag**	**Years (S)**	**corr. (S)**	**Years (T)**	**corr. (T)**
Eastern Norwegian Sea	**–2**	64	0.48	64	0.70
(10-50m)	**–1**	64	0.72	64	0.81
	0	64	0.82	64	0.87
	+1	64	0.74	64	0.83
(BINS leads)	**+2**	64	0.54	64	0.72

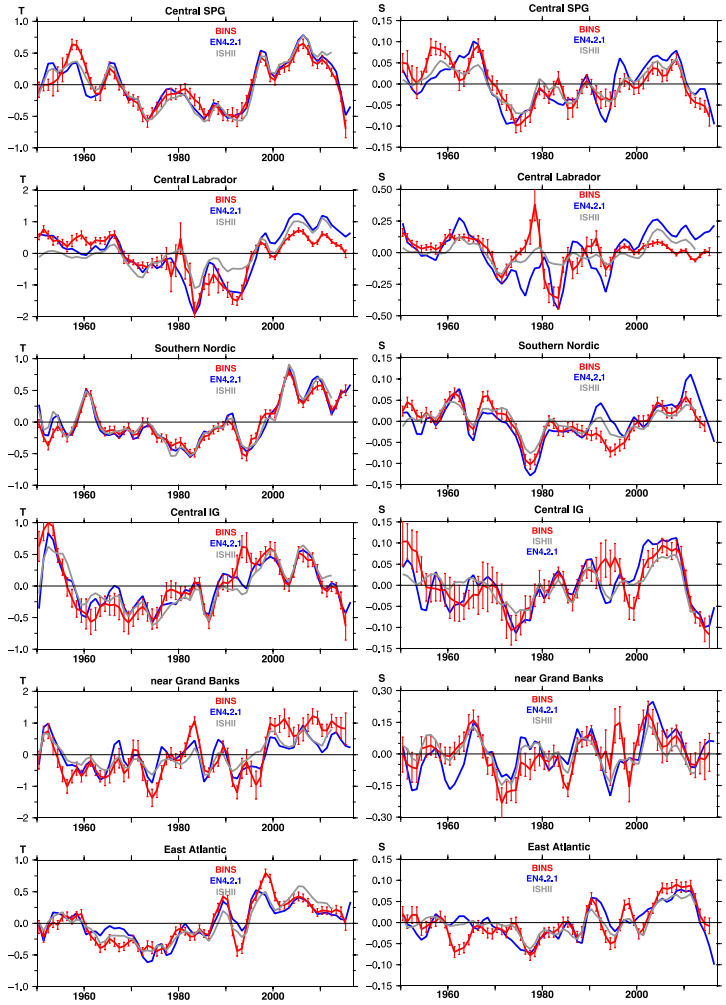


Figure 5. Comparison of regional T (left, °C) and S (right, PSS-78) time series from BINS with EN4 (1950-2015) and ISHII (1950-2012) surface products. The regional domains are the ones presented in Fig. 3. Here, the SPG region includes both the central SPG and the north-east SPG (and the south Greenland/southwest Irminger Sea box). One standard deviation rms error bars are added on the BINS series.

Table 5. RMS variability and correlation with BINS: (a) EN4, (b) ISHII.

(a) EN4 / BINS		**RMS (S) (PSS-78)**		**corr (S)**	**RMS (T) °C**		**corr (T)**
Region	**Time period**	**EN4**	**BINS**		**EN4**	**BINS**	
SPG	1950–2015	0.050	0.046	0.78	0.36	0.34	0.93
Central Labrador/ West Greenland	1950–2015	0.173	0.121	0.49	0.76	0.63	0.87
Central Labrador/ West Greenland	excluding 1976–1979	0.169	0.112	0.67	0.78	0.64	0.87
Southern Nordic S	1950–2015	0.045	0.039	0.75	0.34	0.32	0.96
Intergyre	1950–2012	0.057	0.057	0.78	0.33	0.43	0.85
Grand Banks	1950–2012	0.098	0.114	0.40	0.47	0.69	0.42
East Atlantic	1950–2012	0.037	0.042	0.73	0.26	0.32	0.84

