Seasonality of Trends and Modal Variability in Sea Surface Temperature in the Northwest Atlantic

John W. Loder and Zeliang Wang

Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive, Dartmouth, Nova Scotia, B2Y 4A2

2024

Canadian Technical Report of
Hydrography and Ocean Sciences 384

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of Hydrography and Ocean Sciences

Technical reports contain scientific and technical information of a type that represents a contribution to existing knowledge but which is not normally found in the primary literature. The subject matter is generally related to programs and interests of the Oceans and Science sectors of Fisheries and Oceans Canada.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts.*

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Regional and headquarters establishments of Ocean Science and Surveys ceased publication of their various report series as of December 1981. A complete listing of these publications and the last number issued under each title are published in the *Canadian Journal of Fisheries and Aquatic Sciences*, Volume 38: Index to Publications 1981. The current series began with Report Number 1 in January 1982.

Rapport technique canadien sur l'hydrographie et les sciences océaniques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles mais que l'on ne trouve pas normalement dans les revues scientifiques. Le sujet est généralement rattaché aux programmes et intérêts des secteurs des Océans et des Sciences de Pêches et Océans Canada.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page de titre.

Les établissements de l'ancien secteur des Sciences et Levés océaniques dans les régions et à l'administration centrale ont cessé de publier leurs diverses séries de rapports en décembre 1981. Vous trouverez dans l'index des publications du volume 38 du *Journal canadien des sciences halieutiques et aquatiques*, la liste de ces publications ainsi que le dernier numéro paru dans chaque catégorie. La nouvelle série a commencé avec la publication du rapport numéro 1 en janvier 1982.

Canadian Technical Report of
Hydrography and Ocean Sciences 384

2024

Seasonality of Trends and Modal Variability in Sea Surface Temperature in the Northwest Atlantic

by

John W. Loder and Zeliang Wang

Fisheries and Oceans Canada
Bedford Institute of Oceanography
1 Challenger Drive, Dartmouth, Nova Scotia, B2Y 4A2

Cat. No.: Fs97-18/384E-PDF ISBN: 978-0-660-72837-7 ISSN: 1488-5417

Correct citation for this publication:

Loder, J.W. and Wang, Z. 2024. Seasonality of Trends and Modal Variability in Sea Surface Temperature in the Northwest Atlantic. Can. Tech. Rep. Hydrogr. Ocean Sci. 384: vi + 70 p.

TABLE OF CONTENTS

ABSTRACT

Loder, J.W. and Wang, Z. 2024. Seasonality of Trends and Modal Variability in Sea Surface Temperature in the Northwest Atlantic. Can. Tech. Rep. Hydrogr. Ocean Sci. 384: vi + 70 p.

The seasonality of long-term change and decadal-scale variability in sea surface temperature (SST) in the Northwest Atlantic (NWA) during 1870-2015 is examined using two datasets: ERSST and HadISST1. Suspect features in both interpolated datasets prior to 1940 are identified. Trends vary seasonally, spatially and with historical duration, associated with an apparent spatially-varying anthropogenic trend, modes of large-scale variability on decadal time scales with spatially-varying amplitudes, and the dataset discrepancies. The "warming hole" south of Greenland, with net cooling on the century scale, has reduced extent in summer. Trends on this scale are largest in summer - about 0.06ºC and 0.03ºC per decade averaged over the western and eastern NWA, respectively. Since 1950, annual trends in these regions are 0.09ºC and 0.5ºC per decade, respectively; in the western NWA, largest magnitudes are in summer and smallest ones in winter. Annual trends from Atlantic Canadian shelf monitoring are slightly larger than from the two datasets, with sub-regional differences in seasonality. Trends during the satellite era (since 1979) are larger by a factor of 3-4, in part due to a contribution from a warming phase of the AMO, but also apparently from increasing anthropogenic warming.

RÉSUMÉ

Loder, J.W. and Wang, Z. 2024. Seasonality of Trends and Modal Variability in Sea Surface Temperature in the Northwest Atlantic. Can. Tech. Rep. Hydrogr. Ocean Sci. 384: vi + 70 p.

Le caractère saisonnier du changement à long terme et la variabilité à l'échelle décennale de la température à la surface de la mer dans l'Atlantique Nord-Ouest de 1870 à 2015 sont examinés à l'aide de deux ensembles de données : l'ensemble de données étendu et reconstitué de la température à la surface de la mer (ERSST) et l'ensemble de données de la température à la surface de la mer et de la glace de mer du Hadley Centre (HadISST1). Les valeurs douteuses dans les deux ensembles de données interpolées antérieures à 1940 ont été repérées. Les tendances varient en fonction des saisons, de l'espace et de la durée historique. Elles sont associées à une tendance anthropique apparente variant dans l'espace, à des modes de variabilité à grande échelle sur des échelles décennales avec des amplitudes variant dans l'espace, et à des divergences entre les ensembles de données. Le « trou du réchauffement » au sud du Groenland, avec un refroidissement net sur un siècle, présente une étendue réduite en été. Les tendances à cette échelle sont plus importantes en été – environ 0,06 °C et 0,03 °C par décennie en moyenne pour l'ouest et l'est de l'Atlantique Nord-Ouest, respectivement. Depuis 1950, les tendances annuelles dans ces régions sont respectivement de 0,09 °C et de 0,5 °C par décennie; dans l'ouest de l'Atlantique Nord-Ouest, les plus grandes amplitudes se produisent en été et les plus petites en hiver. Les tendances annuelles issues de la surveillance du plateau du Canada atlantique sont légèrement supérieures à celles des deux ensembles de données, avec des différences sous-régionales dans le caractère saisonnier. Les tendances observées au cours de la période de mesures satellitaires (depuis 1979) sont 3 à 4 fois plus importantes, en partie en raison de la contribution d'une phase de réchauffement de l'oscillation atlantique multidécennale, mais aussi, apparemment, de l'augmentation du réchauffement anthropique.

1 Introduction

Sea surface temperature (SST) is an ocean variable with the goodhistorical data coverage that is frequently used in climate change and variability studies (e.g., Yasunaka & Hanawa, 2011; Hartmann et al., 2013). However, spatial and temporal gaps in coverage, and different sampling and interpolation methodologies are major issues in its use, such that adjustments for sampling changes may be needed and careful analyses are required (e.g., Smith & Reynolds, 2004; Rayner et al., 2006; Kennedy, 2014; Karl et al., 2015). Further, considering the predominant seasonality at mid- and high-latitudes, the annual-mean SSTs commonly used for climate indices may not reflect the actual temperatures affecting the atmosphere, marine biogeochemistry, and ecosystems on seasonal and shorter timescales (e.g., Mann and Lazier, 2006; Drinkwater et al., 2010; Sheridan and Lee, 2018).

In this paper a "seasonality" follow-up to Loder and Wang's (2015; hereinafter LW2015) examination of long-term SST trends and variability in the Northwest Atlantic (NWA) is provided, which used three interpolated gridded historical datasets (up to 2011) complemented by time series from Fisheries and Oceans Canada (DFO) *in situ* monitoring at four sites off Atlantic Canada. Specifically, analysis is extended to examine seasonal long-term and modal variability up to 2015 in the latest versions of the two most widely-used interpolated gridded datasets: version 4 (ERSST.v4) of the Extended Reconstructed SST (Huang et al., 2015), and version 1.1 of the Hadley Centre Interpolated SST (HadISST1) database (Rayner et al., 2003). Discrepancies and scientific issues are discussed in relation to: (i) seasonal *in situ* monitoring time series at four DFO sites; (ii) annual series from two US coastal monitoring areas (Shearman and Lentz, 2009); (iii) the most recent release of the primary non-interpolated gridded global monthly SST dataset used in climate variability studies, namely, ICOADS R3.0 of the International Comprehensive Ocean-Atmosphere Data Set (Freeman et al., 2017); and (iv) the Simple Ocean Data Assimilation (SODA) v2.2.4 global ocean reanalysis simulation for the period 1871-2010 (Chepurin et al., 2014).

As described in LW2015 and elsewhere, there are many motivations for further investigation of SST variability in the North Atlantic (NA) and in the NWA in particular. These range from:

its established and potential roles in the global and basin-scale atmosphere-ice-ocean climate systems (e.g., Stocker et al., 2013; Rahmstorf et al., 2015; Chen et al., 2017), distinct regional modes of natural variability associated with the Atlantic Multi-decadal Oscillation (AMO; e.g., Enfield et al., 2001; Drinkwater et al., 2014) and North Atlantic Oscillation (NAO; e.g., Deser et al., 2010; Hurrell & Deser, 2010), and emerging indications of substantial anthropogenic change (e.g., Terray, 2012; Bakker et al., 2016; Saba et al., 2016; Thibodeau et al., 2018); to

observational datasets with notable discrepancies (e.g., Yasunaka and Hanawa, 2011; Hausfather et al., 2017), and limited information on key aspects such as interactions of the NA subpolar and subtropical gyres' western boundary currents (e.g., Loder et al., 1998; Joyce and Zhang, 2010; Buckley and Marshall, 2016; Foukal and Lozier, 2016); to

uncertainties in the SST variability's controlling dynamics, and future changes and impacts (e.g., Zhang et al., 2013; Ting et al., 2014; IPCC, 2014; Brown et al. 2016; Zhang, 2017).

In addition, the NWA is a region with a significant observed natural variability in SST (e.g., Thompson et al., 1988; Deser et al., 2010), a primary deep convection zone affecting the Atlantic Meridional Overturning Circulation (AMOC; e.g., Rhein et al., 2015; Lozier et al., 2016) contributing to decadal-scale variability (e.g., Danabasoglu et al., 2012; Yashayaev and Loder, 2016), and the largest spread among climate models in projected future changes (e.g., Collins et al., 2013; Knutti & Sedláček, 2013).

LW2015 carried out trend, empirical orthogonal function (EOF) and correlation analyses on the annual and summer means from version 3b of the ERSST (Smith et al., 2008), version 1 of the HadISST1, and the Centennial *in situ* Observation Based Estimates (COBE; Ishii et al., 2005) interpolated gridded datasets. They found good overall agreement among the various datasets but notable discrepancies in some areas and periods (particularly in the trends), strong decadal-scale variability associated in part with the AMO and NAO, pronounced spatial structure in the trends and modal variability patterns, and similarity between the annual and summer results (but with some summer intensification).

LW2015's findings together with those of other studies point to a number of important issues to which an analysis of the seasonality of sub-annual SST variability is relevant. These include:

- the seasonality, spatial structure and origin of the strong natural variability on decadal time scales;
- the extent to which larger-scale oceanic variations like the AMO and those in the AMOC (e.g., Buckley & Marshall, 2016) and Gulf Stream (e.g., Watelet et al., 2017; McCarthy et al., 2018) are affecting SST in shelf and coastal areas (such as at DFO monitoring sites);
- the seasonal and spatial structure of anthropogenic global warming in NWA SST, and its importance relative to the strong decadal-scale SST variability in the region (e.g., Loder et al., 2015; Josey et al., 2018);
- the extent to which recent indications of amplified warming in the NWA's mid-latitude transition zone (between the gyres) are related to anthropogenic versus natural variability (e.g., Brock et al., 2012; Chen et al., 2015; Pershing et al., 2015; Saba et al., 2016; Kavanaugh et al., 2017; Claret et al., 2018; Thibodeau et al., 2018); and

the pitfalls of regional (or local) climate change studies using limited-duration ocean temperature indices in situations with strong decadal-scale variability (e.g., Friedland & Hare, 2007; Belkin, 2009; Mills et al., 2013; Pershing et al., 2015; Larouche & Galbraith, 2016; Thomas et al., 2017), such that trends depend on observational record lengths and hence can be misleading with respect to anthropogenic change.

Because of limited data coverage (in duration, space and time), the issues of data quality, natural (e.g., decadal-scale) variability and long-term trends are intertwined such that a standard serial approach of establishing quality and then addressing variability is problematic. Consequently, we draw on various variability analyses throughout the paper for indications of data quality and consistency, while discussing the above scientific issues and updating our results to 2015. Our analysis of DFO monitoring *in situ* SST datasets that have not been adjusted for sampling inhomogeneities is also relevant to concerns that artificial trends may have been introduced in historical datasets through adjustments intended to account for different methodologies (e.g., Soon et al., 2015).

We note that our approach to seasonality (i.e., the differences among temperature variations in seasons of fixed dates) is different from (but complementary to) those in the recent studies of Friedland and Hare (2007) and Galbraith and Larouche (2013) in which threshold temperatures were used to define the seasons (see Thomas et al., 2017 for discussion of the inter-relation of these approaches). Also, while we discuss our results in terms of the commonly-discussed AMO (sometimes alternatively referred to as the Atlantic Multi-decadal Variability) in NA SST, we recognize that there are alternative interpretations of the NA's variability on this time scale (e.g., related to atmospheric blocking or global air temperature variations associated with aerosols; Häkkinen et al., 2011; Booth et al., 2012; Zhang et al., 2015).

In Section 2, a description of the datasets and methods is provided. Following, in Section 3, with direct inter-comparisons of various datasets – both between the interpolated gridded datasets and with time series from the non-interpolated gridded dataset, dedicated *in situ* monitoring programs and the reanalysis product– involving difference computations, correlation analyses and comparisons of trends over the length of the (shorter) monitoring programs. Some discrepancies among the datasets are identified with suggestions provided for their origin. Then, in Section 4, a focus on the identification of SST trends and their seasonality and spatial structure is completed, examining specific historical periods of the gridded datasets corresponding to eras with different levels of data coverage (and probably quality).

The natural variability is provided in Section 5, initially based on EOF analyses of the gridded datasets over their longest common period (back to 1870) without prior de-trending (in order to examine all variability), and then on EOF analyses of the datasets since 1900 after de-trending, intended to capture the portion of the current "Anthropocene" (e.g., Steffen et al., 2016) with better data coverage and nearly two AMO cycles. Then, in Section 6, a discussion of the main results and some outstanding issues is summarized.

A glossary of abbreviations and acronyms is provided in Appendix A, and results complementing those in the main text are provided in the other appendices and the Supporting Information (SI).

2 Data and methods

a *Gridded SST Datasets*

The focal data are two interpolated gridded (global) historical datasets of monthly SST used widely in model hindcast and climate studies:

ERSST.v4 (hereinafter, ERSST) which is the latest Extended Reconstructed SST dataset of the US National Centers for Environmental Information, with 2º x 2º resolution extending back to 1854 (Huang et al., 2015, 2016; Liu et al., 2016; https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v4). Its primary data sources are Release 2.5 of ICOADS (Woodruff et al., 2011) and, since 2007, the Global Telecommunications System (GTS). Other available SST, surface air temperature and sea ice datasets were used (by previous authors) in quality control and adjustments of ERSST. Gaps were filled using EOFs for high-frequency variability and spatial interpolation for low-frequency variability.

HadISST1 v1.1 (hereinafter HadISST1) which is the UK Met Office's Hadley Centre Interpolated (Sea Ice and) SST dataset, with 1º x 1º resolution extending back to 1870 (Rayner et al., 2003; http://www.metoffice.gov.uk/hadobs/hadisst/; also see https://climatedataguide.ucar.edu/climate-data/sst-data-hadisst-v11). Its primary data sources are the UK Met Office's Marine Data Bank, ICOADS releases and, since 1982, the GTS. Reduced-space optimal interpolation was used in filling gaps. Non-interpolated products from its sources, such as the HadSST2 monthly averages over 5º x 5º grid squares (e.g., Rayner et al., 2006; Kennedy et al., 2011b), were used in our quality assessments.