(b) ISHII / BINS		**RMS (S) (PSS-78)**		**corr (S)**	**RMS (T) °C**		**corr (T)**
Region	**Time period**	**ISHII**	**BINS**		**ISHII**	**BINS**	
SPG	1950–2012	0.034	0.045	0.90	0.38	0.34	0.95
Central Labrador/ West Greenland	1950–2012	0.075	0.124	0.55	0.50	0.64	0.74
Central Labrador/ West Greenland	excluding 1976-79	0.077	0.115	0.61	0.52	0.65	0.75
Southern Nordic	1950–2012	0.028	0.040	0.91	0.35	0.32	0.9
Intergyre	1950–2012	0.034	0.045	0.71	0.34	0.34	0.62
Grand Banks	1950–2012	0.068	0.124	0.45	0.41	0.65	0.42
East Atlantic	1950–2012	0.030	0.040	0.73	0.28	0.32	0.65

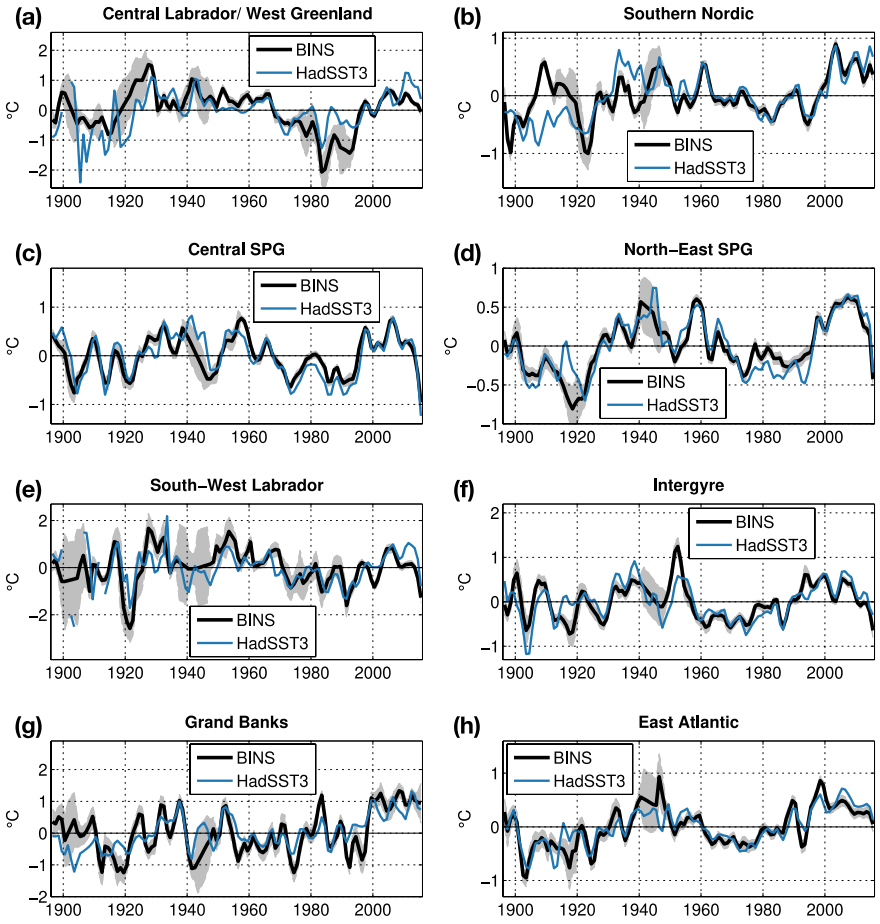


Figure 6. Area-averaged BINS and HadSST3 SST anomalies: (a) central Labrador / West Greenland, (b) southern Nordic, (c) central SPG, (d) north-east SPG, (e) south-west Labrador, (f) intergyre, (g) Grand Banks / Labrador Current, (h) east Atlantic. Shading indicates ±2 BINS error terms. The corresponding regions are shown in **Fig. C1**.

Table 6. HadSST3 correlation with BINS SST

Region	Time period	corr	Time period	corr
Central SPG	1896–2015	0.84	1950–2015	0.95
N-E SPG	1896–2015	0.87	1950–2015	0.93
Central Labrador/ West Greenland	except 1900–01, 1910	0.57	1950–2015	0.72
SW Labrador	except 1900–01, 1904–1905, 1912	0.56	1950–2015	0.71
Southern Nordic	1896–2015	0.65	1950–2015	0.93
Intergyre	1896–2015	0.74	1950–2015	0.84
Grand Banks	1896–2015	0.67	1950–2015	0.82
East Atlantic	1896–2015	0.82	1950–2015	0.89

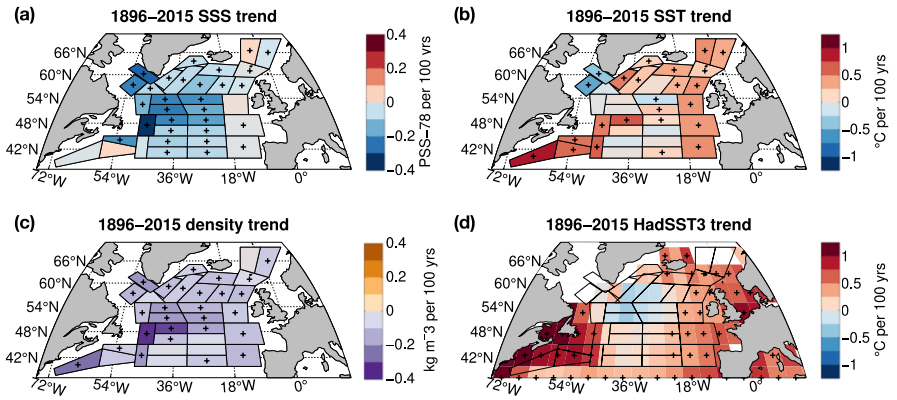


Figure 7. Trend over 1896–2015 (a) SSS; (b), SST; and (c) density; per 100 years. Pluses indicate where the slope magnitude is larger than twice the estimated error. (d) HadSST3 trend, per 100 years.

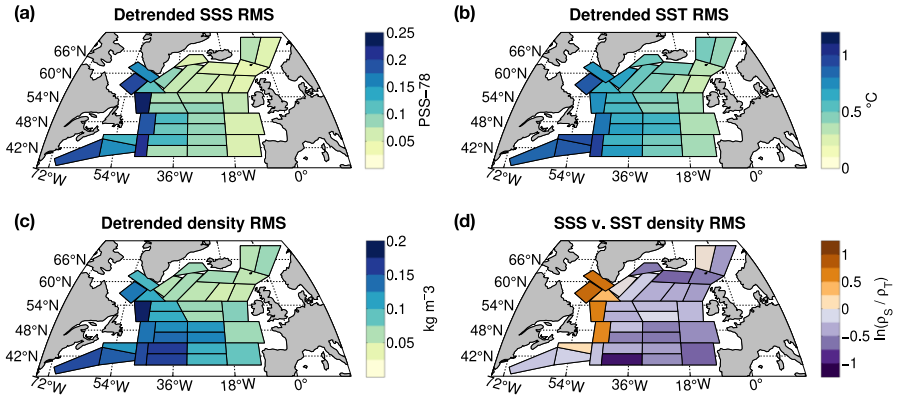


Figure 8. Detrended interannual (1-2-1 filtered) RMS variability for (a) SSS, (b) SST, and (c) density (ρ). (d) Ratio of the contributions of detrended SSS and SST to density RMS variability (plotted as a logarithm).

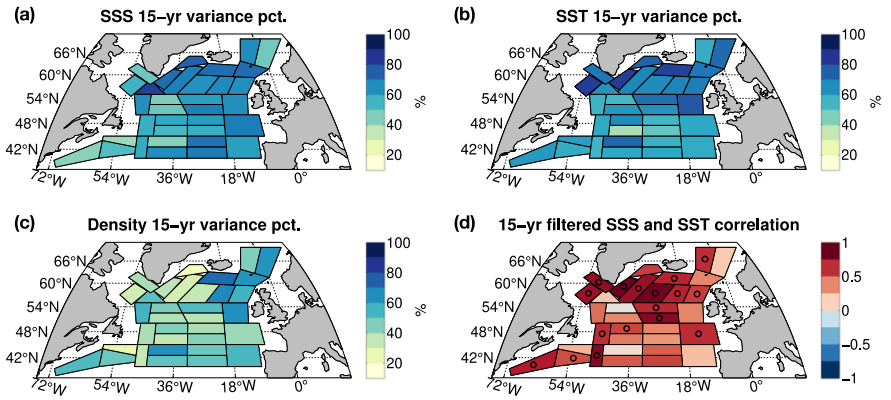


Figure 9. Percentage of detrended interannual (1-2-1) variance explained by the 15-year low-pass filtered time series for (a) SSS, (b) SST, and (c) density. (d) Detrended correlation coefficient between the low-pass filtered SSS and SST. Circles indicate where the correlation is significant at $p<0.05$, estimated with a random-phase bootstrap test to account for serial autocorrelation (Ebisuzaki, 1997).