HadISST1 is basically an update to 2015 of its version 1 used in LW2015, while ERSST here is both a similar update and an upgrade on its version 3b. Extensive work has gone into the development, comparison, improvement and assessment of these interpolated gridded historical SST datasets (e.g., Smith and Reynolds, 2003, 2004; Ishii et al., 2005; Smith et al., 2008; Rayner et al., 2010; Kennedy et al., 2011a,b; Yasunaka and Hanawa, 2011; and other references indicated above). However, it is generally recognized (e.g., Kennedy, 2014; Hausfather et al., 2017) that some significant discrepancies remain between the datasets associated with observational data quality and sparsity, and different sampling, adjustment and interpolation methods. A notable difference between the two interpolated datasets used here is that, in addition to ERSST having coarser resolution (including coastline and island representations) than HadISST1, its version 4 has additional spatial smoothing beyond version 3b.

For the present study, a domain (31 to 67°N, 3 to 81°W) focussed on the Western North Atlantic (WNA) off northeastern North America was chosen (Fig. 1), extending northward to Davis Strait, eastward to Europe and southward past Cape Hatteras (but excluding Hudson Bay and Strait in the west). This domain is smaller than the primary NA domain, but larger than the zoomed NWA one, used in LW2015. Derived datasets of annual and seasonal means of SST were computed for each interpolated dataset's native grid and their common period (since 1870), taking January-March as winter, April-June as spring, July-September as summer and October-December as fall. Since HadISST1 has missing or dubious values in ice-covered areas of Davis Strait and the Labrador Shelf in winter and spring, we omitted those grid points from the analyses for those seasons.

As a first examination of the seasonality of the subannual SST variability on a broad scale, time series of annual and seasonal means were computed for the three major "oceanographic" regions outlined in Fig. 1, which we refer to as the northwest WNA (NW-WNA), mid-latitude WNA (ML-WNA) and northeast NA (NE-NA). These regions (taken as rectangular for convenience) were chosen to be within the domains of both interpolated datasets and to avoid areas with seasonal sea ice.

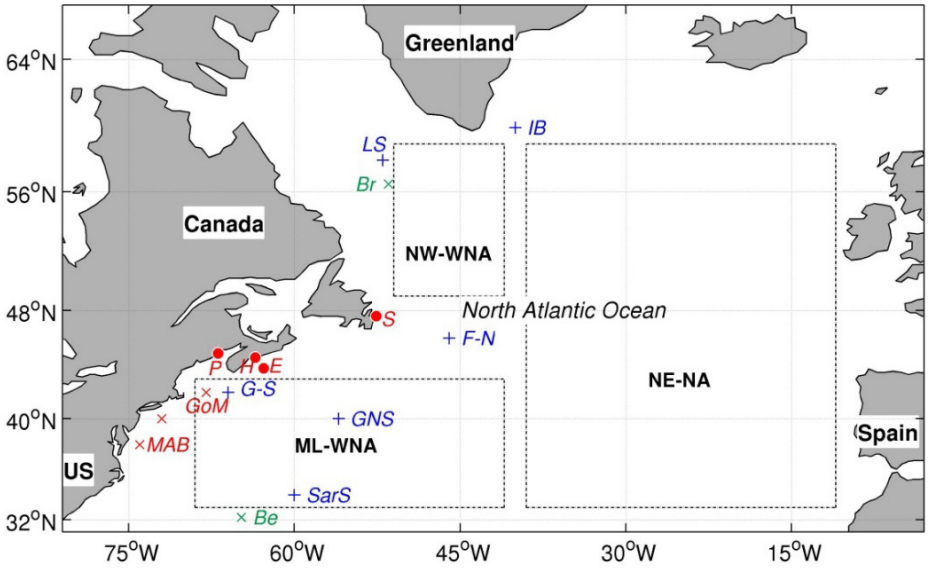

Figure 1. Location map showing: the entire domain used in the gridded data analyses, the three major regions (NW-WNA, ML-WNA and NE-NA) outlined by the dotted lines, the center positions (+'s in blue) of the six areas with persistent ER minus HD differences (*IB*, *LS*, *F-N* for FC-NB, *G-S* for GoM-SS, *GSN*, and *SarS*), the four DFO monitoring sites (solid circles in red: *P*=P5, *H*=HH, *E*=EB, and *S*=S27), the center points (x's in brown) of the *MAB* and *GoM* grid boxes, and the locations of Bravo (*Br*) and Bermuda (*Be*) (x's in green). See text or Appendix A for the definition of acronyms.

Following LW2015, four historical periods were considered for trend and EOF analyses: 1870-2015, the longest common period; 1900-2015, for a century-scale view; 1950-2015, the

modern era of good data coverage; and 1979-2015, the satellite era of best data coverage. Standard linear regression techniques were used to compute trends in the SST time series for different positions, regions, seasons and periods, with p values used to provide an indication of statistical significance at the 90 and 95% confidence levels. EOF analyses were carried out on the ERSST and HadISST1 annual- and seasonal-mean series across the full domain for two periods: 1870-2015 with prior de-meaning but without prior de-trending, and 1900-2015 with both prior de-meaning and de-trending (as a first approximation to removal of an anthropogenic trend). The EOF values in each analysis were normalized with respect to their maximum absolute value in the domain, and the corresponding principal component (PC) time series in each case were then scaled by multiplying each of their values by the same maximum EOF value that was used in the EOF normalization.

In examining the prominent (SST) discrepancies between ERSST and HadISST1 in some areas and periods, we also briefly consider SST time series from:

- the non-interpolated ICOADS Release 3.0 dataset which has monthly statistics on a 2^o x 2^o grid during the period 1800-2014 (Freeman et al., 2017; http://icoads.noaa.gov/products.html); and
- the SODA 2.2.4 reanalysis product which has monthly means on a 0.25^o x 0.4^o grid during the period 1871-2010 (Carton & Giese, 2008; Giese & Ray, 2011; Chepurin et al., 2014; http://apdrc.soest.hawaii.edu/datadoc/soda_2.2.4.php).

For each of six areas with persistent discrepancies, we computed seasonal means in each year as the simple average of the available monthly medians (ICOADS) or means (SODA) over grid squares or points in representative subareas (see the SI for details).

b *DFO and Other Coastal Monitoring Time Series*

Time series of monthly near-surface (upper 2m) ocean temperature from *in situ* measurements at three of the four DFO monitoring sites considered in LW2015 were used to create annual- and seasonal-mean time series for comparison with those from nearby grid positions in the gridded datasets. These sites are Prince 5 (P5) in Passamaquoddy Bay in the outer Bay of Fundy adjoining the Gulf of Maine (GoM; data since 1924), Emerald Basin (EB) on the central Scotian Shelf (data since 1947), and Station 27 (S27) on the Grand Bank off St. John's, Newfoundland (data since 1946). In addition, a newly-"adjusted" coastal SST time series from Halifax Harbour (HH; data since 1926) is examined as a complement to the EB time series for the Scotian Shelf (see Fig. 1 for site locations). These data were available through DFO's Atlantic Zone Monitoring Program (AZMP, http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/azmp-pmza/index-eng.html; see Colbourne et al., 2016, and Hebert et al., 2016, for recent summary reports). The adjustment of the HH time series (Roger Pettipas, personal communication, 2017) is described in Appendix B. Comparison with *in situ* time series from the

Labrador Sea DFO-Bravo site used in LW2015 is not included here because of very limited seasonal coverage other than during 1962-1978 and since 2002.

Since limited sampling (typically only 0-2 observations per month) is an issue for EB in all years and for S27 in some early years, seasonal and annual means were computed by adding the individual-year seasonal and annual anomalies to their corresponding record means, with these anomalies computed as averages of the appropriate monthly anomalies (about their record means). As a coarse screening for temporal coverage, seasonal means were only retained in the main body of the paper when there were observations from at least two (of the possible three) months, and annual means when there were observations from at least two seasons (there were observations from less than three seasons in only 2% of the years retained). Seasonal means with data from only one month (in a few years) are included in Appendix B which examines EB and HH SST in more detail.

Additional long annual-mean SST time series based on *in situ* observations in the study region have been reported by Shearman and Lentz (2010; hereinafter, SL2010), for 1875 to 2007 from lighthouses and lightships between the GoM and Florida. Their computations included spatially-averaged annual SST time series for the GoM and Middle Atlantic Bight (MAB), displayed in their Fig. 8 as anomalies (together with their trends). For comparison of the variability in these time series with those in the ERSST and HadISST1 datasets, we have added approximate estimates of the climatological means for these areas (from Fig. 4a in SL2010) to the annual anomalies to obtain time series of annual SST.

c Potential Forcing Indices

In examining the relation of SST variability to that in larger-scale and potential forcing indices, we use four observational indices:

- global mean (annual, land-sea) surface temperature (GMST), taken from the US National Aeronautics and Space Agency website http://data.giss.nasa.gov/gistemp/ (Hansen et al., 2010; GISTEMP Team, 2016);
- the (de-trended) AMO index (the area-averaged annual SST in the NA north of the equator), taken from the US National Oceanic and Atmospheric Administration (NOAA) website https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/AMO/ (e.g., Enfield et al., 2001; Knight et al., 2005);
- the de-trended winter (January-March) NAO index, computed from monthly data taken from the same NOAA website https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/NAO/ (e.g., Hurrell & Deser, 2010); and
- the annual Bermuda sea surface height (SSH) time series taken from the Permanent Service for Mean Sea Level website http://www.psmsl.org/data/ (e.g., Holgate et al., 2013).

We have also considered a number of annual indices from the Wang et al., (2015) model hindcast study of ocean circulation variability in the NA during 1958-2004; namely, the first three PCs of SSH, the first three PCs of the depth-integrated transport streamfunction (ψ), the

Bermuda minus Bravo (Be-Br) SSH difference, and the first PC of the wind stress curl (WSC) over their model domain. In this and other investigations we use standard correlation analyses, with p-values used to provide an indication of statistical significance at the 90 and 95% confidence levels (since autocorrelation in the various time series has not been accounted for, the significance may be overestimated by these metrics).

Finally, in examining SST variability at EB and HH (Appendix B), we also discuss variability in the annual and seasonal means from two long time series of air temperature (AT) from the Scotian Shelf region, from Environment and Climate Change Canada's (ECCC's) Adjusted and Homogenized Canadian Climate Data (AHCCD; http://ec.gc.ca/dccha-ahccd/Default.asp?lang=En&n=B1F8423A-1; Vincent et al., 2012): for Halifax (1872-2015) and Sable Island (1898-2015). We note that examination of the relation between SST and AT more broadly would be a natural complement to the present study (see, for example, Galbraith et al., 2012, and Galbraith & Larouche, 2013), but is beyond the scope at this point.

3 Intercomparison of SST datasets

a *Differences between ERSST and HadISST1 in Major Oceanographic Regions*

To investigate the broad-scale temporal variability of SST in the study domain and the discrepancies between the ERSST and HadISST1 datasets, we first examine time series of spatially-averaged annual-mean (Fig. 2a) and seasonal-mean (Fig. 3) SST for the three major regions (Fig. 1):

- a northwestern, largely subpolar region which we refer to as the NW-WNA, taken to be the ocean grid points within 49-59°N, 41-51°W (avoiding sea ice areas);
- a mid-latitude western region which we refer to as the ML-WNA covering the southwestern portion of the subpolar-subtropical transition zone in the WNA (e.g., Brock et al., 2012), taken to be the points within 33-43°N, 41-69°W; and
- a larger eastern region which we refer to as the NE-NA in which the Gulf Stream extension bifurcates into the subpolar and subtropical gyre, taken to be 33-59°N, 11-39°W.

To place these time series in an oceanographic perspective we also refer to maps (Fig. 4) of the SST differences (ERSST minus HadISST1) between the two datasets, for the annual and seasonal means averaged over four successive sub-periods which were identified from the time series in Figs. 2 and 3 (see below; note that these sub-periods differ from the historical periods dating back from 2015 that were identified in the previous section for the trend comparisons). The difference patterns in these sub-periods will be discussed further in Section 3b, but we note here the pronounced spatial structure of the differences, with both similarities and contrasts across seasons and sub-periods (Fig. 4). Corresponding maps of the SST averages over the sub-periods for the annual and seasonal means and each dataset can be found in the SI (Fig. SI-1).

1 ANNUAL MEANS

Starting with the annual time series (Fig. 2a), a multi-decadal variation and a long-term warming are apparent in both datasets in all three regions. However, there are notable differences between the datasets. In the NW-WNA, there is generally good agreement between the annual means from the two datasets since 1925, but a sub-period between 1895 and 1925 when the ERSST means were low in most years by about 0.5ºC. In contrast (and consistent with the spatial variations in the sign of the ERSST-HadISST1 differences to be discussed in relation to Fig. 4), the ERSST annual means in the ML-WNA were higher than their HadISST1 counterparts by 0.5-1.0ºC prior to the early 1900s and by ~0.5ºC between 1912 and 1940. The differences are generally much smaller in the NE-NA than in the other two regions, but this region's ERSST means were higher than its HadISST1 ones by ~0.3ºC prior to the early 1880s and by ~0.5ºC in the early 1940s. While the above differences may appear to be small, they can be very important in the determination of the magnitudes of trends of order 0.1ºC per decade (or 1ºC per century), further exacerbating complications from the presence of multi-decadal variability and observational records having much shorter duration in many areas (see below).

To put the variability in these datasets in basin- and global-scale contexts before proceeding to seasonality, we show the same annual-mean data in Fig. 2b in the form of their anomalies about their record means, together with the AMO index and the annual anomalies of GMST (relative to their respective record means). In all three regions, the anomalies show a multi-decadal variation similar that in the AMO (the apparent smaller amplitude in the AMO is partly related to it being de-trended). The low-frequency variability in the interpolated datasets' anomalies also has some similarity to that in GMST, both in a long-term warming and a multi-decadal variation, but there are differences in the apparent periodicity and phase of the latter (LW2015). The linkage between the decadal-scale variability in NA SST and that in larger-scale AT is a topic of ongoing exploration by others (e.g., Booth et al., 2012; Zhang et al., 2013; Chen & Tung, 2018; Lian et al., 2018), and will not be pursued here. However, we note recent findings (Chen et al., 2017; Chen & Tung, 2018) that the AMO is the largest contributor to multi-decadal variability in GMST and global mean SST, much larger than that of the Pacific Decadal Oscillation (which is the largest contributor on interannual time scales).