Appendix A. Comparison with Friedman et al. (2017)

The BINS boxes mainly use the same underlying SSS data as the large boxes north of 40°N in Friedman et al. (2017). The datasets cover a similar area (**Fig. A1a**), though as mentioned previously, the BINS boxes more carefully avoid shelf regions (except for southwest Greenland and the southern part of the Grand Banks). Most of the source data used in the two analyses are the same. Additionally, BINS also incorporates two small datasets from the 1900s and 1910s, plus a few recent transects (and 2014–2015). There are more gaps in the 19 BINS time series than in Friedman et al. (2017): in particular in 1918-1921, and during and just after WWII; these gaps are linearly interpolated in the smaller boxes. Also, when sampling is poor, but varying geographically within the larger boxes, it is possible that some spatial variability is aliased in the temporal variability in the larger boxes.

Fig. A1b compares the NATL index from Friedman et al. (2017), area-averaged SSS from 45°–62°N, with SSS averaged over a similar area in BINS. [NATL from Friedman et al. (2017) is only plotted through 2012, as 2013 was subject to endpoint smoothing]. The two products are very highly correlated (r=0.94, 1896–2012), and compatible considering the differences in area and error estimates. Greater differences are found for smaller regions in the first half of the record, particularly during the gap years mentioned above (not shown)

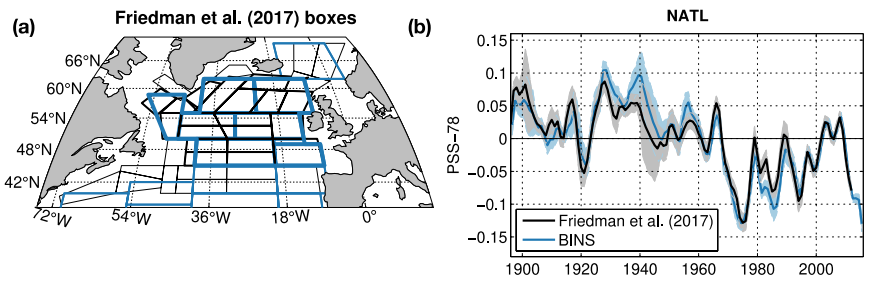


Figure A1. (a) Grid boxes from Friedman et al. (2017) (blue) and BINS (black). Thick lines indicate the NATL region. (c) Comparison of NATL index from Friedman et al. (2017) and BINS grid boxes. Anomalies are from 1896–2012; shading indicates ±2 error terms.

Appendix B. Comparison with Reverdin et al. (2018)

As a check on the BINS time series, we compare the boxes with time series constructed with mostly similar data, monthly binned along two ship routes since mid-1993, intersecting near 59.5°N/32°W: AX02 between Iceland and southern Newfoundland and AX01 between the North Sea and southern Greenland, mostly along 59.5°N (Reverdin et al. 2018), shown in **Fig. B1a**. Time series along AX02 start in July 1993 with few gaps, whereas for AX01 some large data gaps were filled until late 1997. These time series, referred to as B-AX01 and B-AX02, provide increased spatial resolution at seasonal time scales, and portray very coherent variability where they intersect.

We illustrate the comparison of interannual variability of T and S between B-AX01 and BINS for overlapping boxes in the common period 1993-2015 (**Fig. B1b**). For S, the corresponding filtered RMS variability is larger in B-AX01 than in BINS by up to 20%, but with very high correlation coefficients (all larger than 0.95). The smaller RMS amplitudes of salinity in BINS probably result from the larger box sizes and the resulting spatial averaging. RMS variability is more similar for T, but with a slightly smaller correlation in the Iceland Basin (0.80) where gaps in 1993-1997 were the most common in B-AX01. Altogether, the comparisons for the two ship tracks suggest that the method used for BINS in the box averaging to produce interannual variability yields correct results when data coverage is sufficient.