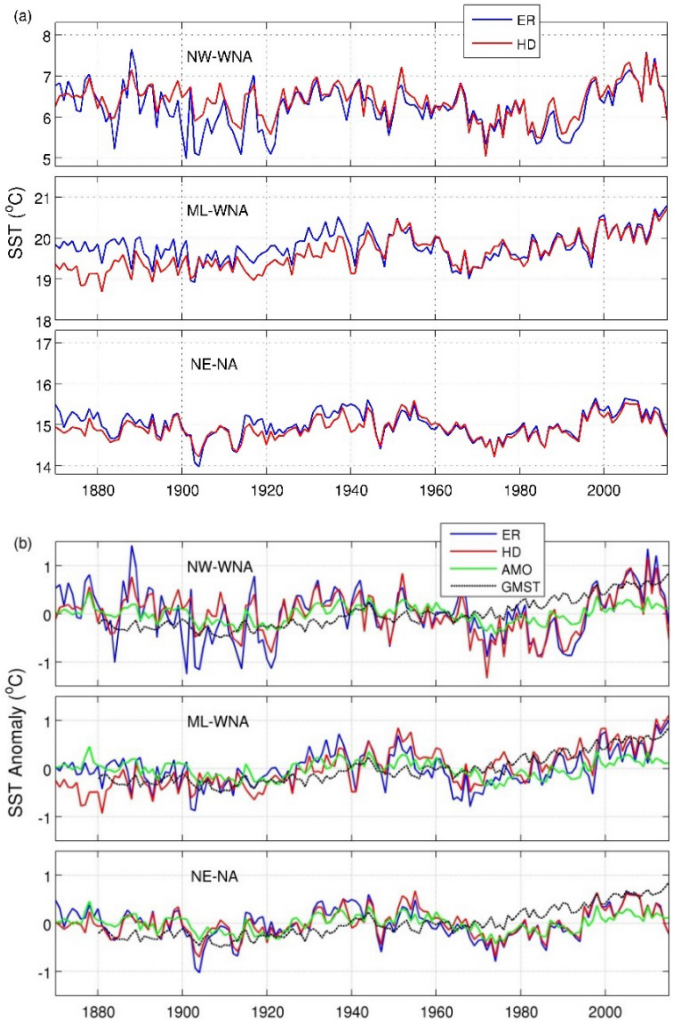

Figure 2. (a) Annual means of SST averaged over the three major regions (NW-WNA, ML-WNA, NE-NA; see Fig. 1 for locations) from ERSST (ER, blue) and HadISST1 (HD, red). (b) Annual SST anomalies (about their record means) for the three regions from the same datasets, with the AMO index (green) and GMST anomalies (dashed black) added to each panel. Note the common SST scales in the different panels in each of (a) and (b).

2 SEASONAL MEANS

To examine further the seasonality of the variability in, and differences between, the two interpolated datasets and also the spatial variability of this seasonality, we show in Fig. 3 time series of their seasonal means for the three major regions, complementing the seasonal difference maps for the sub-periods (Fig. 4). We also refer to the trends in these time series, shown in Table 1 for the annual as well as the seasonal means, and for the four historical periods (extending up to 2015) identified above.

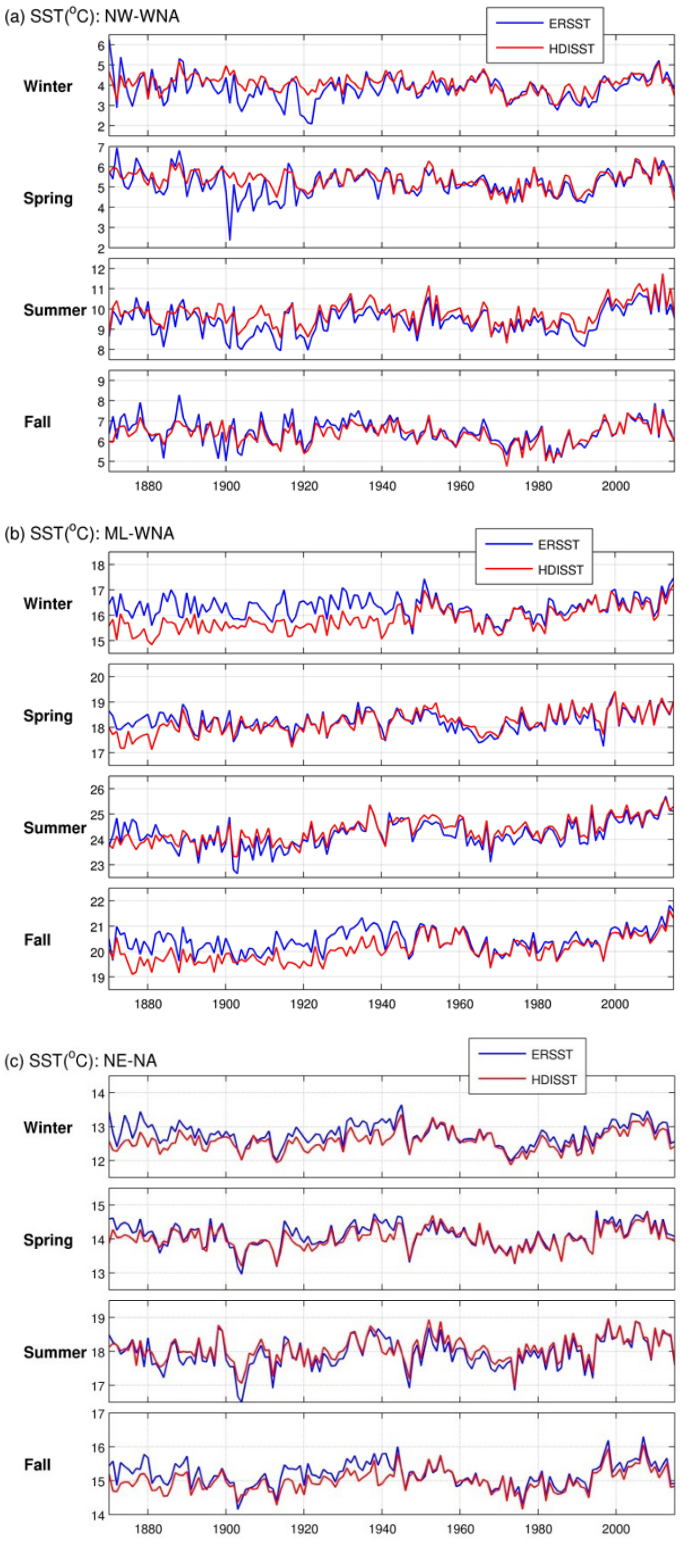

Figure 3. Seasonal means of SST from ERSST (blue) and HadISST1 (red) averaged over the three major regions (Fig. 1): (a) NW-WNA, (b) ML-WNA and (c) NE-NA. Note the common (different) SST scales for different seasons (regions).

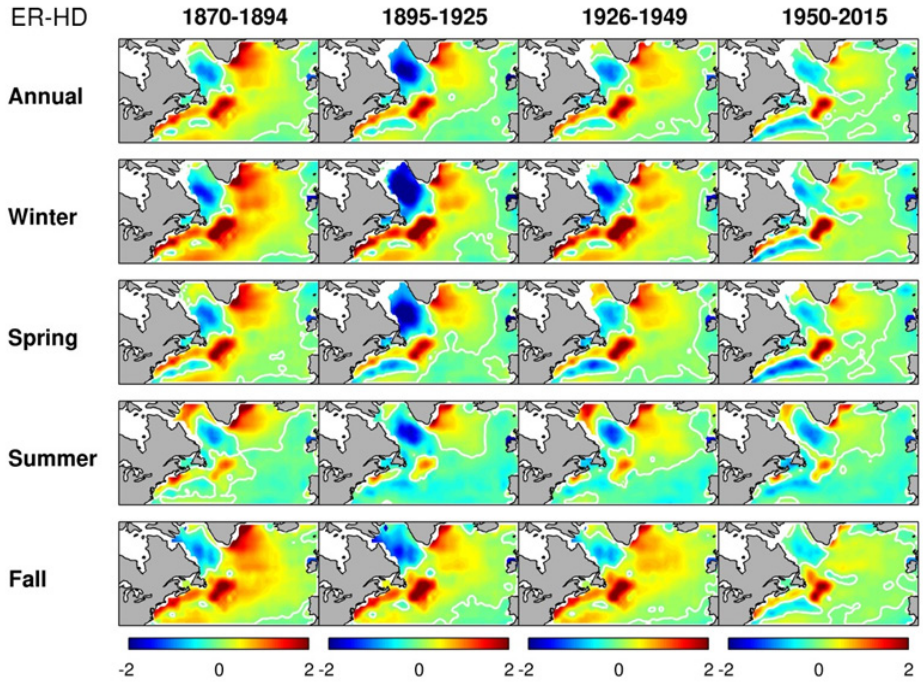

Figure 4. SST difference (ERSST-HadISST1 or ER-HD) maps for annual and seasonal time series during four sub-periods: 1870-1894, 1895-1925, 1925-1949 and 1950-2015. The white contour is the 0ºC isoline. See Fig. SI-1 for the ERSST and HadISST1 fields which have been differenced to obtain these fields.

The agreement between the seasonal time series from ERSST and HadISST1 is generally similar to that between the annual means (Fig. 2), although with some seasonal differences (Fig. 3). Since the mid 1940s, there is close agreement in all seasons in all three regions, except for a tendency for ERSST to be slightly lower than HadISST1 in summer most of the time. In the broadest and least complex NE-NA (Fig. 3c), this good agreement generally extends back to 1870, with the primary discrepancy that ERSST was higher in winter and fall before 1890 and between 1915 and 1945 (probably related to sparser data in these periods). In the NW-WNA (Fig. 3a), it can be seen that the relatively-low ERSST annual means during 1895-1925 arose primarily from differences in winter, spring and, to a lesser extent, summer. In particular, the ERSST winter means of only 2-3ºC over this region, which has deep mixed layers such that SST changes imply large heat content changes (e.g., Loder et al., 2015; Yashayaev & Loder, 2016), raises a concern about the reliability of the ERSST interpolations here (see next subsection).

In contrast, in the ML-WNA (Fig. 3b), the ERSST values in winter and fall prior to the mid 1940s were relatively high in general by 0.5-1.0ºC compared to their HadISST1 counterparts; i.e., the differences had the opposite sign to the above-noted ones in the NW-WNA, although they had the same sign as the smaller ones in the NE-NA. This raises the possibility of a methodological change in the 1940s affecting the winter and fall SST in these areas, and in the ML-WNA in particular. Further support for this can be seen in Fig. C1 (see Appendix C) which

shows time series of the ERSST-HadISST1 differences by season and major region, including a large (~0.7°C) change in the fall and winter differences in the ML-WNA around 1940.

Table 1. Trends in annual- and seasonal-mean SST (°C/decade) averaged over the (a) NW-WNA, (b) ML-WNA and (c) NE-NA, from the two gridded SST datasets (ERSST in blue, and HadISST1 in red) and for the four different historical periods (see Figs. 2 and 3 for plots of these time series). Trends significant at the 95% confidence level are shown in **boldface**, and those at the 90% level in *italics*. The average of the trends from the two datasets is included in black, with an assigned confidence level corresponding to the lower of the individual-dataset ones.

(a)

NW-WNA	1870-2015	1900-2015	1950-2015	1979-2015
Annual	0.01 / -0.00 / 0.00	**0.05** / 0.01 / 0.03	**0.09** / **0.09** / **0.09**	**0.39** / **0.34** / **0.37**
Winter	-0.00 / **-0.02** / -0.01	**0.05**/ -0.02 / 0.02	0.05 / 0.04 / 0.04	**0.38** / **0.29** / **0.33**
Spring	0.00 / **-0.02** / 0.01	**0.07** / -0.00 / 0.03	**0.08** / 0.06 / **0.07**	**0.31** / **0.23** / **0.27**
Summer	**0.04** / **0.03** / **0.03**	**0.08** / **0.06** / **0.07**	**0.13** / **0.14** / **0.14**	**0.50** / **0.46** / **0.48**
Fall	-0.02 / -0.01 / -0.01	0.00 / 0.01 / 0.01	**0.09** / **0.11** / **0.10**	**0.40** / **0.40** / **0.40**

(b)

ML-WNA	1870-2015	1900-2015	1950-2015	1979-2015
Annual	**0.03** / **0.07** / **0.05**	**0.05** / **0.08** / **0.06**	**0.10** / **0.08** / **0.09**	**0.27** / **0.26** / **0.26**
Winter	0.01 / **0.07** / 0.04	*0.02* / **0.08** / *0.05*	**0.09** / **0.09** / **0.09**	**0.24** / **0.26** / **0.25**
Spring	**0.02** / **0.06** / **0.04**	**0.04** / **0.06** / **0.05**	**0.12** / **0.08** / **0.10**	**0.16** / **0.13** / **0.14**
Summer	**0.05** / **0.07** / **0.06**	**0.09** / **0.08** / **0.08**	**0.13** / **0.10** / **0.11**	**0.36** / **0.32** / **0.34**
Fall	**0.03** / **0.08** / **0.05**	**0.04** / **0.08** / **0.06**	**0.08** / **0.06** / **0.07**	**0.31** / **0.32** / **0.31**

(c)

NE-NA	1870-2015	1900-2015	1950-2015	1979-2015
Annual	0.01 / **0.02** / 0.01	**0.03** / **0.03** / **0.03**	**0.07** / 0.03 / 0 05	**0.18** / **0.16** / **0.17**
Winter	-0.00 / **0.02** / 0.01	*0.02* / **0.02** / *0.02*	**0.06** / 0.02 / 0.04	**0.17** / **0.17** / **0.17**
Spring	0.01 / **0.01** / 0.01	**0.03** / **0.03** / **0.03**	**0.05** / 0.02 / 0.04	**0.18** / **0.15** / **0.17**
Summer	**0.02** / **0.02** / **0.02**	**0.05** / **0.04** / **0.04**	**0.08** / 0.04 / 0.06	**0.18** / **0.14** / **0.16**
Fall	0.01 / **0.02** / 0.01	*0.02* / **0.03** / *0.02*	**0.08** / 0.04 / 0.06	**0.21** / **0.19** / **0.20**

Inter-regional differences in the trends, which have overall similarity between the datasets and among the seasons and an increasing magnitude in the more recent periods, can be seen in Table 1. There are notable differences between the ERSST and HadISST1 trends for the two longest periods in the NW-WNA and ML-WNA in most seasons; in particular, for 1900-2015, there are large differences in winter and spring in the NW-WNA and in winter in the ML-WNA. These trends will be discussed further in Section 4, after we examine comparisons between the interpolated and selected other datasets in the next three subsections.

b *Areas with Persistent Differences between ERSST and HadISST1*

Examining Fig. 4 more closely, there are at least six areas in the NWA with notable SST differences between ERSST and HadISST1. These differences are present in all four sub-periods (although with varying magnitude) and have significant persistence across the four seasons. Proceeding from north to south, the areas are (see Fig. 1 and Table SI-1 for their approximate center positions):

The Irminger Basin (IB) to the southeast of Greenland where ERSST is higher than HadISST1 during all sub-periods and seasons, with the difference magnitude notably lower in the last sub-period. This area is somewhat similar to the so-called "warming hole" in the northern NA where historical SST and AT indicate little warming (or weak cooling) since 1900 (e.g., Drifjhout et al., 2012; Hartmann et al., 2013).

The Labrador Sea (LS) between the Labrador and Greenland Shelves where, in contrast to the IB, ERSST is lower than HadISST1 during all sub-periods and seasons. In all four seasons, the largest differences are in the 1895-1925 sub-period and, in the first three sub-periods, the largest differences are in winter. The differences in the last sub-period are smaller and have less seasonality. This subpolar area is centered on the primary deep convection zone in the NWA (e.g., Yashayaev & Loder, 2016).

The Flemish Cap - Newfoundland Basin (FC-NB) area where ERSST is also higher than HadSST1 during all seasons and sub-periods. This area is notably smaller and has smaller difference magnitudes in summer, and somewhat smaller in the last sub-period than in the earlier ones. It is in the portion of the broad transition zone between the subpolar and subtropical gyres where a branch of the Gulf Stream Extension turns northeastward, and includes the more localized deep-water "transition zone" that Buckley and Marshall (2015) have suggested to be the "pacemaker" of AMOC variability.

Three adjacent parallel band-like areas aligned approximately southwest-to-northeast between Cape Hatteras and the Tail of the Grand Bank which we will refer to as the Gulf of Maine - Scotian Slope (GM-SS), Gulf Stream North (GSN) and Sargasso Sea (SarS) areas. ERSST is relatively high in the GM-SS and SarS (compared to HadISST1), and relatively low in the GSN in general. Their spatial structure indicates that they are related to sampling and/or interpolation challenges in the WNA's subpolar-subtropical transition zone along the northern edge of the Gulf Stream (e.g., Watelet et al., 2017). The extents and SST differences of these areas vary with season and sub-period, with their signatures being weakest in summer. Their tri-polar pattern is apparent in the other seasons in all four sub-periods, especially in the most recent ones with the best data coverage.

To examine further the temporal variability of the ERSST-HadISST1 differences in these areas, time series for each of the annual and seasonal means have been extracted for their center positions. The annual series are shown in Fig. 5, and the seasonal series for the LS and SS-GM in Fig. 6, for the IB and FC-NB in Fig. SI-2, and for the GSN and SarS in Fig. SI-3.

To provide an indication of the observational data coverage in the six areas by season and year (and of SST derived directly from observations), we also include in the seasonal plots (Figs. 6, SI-2 and SI-3) estimates of the seasonal means (up to 2014) from the non-interpolated ICOADS R3.0 monthly statistics for the four 2° x 2° grid boxes around each center position (because of no data in some seasons, estimates of the annual means are not included in Fig. 5). Further, to explore whether one of the interpolated datasets appears to be more realistic in cases of discrepancies, we include SST for each area's center position from the SODA 2.2.4 reanalysis product (up to 2010). However, since there have not been many evaluations of this product in the NA, our intercomparisons with ERSST and HadISST1 should not be viewed as a validation of either SODA or these interpolated datasets.

1 ANNUAL MEANS

The annual-mean comparisons (Fig. 5) confirm the persistence of the ERSST-HadISST1 differences throughout the 1870-2015 study period in the IB and FC-NB (both positive), and up to 2008 in the LS (negative) and to the 1990s in the GM-SS (negative; excepting a few years of no difference in the 1940s and 1970s). The differences are larger than for the three broad oceanographic regions (Fig. 2; as expected), with magnitudes ~2°C in the IB up to the 1980s and in the FC-NB up to 1990s, and ~1°C in the GM-SS up to the 1940s. The differences vary more with time in the LS and GSN, with magnitudes 1-3°C in the LS between the 1890s and 1930s, and up to ~1°C in the GSN at times during the 1950s to 1990s. Differences in the SarS are generally positive and smaller, but have magnitudes approaching 1°C prior to the 1940s and 0.5°C since the 1980s. Correlations between the ERSST and HadISST1 annual means are generally only moderate, ranging from r = 0.55 for the IB to r = 0.83 for the GM-SS (where r is the correlation coefficient; see Table SI-2).

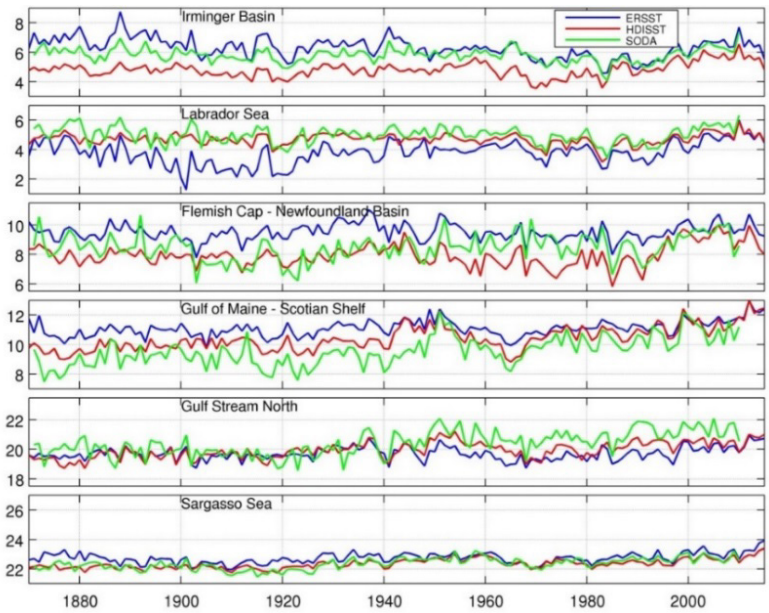

Figure 5. Time series of annual SST from ERSST (blue), HadISST1 (red) and SODA (green) for the approximate center positions (Table SI-1) of six areas: IB, LS, FC-NB, GM-SS, GSN and SarS.

The SODA SSTs are generally between or in the vicinity of those from the interpolated datasets in the three northernmost areas (IB, LS and FC-NB) and in the SarS, but are persistently lower than either in the GM-SS, especially prior to 1940, and higher in the GSN since 1950 (Fig. 5). The poorer SODA agreement in these two areas (also see Figs. SI-2 and SI-3) is consistent with Chepurin et al.'s (2014) findings of poorer agreement with SSH observations from tide gauges along the east coast of North America than elsewhere.

2 SEASONAL MEANS

Comparisons of the ERSST and HadISST1 seasonal time series with the ICOADS seasonal means point to data sparsity being a contributing factor to the ERSST-HadISST1 differences in some seasons and areas. In the LS (Fig. 6a), for example, there are virtually no observations in winter prior to the 1940s, and only scattered observations in spring prior to the 1920s and in fall prior to the 1940s. The relatively-low annual ERSST means during the 1890s to 1930s (Fig. 5) appear to be primarily due to relatively-low ERSST winter means during this period of no winter observations (see the ICOADS time series), supplemented by relatively-low values most of the time in fall and spring as well. The plots (Fig. 6a) and correlations (Table SI-2) indicate that the best agreement between the gridded and ICOADS data occurs in summer when the data coverage (especially prior to 1940) is best and the correlation coefficients for both ERSST and HadISST1 are highest. The SODA seasonal SSTs follow the interpolated series most closely in summer (r ~0.6 for both), and the HadISST series more closely (than the ERSST series) during the data-sparse winters prior to 1940 and springs prior to 1918. Since the shelf waters surrounding most of the LS in fall, winter and spring have lower SST (e.g. Loder et al., 2015; and probably more

observations) than the off-shelf area with the peak differences, it seems likely that the ERSST-HadISST1 differences in this area primarily arise from the winter (and to a lesser degree, fall and spring) ERSST means being too low because of an interpolation artifact. Considering our minimal screening in estimation of the ICOADS seasonal means (see SI), data sparsity may also be a factor in the occasional large ICOADS anomalies in winter and spring in the 1980s and 1990s (Fig. 6a).

A situation somewhat similar to LS in relation to a data-sparsity origin, but notably different in that the differences are of the opposite sign (Fig. 5), occurs in the IB prior to 1980 (Fig. SI-2a). In that case the ERSST seasonal means are consistently higher than the HadISST ones, and more closely follow the ICOADS ones (prior to the 1990s), and the ERSST-HadISST1 differences are largest in winter which has virtually no observations prior to the 1930s. The IB SODA means track both the ICOADS and ERSST means well since the 1920s, and lie between the ERSST and HadISST1 ones in winter between the 1880s and 1920s when there were no observations. ERSST appears to be more reliable than HadISST1 since 1940 in the IB, but the more reliable dataset prior to then (especially in winter) and the source of the differences are unclear.

The FC-NB (Figs. 5 and SI-2b) is similar to the IB in that its ERSST means are consistently larger than the HadISST1 ones and follow the ICOADS means more closely since 1910 in all seasons, except summer in FC-NB when all the means track closely. On the other hand, these areas differ in that the ICOADS means for the FC-NB tend to be larger than those from both interpolated datasets in winter and spring before 1910, and the SODA means for the FC-NB in spring and summer prior to 1930 have substantial differences from the others. These areas are discussed further in the SI.

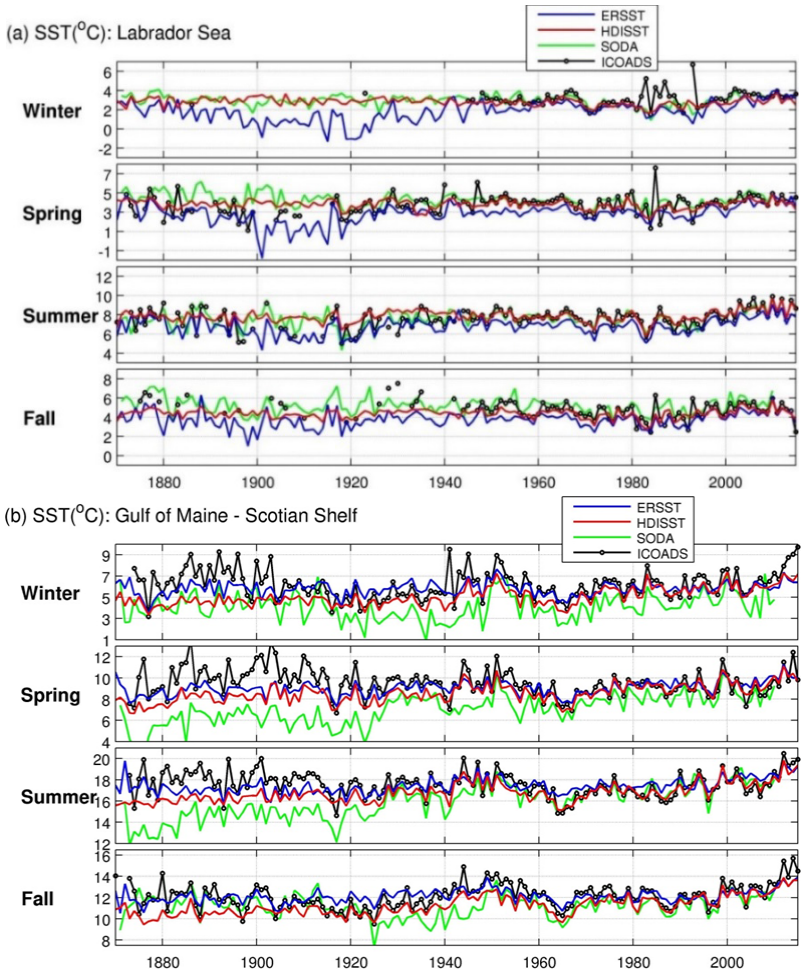

Figure 6. Time series of seasonal SST from ERSST, HadISST1, ICOADS and SODA for the (a) LS and (b) GM-SS.

The GM-SS area (Fig. 6b) contrasts the LS (Fig. 6a) in having observational (ICOADS) data back to the 1870s in all seasons, ERSST-HadISST1 differences of opposite sign, and a reduced seasonal variation in the differences. The latter may reflect that the differences' origin is related to a combination of the spatial distribution of available data in relation to the permanent shelf-slope front (e.g., Loder et al., 1998) and the different interpolation methodologies. Like the FC-NB in winter and spring (Fig. SI-2b), there is a tendency prior to 1920 for the ICOADS GM-SS means to be higher than the ERSST and HadISST1 ones in winter, spring and summer, with enhanced interannual variability which probably reflects inadequate observational data coverage of these high-gradient and high-variability areas (e.g., Fig. SI-1). Going back to one of our original motivations (Loder et al., 2012; LW2015), namely why trends from HadISST1 for this area were larger than those from ERSST, the comparisons with ICOADS in Fig. 6b indicates that underestimated HadISST1 values prior to 1940 is the primary reason for the discrepancy. But the comparisons also indicate that there needs to be critical evaluation (e.g., using non-interpolated datasets such as ICOADS) of whether the ERSST values in this area prior to 1960 are also

underestimates (also see next two subsections), which would have implications for the magnitude of long-term trends in the area (e.g., Friedland & Hare, 2007).

Figure 6b also indicates that the relatively-low SODA annual means (Fig. 5) in the SS-GM receive contributions from suspiciously-low SODA SSTs in different seasons during different periods: in particular, in spring and summer prior to the 1930s, and in fall and winter during the 1930s and 1940s. Notable differences between the SODA and observational SST data elsewhere in the ML-WNA are also apparent in the seasonal comparison plots for the GSN and SarS areas (Fig. SI-3). The poor agreement of the SODA time series with the others prior to 1940 in these three areas and in the FC-NB indicate a potential problem with the simulation in that time period at mid latitudes in the WNA (as noted earlier). Figure SI-3 also indicates good approximate agreement of both ERSST and HadISST1 with ICOADS in the SarS area in all seasons since 1940 (but differences earlier), and the same in the GSN except for slightly higher ERSST (than HadISST1 or ICOADS) values in winter and spring.

c *Comparisons with Observations from DFO Monitoring Sites*

Time series of the annual and seasonal SST means from P5, EB and S27 are shown in Fig. 7, together with the corresponding SST means from the closest grid points in the ERSST and HadISST1 datasets. Correlations between each DFO series and each corresponding ERSST and HadISST1 series are presented in Table 2, and the trends in each series are presented in Table 3, all computed over the period of each DFO series. The correlations of each pair of ERSST and HadISST1 series are included in Table 2 as reference levels, and the correlations and trends for the adjusted HH series and the ERSST and HadISST1 series from the same grid boxes as used for the EB comparisons (but now over the duration of the HH records) are included in the tables (see Appendix B for plots of, and more information on, the HH time series).

The annual and summer time series, correlations and trends for P5, EB and S27 are very similar in features and magnitude to those in LW15 (their Fig. 3 and Tables 2 and 3), with a small increase in most of the correlations and trends for the updated time series discussed here. The latter can be attributed to the occurrence of record or near-record warm SSTs in most seasons at all three sites in 2012, and near- or above-average SSTs in most seasons at these sites during 2013-2015 (Colbourne et al., 2016; Hebert et al., 2016). An additional contributor to the increased ERSST trends is probably the adjustment applied in version 4 that resulted in global trends over the last 19 years increasing by ~0.1°C per decade (Hausfather et al., 2017).