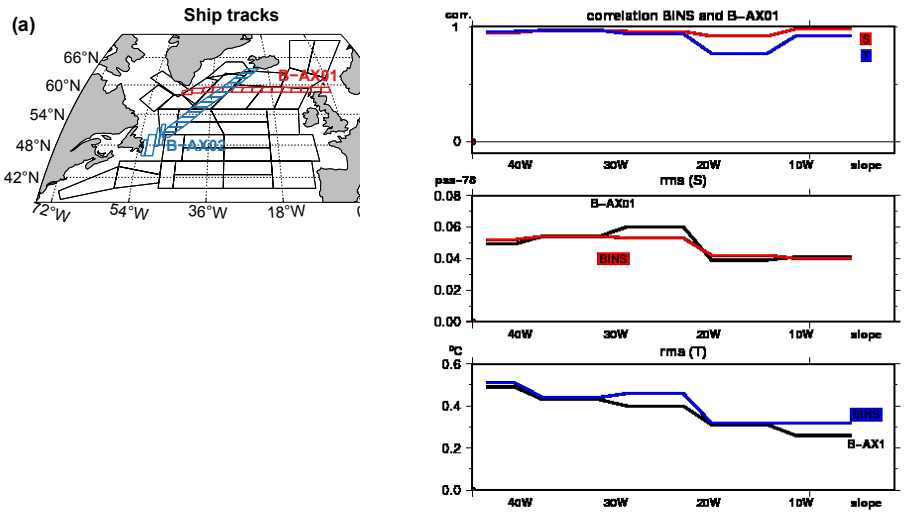


Figure B1. (a) AX01 (red) and AX02 (blue), from (Reverdin et al. 2018). (b) Comparison of B-AX01 with BINS, 1993–2015. The monthly time series of B-AX01 have been yearly-averaged (Dec-Nov), low-pass (1-2-1) filtered over successive years, and then averaged over the bins in BINS: correlation coefficient (top) for T (blue) and S (red), and RMS (middle for S and bottom for T) (B-AX01 in black, BINS in blue (T) and red (S)).

Appendix C: Locations of BINS and HadSST3 grid boxes.

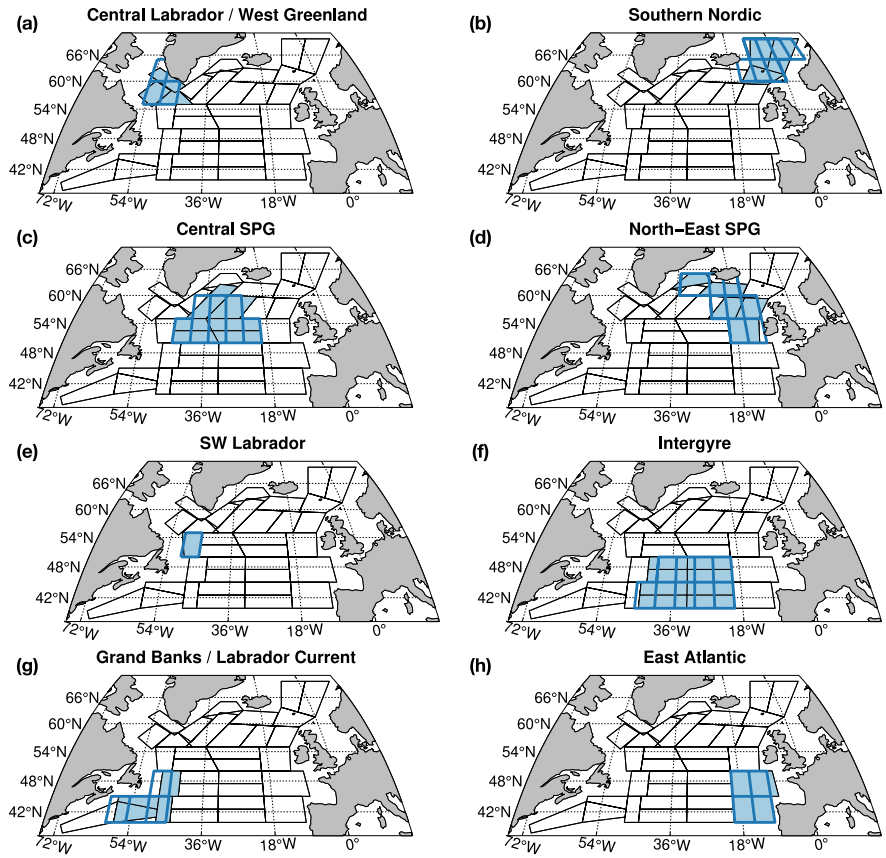


Figure C1. Locations of BINS and HadSST3 grid box regions used in the Section 3.3. Thick blue lines show the 5°×5° HadSST3 grid boxes; shading shows the corresponding BINS grid boxes: (a) central Labrador / West Greenland, (b) southern Nordic, (c) central SPG, (d) north-east SPG, (e) south-west Labrador, (f) intergyre, (g) Grand Banks / Labrador Current, (h) east Atlantic.

authors' response to reviewers' comments
Click here to access/download;authors' response to reviewers' comments;response-reviews.30102018.pdf

Dear reviewers, dear Editor,

Thank-you for your thoughtful reviews.

The time series are made available (up to 2018) at LEGOS/OMP (www.legos.obs-mip.fr/observations/sss/datadelivery/products/tsd-bins-naspg). We are currently discussing with the site managers how they will be referenced on the OMP SEDOO site with a proper doi and this should be available within a week..

We have followed reviewer 1's advice on how to improve figures 4 and 5 (see below). Part of the plot resolution issue is due to conversion in pdf, but that will not be the case in the final, published version.

The response to the different comments is provided below (lines with *Au*).

With best regards, Gilles Reverdin.

Reviewer #1: In this manuscript, the authors present a binned annual product (BINS) of sea surface temperature, salinity, and density for 1896-2015 in the subpolar North Atlantic between 40°N and 70°N. This paper describes the validation of the presented dataset. The methodology for the validation seems reasonable and this is a useful dataset for long-term studies. I recommend accepting this manuscript with minor revisions and look forward to the second part of this study that examines the decadal and multi-decadal variability. My comments are below.

Abstract - Is the water density also a surface-only product? For clarity, I recommend changing likes 27-28 in the abstract from "salinity" to "sea surface salinity" to match the acronym and "water density" to "sea surface density (SSD)"
Au: thank-you we agree that the changes proposed would make this clearer. Yes, the density is a sea surface density (SSD)

Line 29 - "space" to "spatial"
Au: OK

Line 74-75 - Reword this. Upon first read, it gave the impression that hydrographic data have only been used for decadal or longer-term studies, rather than the authors instead describing the temporal extent of the observations.
Au: reworded.

Line 111 - How so? The region presented in this dataset is different than the aforementioned studies.
Au: it is the binning approach that is common with the aforementioned studies. We reworded the sentence.

Line 155 - Citation please regarding the size of the bias. Also, what about regions south of 55°N and north of 65°N that are included in this dataset? There, is this bias significant?
Au: we have reworded the sentence. 'In other parts of the world, this has been documented due to surface heating and evaporation or precipitation (Boutin et al., 2016). In regions of sustained wind conditions, and away from shelves with surface

freshwater sources (melting sea ice or icebergs, outflows from fjords, river plumes), we expect stratification to be small on the order of or less than 0.01 in practical salinity and 0.1°C in temperature. Sustained wind conditions are commonly met in all seasons near 55-65°N, and probably a little less so in particular in spring/summer further south and north.'

Line 182 - Please elaborate. What were the SD criteria for removal?
Au: we did not use n*SD, but a threshold inspired by sigma that varied spatially between 3°C and 6°C.

Line 217 - For a decadal dataset that possesses inadequate sampling of eddies, have you considered removing the eddy component? Otherwise, this will be an inaccuracy that will have to be corrected in any applied study using this research.
Au: This would be very interesting to do, and this can be attempted for recent observations, at least since the 1980s for T, and 1993 for S (or 2010, since SSS products are produced…). Unfortunately, we do not know where eddies were located prior to the satellite data era. Thus, this is unfortunately not feasible over such a long period. Indeed, the binned averages are often undersampling eddy variability, which is a significant component of the error budget in western boxes. Thus, we do not see how we could change that sentence.

Line 228 - Is this a fair assumption? Please cite.
Au: This assumption neglects the part of the error in SST and SSS that might be correlated due to insufficient sampling of eddy variability. On the other hand, the net effect of this on the density error estimate will depend on characteristics of eddies, and cannot be estimated off hand. We are more concerned for that on the seasonal variability and on what we estimate as March density. We don't think that this has been addressed in this way in other studies, and we preferred not to expand on it.

Line 254-255 - Why is this?
Au: This is an interesting point. We suspect that shallower mixed layer and variable atmospheric forcing might be the reason for summer SST deviating from the other seasons and presenting more variance in the central SPG. Here, we don't want to speculate, and will leave this sentence as written.

Line 269 - Awkward wording. Maybe "This is" to "It is"
Au: We changed 'This is for…' by 'For this reason and…, we chose not to…'

Line 282 - Remove "and"
Au: Removed.