The most notable features of the plots (Fig. 7) are (i) the interannual-to-decadal scale variability in all of the time series with amplitude generally exceeding the long-term change over the record, (ii) mean offsets in the DFO series (compared to the interpolated ones) in some seasons, and (iii) differences between the ERSST and HadISST1 series that often are different from those discussed above for the major regions and have magnitudes comparable to those between the DFO and gridded series. In

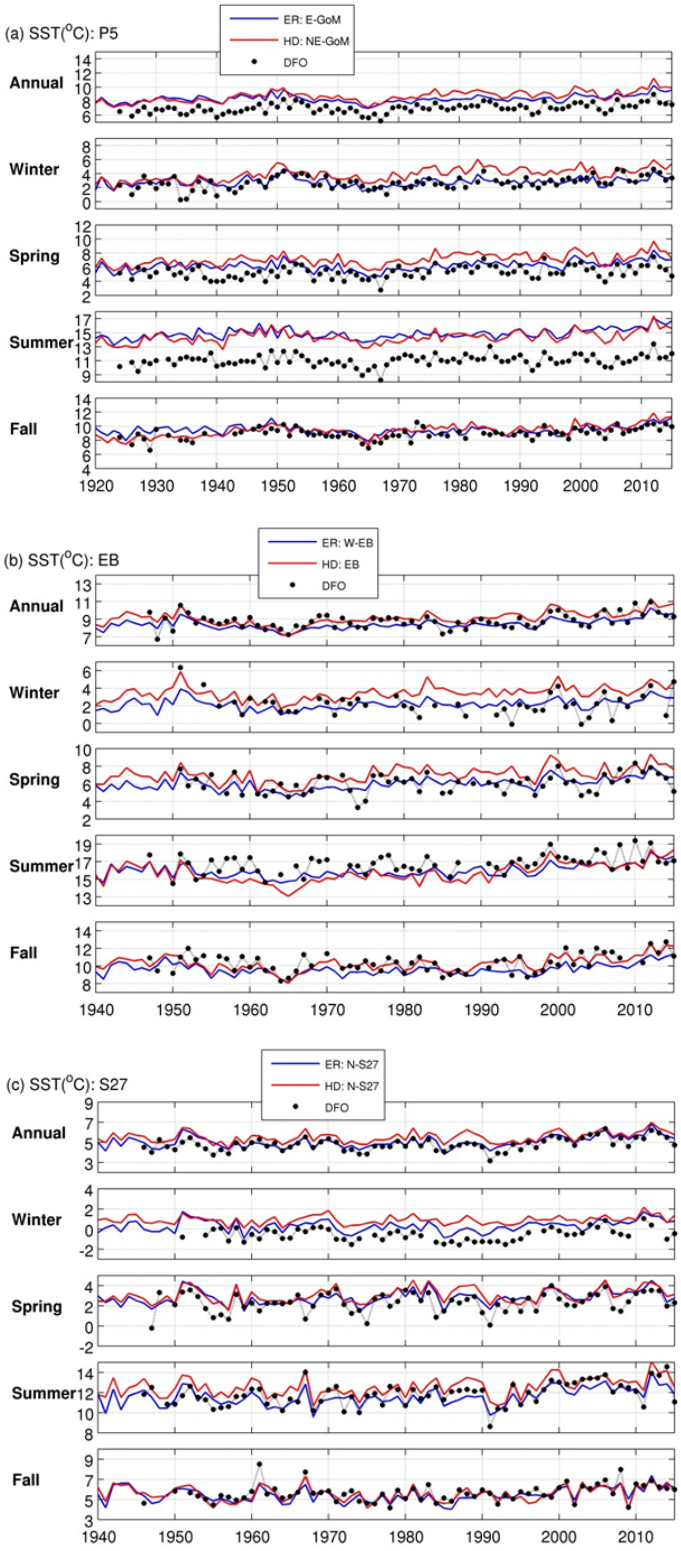

Figure 7. Annual- and seasonal-mean SST from DFO monitoring (black dots) at (a) P5, (b) EB and (c) S27, together with the corresponding time series from the nearest grid boxes in ERSST (ER, blue) and HadISST1 (HD, red). Seasonal DFO means were retained when there were observations in at least 2 (out of 3) months, and annual means for observations in at least 2 seasons.

particular, the ERSST and HadISST1 values for the northeastern GoM grid points used in the P5 comparisons are higher than the P5 values by ~4°C in summer and ~2°C in the spring and annual series, with little mean difference in winter and fall. This is reasonable considering the spring-summer stratification in the offshore GoM and the stronger tidal mixing in the Bay of Fundy and Passamaquoddy Bay resulting in lower SST there (e.g., Loder and Greenberg, 1986). There are also smaller (~1°C) mean offsets at EB in summer and at S27 in winter, with the interpolated values lower and higher than the DFO ones in these respective cases, which may also partly reflect horizontal structure in SST in the real ocean.

Another noteworthy feature of the EB time series plots (Fig. 7b) is that the ERSST values estimated for this position are consistently lower than the HadISST1 values, in contrast to the results of the previous subsection where the ERSST values for the broader GM-SS area are consistently higher than the HadISST1 ones. The low EB values in ERSST are related to the latter's coarse representation of the coastline which does not include the Nova Scotia land mass separating the Scotian Shelf and Gulf of St. Lawrence, thereby affecting the interpolation. This points to a particular problem with the coarser and smoother ERSST dataset in representing the spatially-variable SST in a coastal region with complex geometry like that off Atlantic Canada.

For P5, EB and S27, the correlations of the DFO series are generally higher with HadISST1 (r = 0.62-0.85) than with ERSST (r = 0.49-0.81), and statistically significant in all seasons (Table 2). These correlations are only slightly lower than those between the ERSST and HadISST1 series (r = 0.69-0.94) in most cases, indicating common variability signals in the different series in each case. There is little seasonal variation among the correlations for a particular site, both for those between the interpolated time series and those between them and the DFO ones. In most cases (sites and dataset pairs), the correlations for the annual series are larger than those in any particular season, probably reflecting some averaging over the higher-frequency variability present in the individual seasonal series.

Table 2. Correlation coefficients between annual- and seasonal-mean SST time series from four DFO monitoring sites, and the corresponding time series from each of ERSST and HadISST1 at nearby grid points (see Fig. 7 for plots of the P5, EB and S27 series, and Fig. B1 for the HH series). Correlations between the various ERSST and HadISST1 time series (during the period of the DFO series; Table 3) are also included. The dataset pairs are referenced by the column header abbreviations as follows: E-D: ERSST vs DFO; H-D: HadISST1 vs DFO; and E-H: ERSST vs HadISST1. Correlations significant at the 95% level are shown in **boldface**, and those at the 90% level in *italics*.

Site	Annual			Winter			Spring			Summer			Fall		
	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H
P5	**0.74**	**0.83**	**0.81**	**0.63**	**0.69**	**0.73**	**0.63**	**0.72**	**0.84**	**0.49**	**0.67**	**0.76**	**0.62**	**0.74**	**0.69**
EB	**0.73**	**0.71**	**0.94**	**0.61**	**0.65**	**0.94**	**0.61**	**0.62**	**0.92**	**0.58**	**0.62**	**0.88**	**0.69**	**0.69**	**0.88**
HH	**0.60**	**0.52**	**0.92**	**0.61**	**0.58**	**0.83**	**0.64**	**0.65**	**0.91**	**0.49**	**0.53**	**0.86**	**0.49**	**0.47**	**0.79**

Site	Annual			Winter			Spring			Summer			Fall		
	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H	E-D	H-D	E-H
S27	**0.79**	**0.85**	**0.88**	**0.71**	**0.69**	**0.82**	**0.66**	**0.72**	**0.86**	**0.81**	**0.85**	**0.92**	**0.75**	**0.78**	**0.86**

The correlations between the HH time series and the interpolated ones are generally lower (r = 0.49-0.65) than those for the three primary DFO sites, particularly in fall. Time series plots of the EB and HH annual and seasonal means (Fig. B1) show mean offsets in spring and summer, as expected between coastal and offshore SST, but also substantial differences in their variability and trends. It appears that an additional forcing is affecting the HH SST variability compared to those affecting the offshore Scotian Shelf, probably wind-driven coastal upwelling (e.g., Petrie et al., 1987) and other factors affecting exchange between the HH coastal embayment and the shelf (e.g., Shan et al., 2011).

Table 3. Trends (in ºC/decade) in annual- and seasonal-mean SST time series from DFO monitoring sites and nearby ERSST and HadISST1 grid points, over the period of each DFO series (start year indicated for each site and season). See Figs. 7 and B1 for plots of these series, including earlier years for the gridded datasets. Trends significant at the 95% level are shown in **boldface**, and those at the 90% level in *italics*.

Site/Dataset	Annual	Winter	Spring	Summer	Fall
Prince 5	1924-	1924-	1924-	1924-	1924-
DFO	**0.11**	**0.15**	**0.11**	**0.07**	**0.12**
ERSST	**0.08**	*0.04*	**0.12**	**0.13**	0.03
HadISST1	**0.17**	**0.20**	**0.16**	**0.13**	**0.19**
Emerald Basin	1947-	1951-	1951-	1947-	1947-
DFO	**0.12**	-0.08	*0.14*	**0.21**	*0.13*
ERSST	**0.15**	**0.13**	**0.23**	**0.21**	**0.13**
HadISST1	**0.23**	**0.19**	**0.25**	**0.37**	**0.23**
Halifax Harbour	1926-	1926-	1926-	1926-	1926-
DFO	*0.05*	0.04	0.03	0.06	**0.08**
ERSST	**0.10**	**0.07**	**0.13**	**0.14**	**0.06**
HadISST1	**0.18**	**0.18**	**0.17**	**0.18**	**0.19**
Station 27	1946-	1951-	1947-	1946-	1946-
DFO	**0.12**	0.05	*0.10*	**0.23**	**0.12**
ERSST	**0.12**	0.04	**0.11**	**0.17**	**0.13**
HadISST1	**0.10**	0.05	**0.10**	**0.15**	**0.09**

The HH and EB SST series, and their relation to annual and seasonal AT records from nearby Halifax and Sable Island respectively, are discussed further in Appendix B. There is some similarity in the variability and trends in the series from all four sites (Fig. B1), but also notable differences consistent with an additional process affecting HH SST. The agreement between the HH SST series and the other SST and AT series has been improved with the implementation of an adjustment for a methodological and depth sampling change at HH in the 1990s.

The annual-mean trends at the other three DFO sites are 0.11°C per decade for P5 and 0.12°C per decade for EB and S27, which are statistically significant (Table 3). The seasonality of the trends differs with the site. At P5 (back to 1924), the trends are significant in all seasons and highest in winter when the magnitude (0.15°C per decade) is about twice that in summer. At S27 (back to only 1946), the trends are highly significant in summer and fall only, with the highest (summer) trend (0.23°C per decade) about twice the annual mean and with the weakest trend in winter. At EB, the only highly significant trend is in summer (0.21°C per decade), with marginally significant trends in spring and fall with magnitudes comparable to that in the annual mean, and with a non-significant negative trend in winter.

As described in Appendix B, there are also no significant trends in winter HH SST and winter Halifax and Sable Island AT since circa 1950, in part associated with a period with relatively warm winters in the early 1950s. On the other hand, the winter EB time series from the interpolated datasets show significant positive trends of 0.13-0.19°C per decade over this period (Table 3), but it is questionable whether these are reliable in view of the coastline representations (especially in ERSST) and interpolations used in these datasets. The lack of a winter trend in observed AT and SST from the central and eastern Scotian Shelf (Halifax and offshore) since 1950 at least partly reflects the strong decadal-scale and interannual variability in the region. Note, however, that there are significant trends of 0.09 and 0.12°C per decade, respectively, in the longer records of Halifax AT dating back to 1872 and Sable Island AT back to 1898 (Table B1), consistent with anthropogenic climate change.

d *Comparison with SST from U.S. Coastal Lighthouses and Lightships*

In this subsection, a brief comparison between annual means from ERSST and HadISST1 is presented, and those from SL2010's study of long-term SST variability along the U.S. east coast based on coastal lighthouse and lightship data, supplemented by buoy and shore-based observations. Figure 8 shows the ERSST and HadISST1 time series from grid points in the central GoM and the northeast and southwest parts of the MAB (MAB-NE and MAB-SW, respectively). The GoM grid point approximates the area for which SL2010 show a GoM annual anomaly time series in their Fig. 8, and the MAB grid points approximate its cooler (northeast) and warmer (southwest) parts. Substantial coherence of the interannual to decadal variability in all of these gridded data time series is apparent, although there are clear differences in the record-mean values for the different areas (and to a lesser extent for the different datasets).

Also shown in Fig. 8 are approximations to the SL2010 time series for their GoM and MAB areas, obtained by adding estimates of the long-term mean SSTs for these areas (from SL2010's Fig. 4a) to the anomaly time series in their Fig. 8. The resulting annual SST time series approximate the lower range of the offshore interpolated data series which is not surprising for coastal data (although a precise comparison is not appropriate considering the rough estimation here of the SL2010 long-term means). Encouragingly there is substantial similarity in the pentadal-to-decadal variability in the SL2010 and interpolated data series (although there is an apparent lag in the 1950s peak in the coastal series compared to that in the interpolated data).

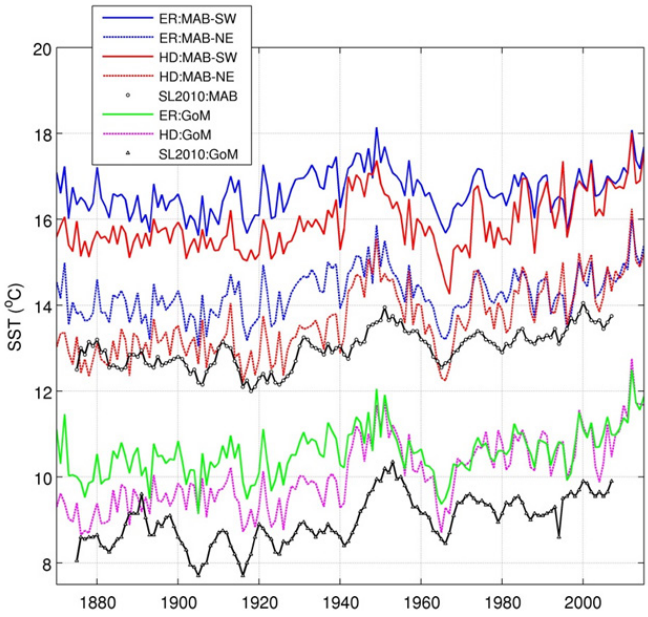

Figure 8. Annual-mean SST (coloured curves, see legend box) for the GoM and two parts (southwest and northeast) of the MAB from each of ERSST (ER) and HadISST1 (HD). The small circles (MAB) and triangles (GoM) connected by black curves are the annual coastal SST anomalies from SL2010's Fig. 4 with approximate climatological annual means estimated from their Fig 7 added (because the latter are only approximate estimates, focus should be on the time series' variability rather than their mean values).

SL2010 found long-term trends over 1870-2007 of 0.1±0.03ºC and 0.07±0.03ºC per decade in their annual anomalies for the GoM and MAB regions, respectively (their Fig. 8). These trends are shown in Table 4, together with those for the same period from the various interpolated time series in Fig. 8, and averages of the MAB-NE and MAB-SW trends (thereby approximating the MAB) and of the ERSST and HadISST1 trends for each area. For the GoM, the trends over 1870-2007 are 0.05ºC and 0.11ºC per decade from ERSST and HadISST1, respectively, bracketing the SL2010 value per decade from ERSST and HadISST1, respectively. For the MAB, the trends are 0.05ºC and 0.10ºC per decade, again bracketing the SL2010 value.

Table 4. Trends (in ºC/decade) in annual coastal SST for the GoMand MAB from SL2010, and in annual SST for the GoM and the southwest (SW) and northeast (NE) parts of the MAB from each of ERSST and HadISST1 (see Fig. 7 for plots of the time series). Averages of the ERSST and HadISST1 trends (all of which are statistically significant at the 95% level) are shown in the last row, and averages of the MAB-SW and MAB-NE trends in the last column for comparison with the MAB trends from SL2010.

Source	GoM	MAB-SW	MAB-NE	MAB
SL2010	0.1 ±0.03	-	-	0.07 ± 0.03
ERSST	0.05	0.04	0.05	0.05
HadISST1	0.11	0.07	0.13	0.10
(ER+HD)/2	0.08	0.05	0.09	0.08

While the trend differences are not statistically significant, the higher trends in the HadISST1 time series (by a factor of 2 compared to those in the ERSST time series) are consistent with those found across the ML-WNA region in Section 3a (Table 1) and with the relatively-low HadISST1 annual means prior to 1940 in the SS-GM area (Figs. 4 and 5) in Section 3b. This corresponds to the HadISST1 trends being high compared to those found by SL2010 (by up to 0.03ºC per decade, for the MAB), while the ERSST trends are low by comparable amounts (by up to 0.05ºC per decade, for the GoM). On the other hand, with averaging across the two gridded datasets, the trends are well within the uncertainties in the SL2010 values.