The description of the datasets in 3.1 should go in the "Data and Methods" section.
Au: We thus moved in a section 2.3titled 'Other products' the presentation of the other datasets and products used in section 3 (first sentences in 3.1, 3.2 and 3.3)

Figure 1 - Does "approximately" refer to the variation in temporal coverage across the selected boxes? If so, please change the climatology to span the same number of years for a more accurate climatology. Are the SST and SSS climatologies over the same years?

Au: yes, SST and SSS climatologies are estimated over the same years and encompass the period 1896-2000. On the other hand, years with no data are not always the same in different places, but we agree that it is not necessary to add 'approximatively' in the caption. We thus removed 'approximately'

Figures - add error bars to time series.
Au: this is a good idea, in particular for BINS averages (we do not have the equivalent values for the curves from other products in figure 4 and 5). On the other hand, we find it difficult to apply, when we overlay different curves for figure 3. We thus only added error bars to the red curves (BINS) for figures 4 and 5.

The authors did well explaining the various sources of bias in all the validated datasets.

Figure 3 - I am confused by panel (a) being on a different page. How will this look in the publication? The text on panels (b) and (c) is hard to read. Also, please make the panels wider to better see the data and reduce white space. The data for "T Central Labrador" and "S Central Labrador" is touching the Y-axis.
Au: Figure 3 will be of a higher quality in the final version. We also apologize for panel (a) being on another page, in this version. This is not how it will be in the published paper. We left the same vertical scale for the T-panels (b) as the curves for 'Central Labrador' barely touch the lower axis. We changed the scales for S for central Labrador (-0.5 to 0.5). We also tuned a little the scales for S (second and fourth panels), and for T (first, second and fourth panels) to minimize white space.

Figure 4b - "central" to "Central"
Au: Corrected

Figure 4b - again, data is cut off by Y-Axis.
Au: We changed 'Central Labrador Sea' scale to avoid the slight cut-off, and we added error bars on BINS curve.

Figure 5 - Font is too small, data is cut off, and there are no error bars/shading
Au: Font is too small only because of the issue of conversion in pdf. The font quality is much better in the original. Data cut off (slightly) for Central Labrador Sea corrected by extending slightly the scale. Some of the scales were modified on other panels to minimize white space, and we added error bars on BINS curve.

Figure 6 - This is a much more appealing figure. Please make your other figures like this one.

Figure B1 (b) - Change font and font size for readability
Au: This is just an issue of conversion. Font is quite readable in the original eps file.

Reviewer #2: Our ability to look into the past ocean state is restricted by data coverage that greatly reduces back in time. Various efforts have been put forward to homogenize data time series. The authors focus on box averaging approach that has

been proven useful by the team's previous efforts. This paper extends the approach described by Friedman et al. (2017) by including both, SST and SSS anomaly data and by the better shaping of averaging boxes. This approach provides reasonable agreement with alternative compilations on rather large spatial scales and multi-year periods. Given sparse historical data coverage, especially for salinity, the box averaging of anomalies is a practical approach among a few. To some extent, it has been explored before (say in Grodsky et al., 2006), but only in terms of significantly shorter, regional (tropical Atlantic box) study.

The paper mostly focuses on data description. But, to my surprise, there is no link to the dataset described in the paper.

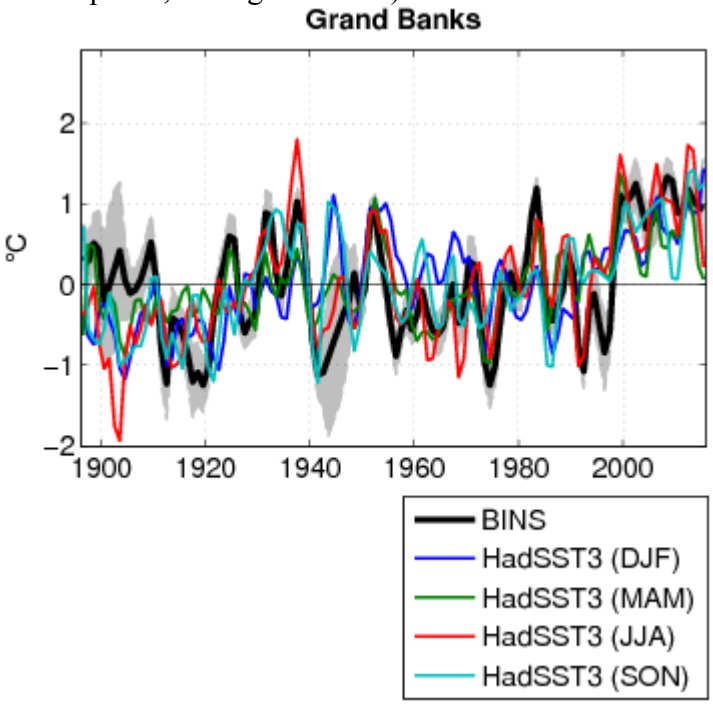
Au: thank-you very much for your 'positive' appreciation of the paper. We decided to place the dataset on the same site that we used for the time series of Friedman et al (2017). For the convenience of the users, we decided to slightly extend in time the time series of the paper to 2018 (which implies that the 1-2-1 running-averaged time series is really updated to 2017). However, we decided to introduce a small inconsistency in our criteria in the western part of the Atlantic to allow larger 'outliers' (as we are quite confident for these recent years that there are not 'errors' in the data). This results in larger error estimates in years 2015-2018 in this region.