SL2010 also computed a trend of 0.11±0.04ºC per decade during 1880-2005 as a proxy for that in LS SST. This contrasts the non-significant weak trends shown in Table 1 for the NW-WNA and found in Section 3b for the LS during 1870-2015 from HadISST1 (which appears to be the more reliable of the gridded datasets), and the non-significant negative trend found by LW2015 for 10-30m temperature in the central LS during 1928-2011. These differences among upper-ocean LS temperature trends from various sources raise questions about the robustness of SL2010's conclusions that advection of SST anomalies from the LS is the primary factor influencing SST variability in the GoM-MAB. This is reinforced by recent studies of a more local origin for SST variability in the latter area (Chen et al., 2015; Brickman et al., 2018).

4 Long-term and recent trends in different seasons

The ERSST, HadISST1 and ICOADS intercomparisons, and the comparisons with the DFO and US monitoring time series, in the previous section and SI indicate reasonable overall agreement in the different time series, but also significant differences in some areas and periods which vary with season and appear to be at least partly related to data sparsity. With these in mind, we examine the patterns in the trends in this section, referring to spatial maps for the different historical periods and seasons (Fig. 9), and the trend values in Tables 1, 3 and 4.

In 1870-2015, there is little apparent seasonal difference among the trend patterns (Fig. 9) in each of ERSST and HadISST1, with the exceptions that the offshore area of cooling southeast

of Greenland does not extend as far south in summer as in the other seasons and, in ERSST, the warming area off Labrador extends further south in summer. The trend magnitudes in the major regions (Table 1) have a generally consistent pattern of being largest in summer, and smallest in winter and spring. The primary differences between the datasets are larger trends in ERSST in the LS which are an artifact of the ERSST underestimation there prior to 1940, and larger trends in the GM-SS in HadISST1 which are at least partly an artifact of the HadISST1 underestimation there prior to 1940. Using HadISST1 for the NW-WNA, and the averages of ERSST and HadISST1 for the ML-WNA and NE-NA, respectively, rough estimates of the annual SST trends over this period are 0.05ºC per decade in the ML-WNA and not significantly different from zero in the NW-WNA and NE-NA.

In 1900-2015, the enhanced summer warming and reduced winter warming (including offshore cooling) are also the primary seasonal changes in both datasets. The differing trends between the datasets in each of the LS and GM-SS, due to methodological issues as just noted, are again the primary differences. Using the same representative datasets as for 1870-2015, rough estimates of the annual trends since 1900 are 0.06ºC per decade in the ML-WNA, 0.03ºC per decade in the NE-NA, and not significantly different from zero in the NW-WNA.

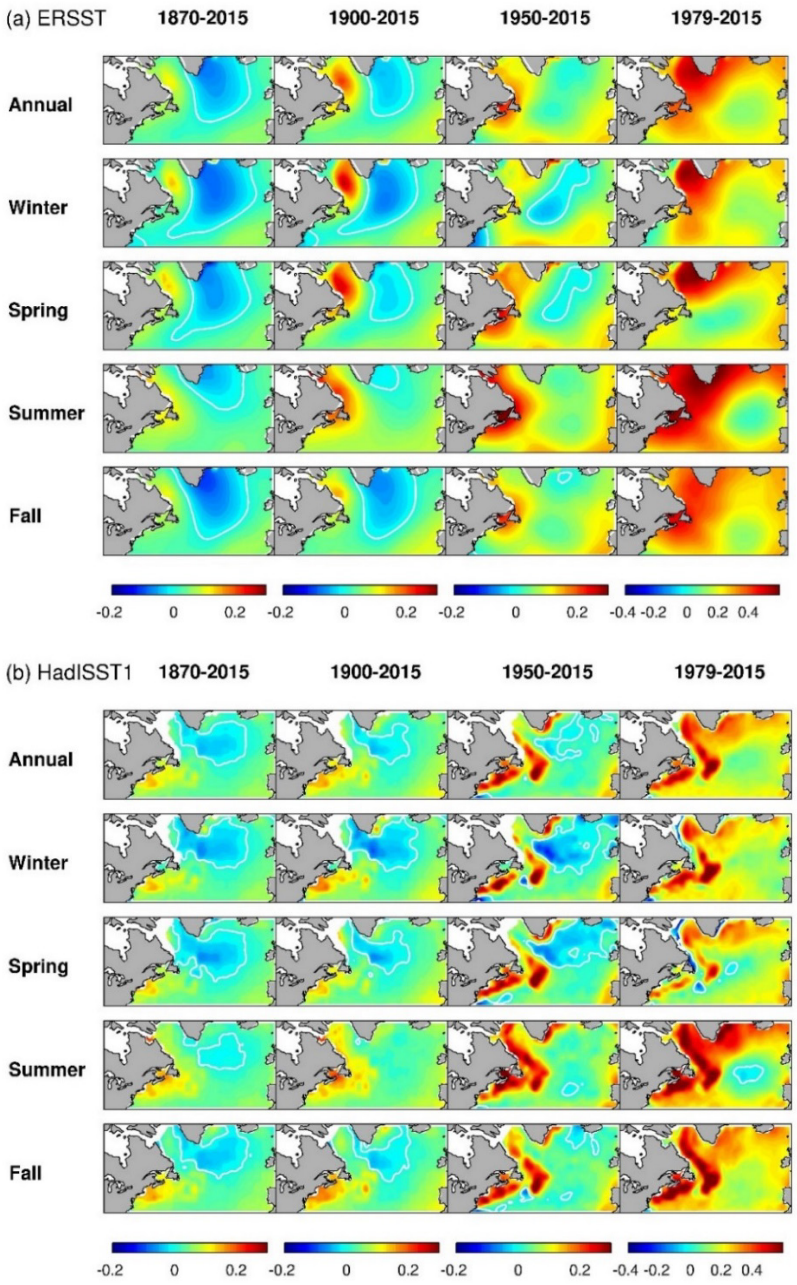

Figure 9. Maps of SST trends (in ºC per decade) in annual- and seasonal-mean (a) ERSST and (b) HadISST1 in four different periods. Note the different trend scale in the 1979-2015 period (right column) in which the trends are larger than in the three longer periods. The white contour is the 0 isoline. The domain shown here is slightly reduced on the north, east and south sides, and significantly expanded on the west side compared to the analysis domain (Fig. 1), for display purposes.

In 1950-2015 (when the data sparsity problems are reduced), the most prominent seasonal difference in both datasets is the more widespread off-shelf warming in summer and fall, in contrast to the widespread off-shelf cooling in winter and spring southeast of Greenland

(Fig. 9). There is also enhanced warming pattern over the entire shelf between the MAB and LS in summer in both datasets, and an indication of greater seasonality in ERSST than HadISST1 in some shelf areas – notably little or no warming in ERSST in the MAB in winter and fall. The reliability of this feature is unclear (note that only annual means were available from SL2010). Because of winter-spring sea ice in the Labrador Shelf region, it is not possible to comment reliably on differences between ERSST and HadISST1 there, but there is a clear indication that the offshore cooling area in winter and spring extends further into the LS in HadISST1 than in ERSST. Using the averages of the trends from the two gridded datasets as indicators (Table 1) , rough estimates of the annual trends since 1950 are 0.09ºC per decade in the NW-WNA and ML-WNA, and 0.05ºC per decade (but only marginally significant) in the NE-NA. On the scale of these regions, the seasonality is greatest in the NW-WNA and weakest in the NE-NA. The DFO observations from S27 are consistent with this, but those from P5 and EB indicate seasonal differences locally, with P5 indicating that the trend in the eastern GoM is largest in winter, and EB (and the ATs) indicating no trend in winter since 1950.

In 1979-2015, the differences between the interpolated datasets' trend patterns can be largely attributed to the greater smoothing in ERSST.v4 (Fig. 9). The primary seasonality in the WNA is lower trends to the west and south of the Grand Bank in spring (in both datasets). There is also notably reduced warming west of the Scotian Shelf in winter in ERSST (compared to other seasons or HadISST1), indicating that some regional differences between the datasets exist in recent decades (as well as prior to 1950) (also see Hausfather et al., 2017). During this period, the average annual trends are notably larger than for the longer periods: ~0.4ºC per decade in the NW-WNA, ~0.3ºC per decade in the ML-WNA, and ~0.2ºC per decade in the NE-NA. While the larger trends in this period probably partly reflect increased anthropogenic global warming, a contribution from the apparently natural multi-decadal variability seen in Figs. 2, 3, 5, 6, 7 and 8 also needs to be considered (see next section).

5 Modes of variability in different seasons

LW2015 reported on EOF analyses on the annual and summer means from ERSST, HadISST1 and COBE up to 2011 for different start years and domains. They found remarkable agreement among the three leading modes across domains, datasets and periods, with these three modes combined accounting for 60-70% of the total variance in each case.

Using a larger NA domain than the one used here and without prior de-trending, LW2015 found that the leading mode in ERSST and HadISST1 during 1870-2011 in both the annual and summer means accounted for 30-40% of the variance, with its low-frequency (PC) variability resembling a combination of the AMO index and an increasing long-term warming like that in GMST (Fig. 2b). To focus on natural variability, they carried out additional EOF analyses for the annual and summer means for the period 1900-2011, with linear trends removed from the time series. The three leading modes in these cases accounted for 55-65% (41-51%) of the variance

for the annual (summer) means in the three datasets. The modes resembled those from the 1870-2011 analyses (without de-trending), with the exceptions of a notable change in the spatial structure of the leading mode and the disappearance of its trend. LW2015 then used correlation analyses and literature references to suggest that these three modes reflected the AMO, a large-scale response to winter NAO forcing, and an ocean circulation variability mode identified in the Wang et al. (2015) model hindcast study, respectively.

a *For 1870-2015 without De-trending*

Figure 10 shows the EOF patterns and PC time series for the three leading modes in the annual and seasonal means from ERSST and HadISST1 for 1870-2015, without prior de-trending. These are essentially seasonal and northern NA zooms and updates on the analyses in Figs 4 and 5 of LW2015, with the absence of prior de-trending resulting in trends in some of the PCs. LW2015 presented the EOFs and PCs in the order of their variance explained and found this order for modes 2 and 3 in the summer means to be different than in the annual means (for both datasets). In our updated analyses we find the ordering of these modes to have that same annual-versus-summer difference. The ordering of the modes in the other three seasons in HadISST1 is the same as for the annual modes, while the ordering in ERSST is the same in fall as for the annual modes but the same in winter and spring as for the summer modes.

To help in the discussion, the modes in Fig. 10 are displayed in an order based on their physical interpretation which we refer to as "response" modes, using the nomenclature EOF-R1, -R2 and -R3 and PC-R1, -R2 and -R3. Response mode R1, with a low-frequency variation resembling the AMO plus a trend, is the EOF1 (with greatest variance explained) in all seasons in both datasets, accounting for 24-37% of the variance for ERSST and 25-40% for HadISST1 (in the annual and different seasonal means). As expected, the variances explained by the annual and summer means are similar to those (32-40%) in LW2015. In both datasets, the percentage explained is highest in summer, consistent with amplified warming in the presence of shallow summer surface mixed layers, while the lowest percentage is in winter in ERSST and in spring in HadISST1. The most notable differences among the EOF-R1 patterns are the different latitudinal structures between ERSST (subpolar peak amplitude) and HadISST1 (mid-latitude peak amplitude), and the smoother ERSST patterns, rather than any seasonal differences (Fig. 10a,b). The most notable differences among the PC-R1 patterns (Fig. 10c,d) are the reduced long-term trend and the more pronounced multi-decadal cycle in the early decades in ERSST. The peak amplitude of ERSST EOF-R1 in the LS and its minimum PC-R1 values in the early 1900s suggest that the apparent interpolation problem in ERSST in that area and period discussed in Section 3b is contaminating this mode. Similarly, the peak amplitudes of HadISST1 EOF-R1 and EOF-R2 in the ML-WNA raises a question of whether these modes are influenced by the apparent methodological change in HadISST1 around 1940 (see below).

After de-trending (to reduce such contamination), the annual PC-R1s for both ERSST and HadISST1 are highly correlated ($r = 0.8$ and 0.7, respectively) with the de-trended AMO

index (Fig. 2b). Note that the correlations of 0.95 reported for these indices in LW2015 were for the AMO and PCs without de-trending. The de-trended PC-R1s are also moderately correlated (r = 0.6 and 0.7, respectively) with the de-trended GMST anomaly time series, with these values reduced from the LW2015 values of ~0.8 due to the de-trending.

Response mode 2 (R2) in these analyses, resembling the NAO-like mode in LW2015 with out-of-phase peak amplitudes in the ML-WNA and area southeast of Greenland, accounts for 13-19% and 13-21% of the variances in ERSST and HadISST1, respectively (Fig. 10). This is slightly more than the

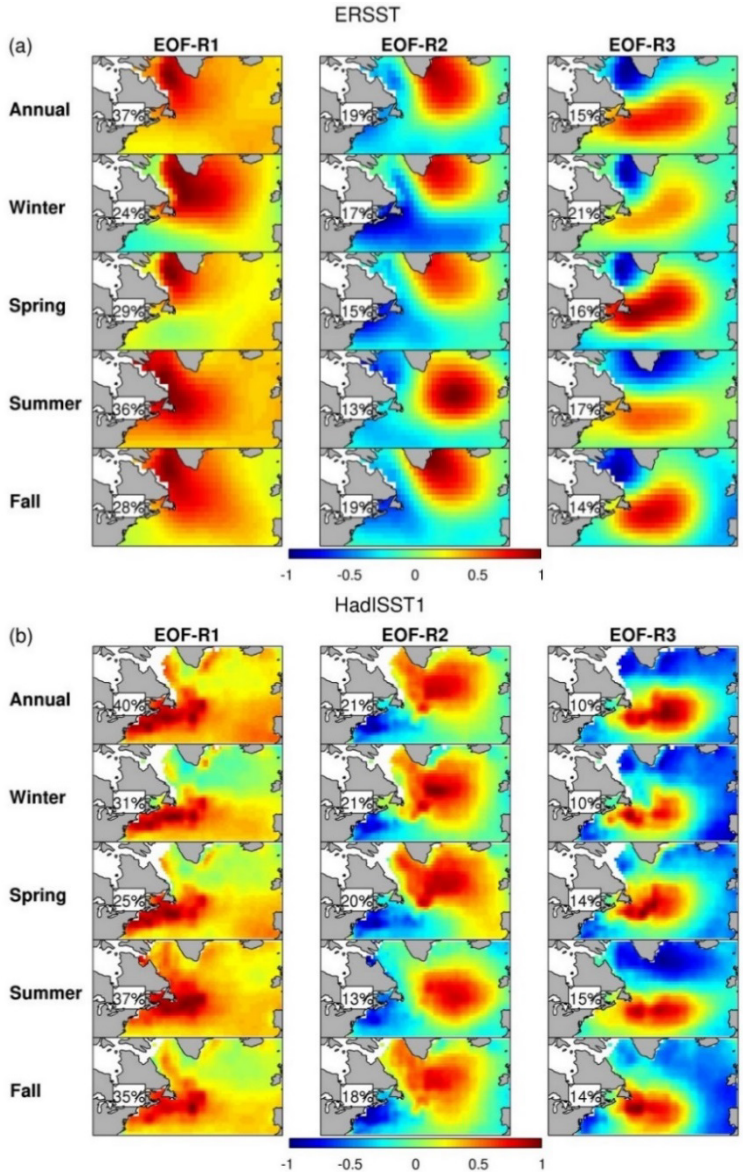

Figure 10. (a,b) EOF patterns (°C normalized by the peak value in each case) for the leading three modes of the annual and seasonal means of (a) ERSST and (b) HadISST1 for 1870 to 2015, with prior de-meaning but without prior de-trending. The modes (columns for EOFs) are organized according to their physical interpretation (see text), and referred to as "response" modes (R1, R2 and R3). The percentage of the total variance in each annual/seasonal dataset explained by each mode is indicated in the small box in each EOF panel. The domain shown here is reduced compared to the analysis domain (Fig. 1).

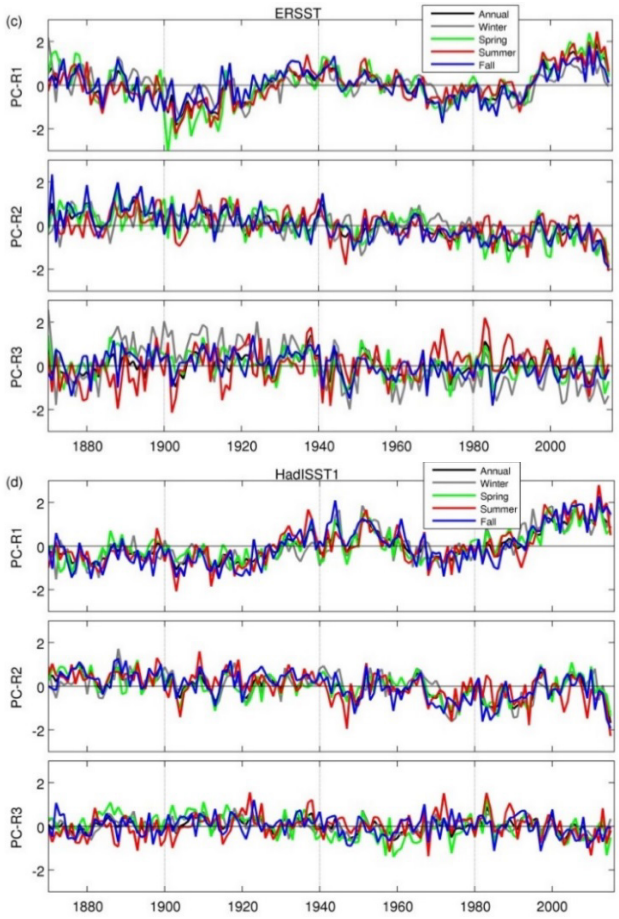

Figure 10. (c,d) Scaled PCs for the leading three modes of the annual and seasonal means of (c) ERSST and (d) HadISST1 for 1870 to 2015, with prior de-meaning but without prior de-trending. The modes (rows for PCs) are organized according to their physical interpretation, referred to as "response" modes.

10-16% in LW2015. The highest percentages are in fall and winter in ERSST and in winter and spring in HadISST1, and the lowest percentages are in summer, generally consistent with the seasonality expected for the winter NAO influence. The PC-R2s for this mode show a long-term negative drift in both datasets which, considering the EOF-R2s' spatial structure, might be interpreted as a contribution to amplified long-term warming in the ML-WNA (e.g., associated with a northward shift of the Gulf Stream; Wu et al., 2012; Yang et al., 2016) and a reduced long-term warming in the subpolar region (e.g., Yashayaev and Loder, 2016). On the other hand, it could be argued that the PC-R2s' drift includes a negative shift around 1940 which might be a reflection of the shift in the winter and spring ERSST-HadISST1 differences at that time in the ML-WNA (Fig. C1), pointing to a methodological influence.

Response mode 3 (R3), accounting for 14-21% of the variance in ERSST with a peak in winter and 10-15% of that in HadISST1 with a peak in summer, has similar spatial structure to the ocean circulation variability mode in LW2015, with a notable centre-of-action in the gyre-gyre transition zone to the east of Newfoundland and the Grand Bank (Fig. 10).

b *For 1900-2015 with Prior De-trending*

We next present the results of similar EOF analyses to those above, with the exceptions that the period is limited to 1900-2015 and linear trends have been removed prior to the analyses. The EOF patterns and PC time series are shown in Fig. 11, with the same response-mode ordering described above. These are essentially seasonal zooms and updates on the results in Figs 8 and 9 of LW2015. The correlations of the annual PCs with selected forcing indices are presented in Table 5, and the more significant ones are broken out by season in Table 6.

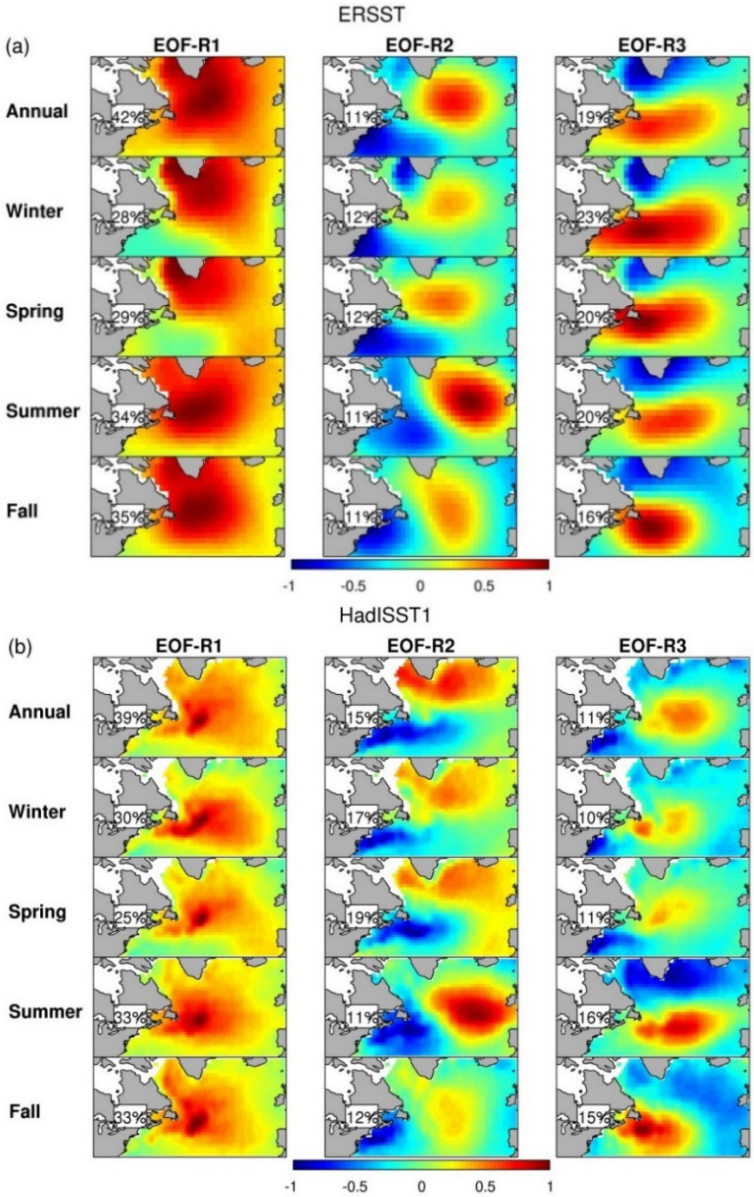

Figure 11. (a,b) EOF patterns (ºC normalized by the peak value) for the leading three modes of the annual and seasonal means of (a) ERSST and (b) HadISST1 for 1900 to 2015, with prior de-meaning and with prior de-trending. The modes are organized according to their physical interpretation (like in Fig. 10). The percentage of the total variance explained by each mode is indicated.

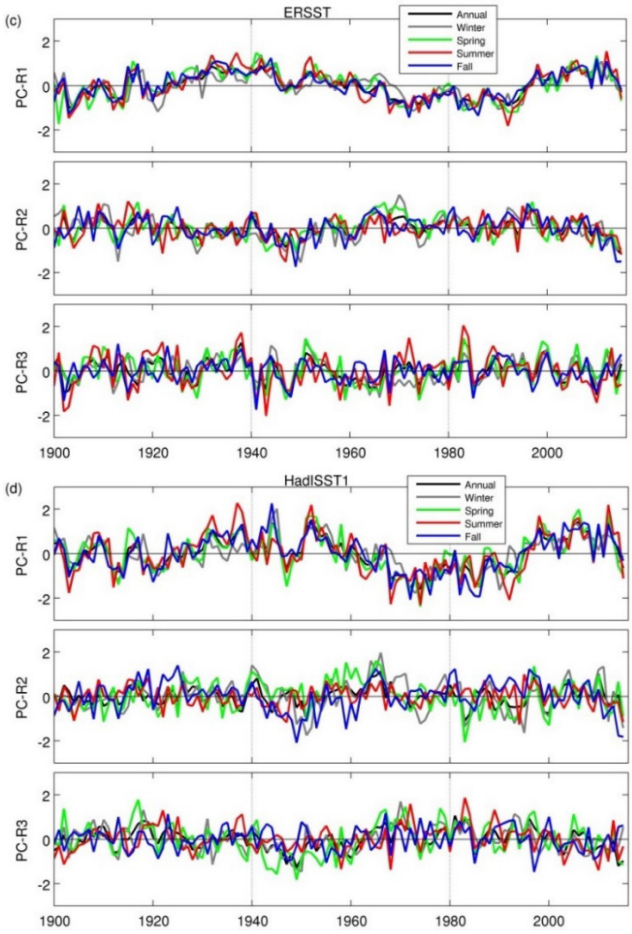

Figure 11. (c,d) Scaled PCs for the leading three modes of the annual and seasonal means of (c) ERSST and (d) HadISST1 for 1900 to 2015, with prior de-meaning and with prior de-trending. The modes are organized according to their physical interpretation.

The leading mode (variance explained) in all cases is again the AMO-like mode R1, accounting for 28-42% and 25-39% of the variance for ERSST and HadISST1 (Fig. 11a,b). The EOF-R1 and PC-R1 patterns in all seasons and both datasets resemble those in LW2015, with the exceptions that the area of the peak amplitude in ERSST is broader than in HadISST1 (because of the greater smoothing and normalization about the peak value), and the range of the multi-decadal variation in ERSST PC-R1 is smaller (consistent with the spatial smoothing).

A notable difference between these EOF-R1s and those from the 1870-2015 analyses without de-trending (Fig. 10a,b), especially for HadISST1 (and also present in the LW2015 results), is the peak amplitude east of the Grand Bank. It extends northward and longitudinally and has a marked decline southwest of the Grand Bank (in the de-trended 1900-2015 analysis), whereas it has reduced northward extent but extends to the MAB region in the 1870-2015 analyses. The seasonally-varying differing structure of EOF-R1 southwest of the Grand Bank in the de-trended analyses is another manifestation of a significant difference in the ERSST and HadISST1 datasets in the ML-WNA, affecting the long-term trends and low-frequency

variability. The implication is that studies of trends and AMO (or other variability) influences in this region need to pay particular attention to which historical dataset they are using.

The PC-R1s (Fig. 11c,d) have the common features of positive values from the early 1930s to the mid 1960s and from the early 1990s to the early 2010s, and negative values in the 1970s and 1980s, like the AMO index (Fig. 2b). Analyses indicate that the PC-R1s are highly correlated across seasons (with highest correlations between adjacent seasons) and between the datasets (see Tables D1 and D2) and, for all seasons, with the AMO index (with higher correlations in spring, summer and fall than in winter; Tables 5 and 6). This provides compelling evidence that this response mode is dominated by the AMO, such that the EOF-R1 patterns in Fig. 11a,b indicate the AMO's spatial patterns and their uncertainties, and the PC-R1s indicate that the recent positive AMO phase may have ended (or be ending). There is a modest positive peak in 2012 which was the recent warmest year in the ML-WNA (e.g., Mills et al., 2013; Hebert et al., 2016), which indicates that this mode contributed to these anomalous conditions.

Table 5. Correlation coefficients between the annual SST PC-Rs in Fig. 11 for both ERSST (ER) and HadISST1 (HD), and potential forcing indices after de-trending. The correlations with the AMO and winter NAO indices are for the period 1900-2015, those with Bermuda SSH for 1933-2015, and those with the SSH and WSC indices from Wang et al. (2015) for 1958-2004. Trends significant at the 95% confidence level are in **boldface**, and those significant at the 90% level are in *italics*.

SST PC-Rs	AMO	Winter NAO	Be SSH	Hindcast SSH Indices				
				Be-Br	PC1	PC2	PC3	WSC
PC-R1 ER	**0.81**	**-0.23**	0.17	-0.03	**0.35**	**0.69**	**0.43**	**0.41**
PC-R1 HD	**0.85**	-0.15	0.17	-0.03	*0.29*	**0.62**	**0.43**	**0.31**
PC-R2 ER	-0.12	**-0.27**	-0.20	**-0.54**	**0.68**	*-0.29*	-0.07	*0.29*
PC-R2 HD	0.08	**-0.40**	*0.24*	0.03	**0.54**	**0.35**	-0.01	**0.71**
PC-R3 ER	0.06)	**0.26**	**-0.29**	*-0.26*	*-0.28*	**-0.38**	0.11	**-0.64**
PC-R3 HD	**-0.20**	-0.12	**-0.37**	**-0.52**	**0.41**	**-0.47**	-0.08	-0.12

The PC-R1s for all seasons are also moderately correlated with the PC2 of SSH in the Wang et al. (2015) study for 1958-2004 (Table 6), which in turn was moderately correlated with the PC1 of the transport streamfunction (ψ) in that study. Wang et al. interpreted the coupling of these modes as mainly representing variability in the strength of the subtropical gyre and in the position of the gyre-gyre interaction region, all of which suggests associated circulation variability. We also note weak-to-moderate correlations of the PC-R1 in some seasons with Wang et al.'s SSH PC3 and with PC1 of their wind stress curl (WSC) over the NA (Table 6).

The NAO-like mode R2 is qualitatively similar to that in LW2015 and in the above 1870-2015 analyses (without de-trending), especially in HadISST1 where it is the second-leading

mode variance-wise (Fig. 11a,b). However, it differs here in being the third-leading mode variance-wise in ERSST, behind the circulation variability mode. Its spatial pattern again has a dipole structure in the NWA with opposite phases north and west of the Grand Bank, consistent with the NAO response identified by Petrie and Drinkwater (1993) and others. Its temporal variability (PC-R2; Fig. 11c,d) includes a period of positive values in the 1960s implying a contribution to the well-documented below-average ocean temperature west of the Grand Bank (e.g., Petrie 2007), although this period is not as anomalous in this mode as in the annual means. The annual (PC-R2) correlations with the winter NAO (Table 5) are lower than in LW2015 and, seasonally, the correlations are only weak-to-moderate in winter and spring (Table 6).