Although geophysical interpretations are set aside for a coming paper, this data descriptive part is important. The paper is carefully written with great attention to details and includes an accurate description of data handling procedures. I believe this paper worth publishing. Because it is mostly data descriptive, my evaluation includes only a few minor points.

Minor:
The Abstract states that BINS resolves space scales of the order of 200 to 500 km. I don't see a confirmation of this statement in the main text. Are these scales dictated by characteristic size of a bin? Which zonal and meridional scales are resolved?
Some of the Figures are only episodically discussed. For an example, rather strong SST anomalies are present in Grand Banks box. Particularly strong warming change of ~2C is present in the BINS in the late 1990s. This change is almost twice as strong as in the HadSST3. Could this particular episode be validated/explained?
Normally, salinity contribution to water density increases at cold temperatures. But, Fig.8d depicts mostly zonal dependency rather than meridional dependency consistent with cooler SST in the north. Could you please come up with more explanation of this result?

Au: The scale of 200 to 500 km is dictated by the number of data available and thus by the bin size necessary to reduce the sampling error. The scales resolved thus correspond to the typical size of the boxes. This is variable but hovers around 200-500 km. Typically, thus we could resolve spatial wavelengths on the order of 3 box size (~1000 km) (thus box size ~scale). Some boxes are distorted zonally, while other are in a different direction, based on considerations on ocean structures (thus by the orientation of oceanic fronts or the direction of currents). Thus, we have changed the sentence in the introduction.
Au: We agree that some features require further discussion. We thank the reviewer for having pointed out the differences for the Grand Banks box time series. We are rather

confident in our time series, but have no explanation on why the SST variability portrayed there is different than in HadSST3 (or in EN4, to some extent) since the mid-1990s. We checked data distribution, which is certainly not ideal in particular over the southern Grand Banks bin for that period. There should not be major issues with seasonal sea ice, and the fact that most of the boxes in BINS that are included do not include the fall season is not the major reason for the discrepancies with HADSST3 (in HadSST3, the different seasons have the same timing for the different recent peaks, see figure below).



We also agree that the spatial distribution in Figure 8d is indeed worth further comments, which we have now introduced. Indeed, the mostly zonal structure in this ratio does not follow the expectation of having a pattern closer to the one in SST (thus with an increase in the north). This is because of the decrease in rms(SSS)/rms(SST) (compare 8 a to 8b) towards the north that compensates the effect of decreasing temperature.

Edits:
Line 343 but that exclude(s) Nansen
Line 356 For SSS, the large positive peak in BINS originate(s)
Line 450 boxes with 5 or less (replace 'less' by 'fewer') years
Line 485 density tends to be positively correlated to (replace 'to' by 'with') SSS
Lines 506-7 Indeed, the low-pass filtered correlation of SSS and SST (Fig. 8d) is positive in most grid boxes. Should it be Fig.9d?
Line 550 Possibly, (the) insufficient accuracy of [remove the] binned time series and the change in
Line 577 We base our evaluation (smth is missing there) that they provide reasonable

Grodsky, S. A., J. A. Carton, and F. M. Bingham (2006), Low frequency variation of sea surface salinity in the tropical Atlantic, Geophys. Res. Lett., 33, L14604, doi:10.1029/2006GL026426

Au : Thank-you very much for the 'Edits' proposed that we have incorporated.