Table 6. Correlation coefficients between the seasonal SST PC-Rs in Fig. 11 for both ERSST (ER) and HadISST1 (HD), and potential forcing indices after de-trending. The correlations with the different indices are for the same periods as those in Table 5. Trends significant at the 95% confidence level are in **boldface**, and at the 90% level are in *italics*. [Single wide table in AO]

Season	PC-R1 vs AMO		PC-R1 vs SSH-PC2		PC-R2 vs NAO		PC-R2 vs SSH-PC1		PC-R2 vs WSC-PC1	
	ER	HD	ER	HD	ER	HD	ER	HD	ER	HD
Winter	**0.63**	**0.66**	**0.59**	**0.42**	**-0.22**	**-0.42**	**0.57**	**0.80**	0.22	**0.58**
Spring	**0.73**	**0.79**	**0.65**	**0.41**	**-0.27**	**-0.31**	**0.56**	**0.37**	0.25	**0.56**
Summer	**0.76**	**0.72**	**0.59**	**0.53**	-0.10	-0.11	0.16	0.20	0.05	0.13
Fall	**0.73**	**0.78**	**0.68**	**0.67**	*-0.17*	*-0.17*	*0.28*	0.24	0.10	0.11

Season	PC-R3 vs Be SSH		PC-R3 vs Be-Br		PC-R3 vs SSH-PC2	
	ER	HD	ER	HD	ER	HD
Winter	-0.06	**-0.42**	*-0.07*	**-0.43**	-0.12	**-0.76**
Spring	*-0.25*	**-0.32**	**-0.36**	**-0.57**	**-0.38**	**-0.42**
Summer	**-0.33**	**-0.34**	-0.23	-0.20	**-0.33**	**-0.39**
Fall	-0.11	-0.11	-0.07	-0.01	-0.14	-0.14

A marked seasonal difference is also apparent in the correlations of the PC-R2s between datasets (low in winter and spring, high in summer and fall; Table D1), and in the inter-seasonal correlations which are generally low or non-significant for non-adjacent seasons, especially in ERSST (Table D2). The annual PC-R2s' correlations with the NAO-related SSH PC1 in Wang et al. (2015) are moderately high (Table 5) but, seasonally, the significant correlations are limited to winter and spring (like those with the NAO). For HadISST1, there is still a moderate correlation with the Wang et al. WSC PC1 in annual, winter and spring but there are no significant correlations with this PC for ERSST. Collectively, these results indicate that there is an NAO-response aspect to this mode with a circulation linkage related to the wind stress curl,

but this mode is of much less importance in summer and fall than in winter and spring. The tendency for increasingly negative PC-R2 values in 2013-2015 indicates that this mode is making a warming contribution west of the Grand Bank and a cooling one southeast of Greenland which, together with its negative winter-spring correlation with the winter NAO, may imply a linkage with the recent recurrence of strong winter convection in the Labrador and Irminger Seas (Yashayaev and Loder, 2016; de Jong and de Steur, 2015) and cold SSTs across the subpolar NA (Duchez et al., 2016).

The ocean circulation variability mode (PC-R3), with its distinguishing spatial structure of a mid-latitude maximum extending eastward from the Grand Bank (Fig. 11a), accounts for 10-15% and 16-23% of the variance in various seasons in HadISST1 and ERSST, respectively. It also has differences in seasonality and detailed spatial structure between the datasets (Table 6). The PC-R3s have high correlations between the datasets in summer and fall only (Table D1), and moderate inter-seasonal correlations primarily between adjacent seasons (and limited ones otherwise for ERSST; Table D2). There is also strong seasonality and dataset-dependence in this PC's weak-to-moderate correlations with the forcing indices (Tables 5 and 6). The correlations are notably higher for HadISST1 than ERSST, with moderately weak to moderately strong values in winter, spring and summer with Bermuda SSH, and in winter and spring with Be-Br SSH and SSH PC2 from Wang et al., (2015). Also, considering the correlation between the PC2 of SSH and the PC1 of ψ in the latter study, these results provide further support for R3 being a circulation variability mode, although there is the unresolved issue of the implications of the discrepancies between the HadISST1 and ERSST datasets for these modes in the real ocean.

6 Discussion

a *Dataset Intercomparisons*

The intercomparisons confirm overall broad-scale consistency in the ERSST and HadISST1 datasets in the NA. However, they also indicate that there are significant discrepancies (e.g., Fig. 4) that need to be considered in applications to climate change, variability and impacts.

The comparisons for the major oceanographic regions indicate close agreement between the spatially-averaged SSTs in all seasons since 1940, except for a tendency for ERSST to be slightly lower than HadISST1 in summer (Fig. 3). The origin of this discrepancy is unclear.

Prior to 1940, the discrepancies are generally larger, especially in some seasons. In the NW-WNA and in the LS area in particular (Fig. 6a), it appears that the ERSST values between 1895 and 1925 are significant underestimates in winter, and also in spring and summer to a lesser degree, as a result of sparse observational data and the ERSST interpolation methodology. In contrast, in the IB area to the east of the LS (and largely outside the major regions), ERSST is consistently higher than HadISST1 in all seasons, especially prior to the 1980s (Fig. 4). It appears that data sparsity prior to 1940, and especially in winter, is also a contributing factor to this discrepancy in the IB (Fig. SI-2a), in addition to interpolation challenges. Comparison of the

interpolated SSTs from the IB with the non-interpolated ICOADS dataset indicates that HadISST1 underestimation in this area is an issue. The SODA simulation provides support for the ERSST and HadISST1 datasets being the more reliable ones in the IB and LS areas, respectively, but it deviates significantly from the interpolated and ICOADS datasets in at least one season in the other four areas where it was examined (Figs. 6, SI-2 and SI-3).

Moving back to the major regions, the broad-scale seasonal means from the two interpolated datasets are in closest agreement in the NE-NA, away from the SST gradient areas associated with the interaction of the NA's subpolar and subtropical western boundary currents (e.g., Fig. SI-1). Nevertheless, there are persistent SST differences in some seasons and periods, including an overall tendency for ERSST to be relatively high in winter, spring and fall prior to 1940, in contrast to the tendency for it to be relatively low in summer (as noted above). The persistent broad-scale but seasonally-varying differences in the NE-NA, although small, point to a systematic bias remaining in one of the interpolated datasets, especially prior to 1940.

The ML-WNA is somewhat similar to the NE-NA in having seasonally-varying ERSST-HadISST1 differences that are generally of the same (positive) sign, but different in that the ML-WNA differences in fall and winter prior to 1940 generally have larger magnitudes (than those in the NE-NA). The ML-WNA also contrasts the NW-WNA in that its large (~1^{o}C) persistent differences in fall and winter prior to 1940 are of the opposite sign to the negative ones in the NW-WNA during winter, spring and summer between 1895 and 1925. This points to a different origin for these two different discrepancies.

The ERSST-HadISST1 difference maps indicate that the positive values in the ML-WNA in fall and winter are associated with positive differences in three particular areas (FC-NB, GM-SS and SarS) which surround an area (NGS) with negative differences. It is remarkable that the differences in these four areas, and those in the LS and IB further north, are present to varying degrees in all four sub-periods and all four seasons (Fig. 4). Nevertheless, the SST time series plots for the ML-WNA (Fig. 3b), GM-SS (Fig. 6b), FC-NB (Fig. SI-2b) and SarS (Fig. SI-3b), as well as the difference plots for the ML-NWA (Fig. C1), all indicate that the ERSST-HadISST1 differences in the ML-WNA and surrounding areas were greater in fall and winter prior to 1940. This suggests that there was a methodological change affecting one of the interpolated datasets at mid latitudes in the WNA around 1940. The comparisons with ICOADS in the above plots indicate that ERSST is in better agreement with non-interpolated observations from these areas in fall and winter prior to 1940, with the HadISST1 values appearing to be underestimates. On the other hand, the comparisons of the trends from the available dedicated monitoring programs in the region that extend back beyond 1940 – DFO's P5 in the eastern GoM back to 1924 and the US lightship data (SL2010) back to 1870 in the MAB and GoM – indicate that, while the HadISST1 trends are high compared to those from the monitoring datasets, the ERSST ones are low. This suggests that there are contributions from both interpolated datasets to the pre-1940

discrepancy. The coastal HH series are not included here because they are not well correlated with the offshore DFO EB series (Appendix C), probably due to coastal processes.

Additional discrepancies in areas with complex coastal geometry can arise from the two datasets' different resolution and their different representations of islands and land promontories, such as the artificial connection of the Gulf of St. Lawrence and Scotian Shelf in ERSST.

The above dataset discrepancies contribute to differences in the trends (Fig. 9) and in the spatial structure of at least EOF-R1, especially without prior de-trending (Fig. 10a,b).

b *Decadal-scale Variability in SST*

The time series plots and EOF analyses show the strong decadal-scale SST variability in the NA since the late 1800s which needs to be considered in examining climate change and its impacts. In particular, the predominance of a multi-decadal AMO-like variation in SST averaged over the major regions (Figs. 2 and 3) and in PC-R1 from the EOF analyses (Figs. 10 and 11), in combination with the longest SST records including only two AMO cycles (and only one cycle in the modern data era), is a major complication. While observational and modelling studies are unraveling the complex patterns and dynamics of coupled atmosphere-ice-ocean climate variability and change in a basin strongly influenced by the AMOC, major upper-ocean gyres, Arctic outflows and atmospheric forcing (e.g., Buckley and Marshall, 2016; Lozier et al., 2017), it is important that impacts studies and future projections take into account the relatively short observational datasets and limited understanding of some climate system interactions.

It should also be noted that the examination here of only the leading modes in the EOF analyses introduces a bias in the results towards the broader-scale and lower-frequency features of the SST variability. This was demonstrated in LW2015 where only a small fraction of the observed *in situ* interannual SST variability at the DFO monitoring sites was explained by the three leading EOF modes, both for annual and summer. For example, the strong influence of the NAO-related extension of the Labrador Current west of the Grand Bank in the 1960s (e.g., Petrie and Drinkwater, 1993; Loder et al., 2003) is under-represented in the NAO-related R2 mode in this paper. This implies that additional local SST variability related to forcings such as the NAO is probably occurring in higher EOF modes that are not examined here.

The present results clearly show the year-round persistence of the AMO, with greatest influence in summer. The latter is consistent with a direct atmospheric contribution (such as from GMST) and/or atmospheric feedback resulting in larger amplitudes in the shallow surface mixed layer in summer, but this does not preclude an AMOC origin or contribution as well (Rahmstorf et al., 2015). In contrast to the AMO, the weaker influences of the NAO (R2) and ocean gyre interactions (R3) are strongest in winter and spring, consistent with greater depth-extent of these signals via deeper mixed layers, such as potential contributions from water mass modification in deep convection zones like the Labrador Sea (e.g., Yashayaev and Loder, 2016).

c *Trends and Anthropogenic Changes*

In spite of the dataset discrepancies, the SST trends across most of the study domain have become increasingly positive during the period covered by the interpolated datasets. A prominent feature, however, is the warming hole southeast of Greenland (and extending southward to mid latitudes) where both datasets indicate overall cooling since 1870, and also since 1900, except in summer. Although the cooling rates in this area have declined in the later historical periods, with warming present almost everywhere in the NA in both datasets since 1979, it is important that this spatial feature be recognized and further understood.

Considering the suspect features in the datasets especially prior to 1940, the trends going back to 1870 and 1900 have greater uncertainty than those since 1950. However, in summer - the season with the best data coverage - there is quantitative consistency between the datasets in the statistically-significant positive trends in average SST in the three major regions since both 1870 and 1900. These (summertime) trends are largest in the ML-WNA (~0.06ºC and 0.08ºC per decade, in those respective periods), next largest in the NW-WNA (~0.03ºC and 0.07ºC per decade), and smallest in the NE-NA (~0.02ºC and 0.04ºC per decade).

The trends since 1950 are generally larger than those since 1900, with increased consistency across seasons but still largest in summer. Further, the trends since 1979 are dramatically larger than those since 1950, typically by a factor of 3-4. The increased trend magnitude and reduced extent of the warming hole since 1950 and 1979 (compared to those since 1870 and 1900) provide strong support for the occurrence of anthropogenic ocean warming in the NA and especially the NWA over the past century.

Nevertheless, the presence of strong multi-decadal variability, and spatial structure associated with both the warming hole and this multi-decadal (and other) variability, dictates special caution in the inference of anthropogenic climate changes from limited datasets. For example, it appears likely that the amplified warming observed during the recent satellite era (since 1979) includes a contribution from the multi-decadal (AMO-like) variability in at least some regions. Thus, recent trends based on satellite datasets (e.g. Pershing et al., 2015; Larouche & Galbraith 2016) can be misleading with respect to past and future climate change trends.

d *Seasonality*

Because of seasonality in data sparsity affecting the quality of the interpolated datasets prior to 1940, it is difficult to draw conclusions on the seasonality of the SST trends and variability modes going back to 1870 or 1900. However, two relevant statements that can be made for these historical periods are that there was (i) statistically-significant warming averaged over all three major regions in summer, and (ii) a broad warming hole with apparent increasingly-reduced extent in summer. An additional notable result from P5 is that the warming trend there since 1924 was largest in winter and smallest in summer (0.15ºC and 0.07ºC per decade, respectively), which is an opposite seasonality to that at EB and S27 (since ~1947; Table 3) and that indicated by the interpolated datasets for the major regions (Table 1). It is unclear whether this difference

is related to P5's location in a bay with an additional influence from seasonally-varying river run-off and/or in an area with strong tidal mixing. It is intriguing that HadISST1 in this area (Table 3) has a similar seasonality in trends (0.20ºC and 0.13ºC per decade, in winter and summer, respectively) to that at P5 but the reliability of this needs further investigation.

The trends since 1950 have strongest seasonality in the NW-WNA with those in summer about twice those in winter, and weaker seasonality of the same sense in the other two major regions. The trends at the two offshore DFO monitoring sites (EB and S27) are qualitatively consistent with this, but with the notable result of no significant trend in both SST and AT in the vicinity of EB in winter since 1947 as a result of decadal-scale variability. During the satellite era (1979-2015) with the best data coverage and amplified trends, the weakest trends in both the NW-WNA and ML-WNA occur in spring and the strongest in summer when their magnitude approximately doubles. There is little seasonal change in the NE-NA. However, because this 36-39 year period has occurred during about a half-cycle of the AMO when SSTs have been increasing, they should not be interpreted as trends in a climate change sense.

The leading variability mode after de-trending, namely the spatially-coherent AMO-like variation with spatially-varying amplitude, accounts for the most variance in summer, but with less seasonal amplification than in the long-term trends. These summer amplifications are consistent with an atmospheric surface heating influence (such as anthropogenic climate change) and shallow summer mixed layers, but this does not rule out an AMOC contribution to (or even primary origin of) the AMO-like variation. In contrast, the weaker variations (R2 and R3) apparently associated with the NAO and gyre interactions east of the Grand Bank generally account for higher fractions of the total variance in winter and spring, consistent with winter origins, but the gyre interaction variation in ERSST is an apparent exception to this (Table 6).

These results suggest that annual means should be useful proxies for SST variability in particular seasons in many, but not all, cases. A notable exception, which also illustrates the pitfall of relatively short records in the presence of decadal-scale variability, is the absence of a warming trend in winter in *in situ* air and ocean surface temperature in the Scotian Shelf region since 1950, whereas the longer AT datasets indicate a century-scale warming in the region in winter (like in other seasons).

e *Concluding Remarks*

In relation to its goal of determining the seasonality of long-term and decadal-scale SST variability in the NWA, this study has only been a partial success. Data sparsity and quality issues prior to 1940, together with spatial and temporal variability (e.g., Figs. 6, 7 and SI-1, -2 and -3) limit the extent to which seasonality can be described quantitatively over the 136-year study period with potential anthropogenic warming and multiple AMO cycles. Nevertheless the available information from this period and the statistically-significant trends for most seasons during 1950-2015 (after some spatial averaging) provide a clear indication of greatest warming in summer and least warming in winter in most areas, consistent with expectations based on the

seasonal variation in ocean mixed layer depth. Further, the EOF analyses point to the AMO making its largest contribution to SST variability during summer, and the NAO and gyre interactions making their largest contributions in winter, all of which is understandable from basic physics. Finally, the presence of positive SST trends of ~0.1°C per decade (since the mid 1900s) at the three DFO monitoring sites off Atlantic Canada (where dataset adjustments have not been applied) provides at least regional support for significant SST warming over the past ~60-80 years in the WNA, consistent with the findings of Hartmann et al. (2013).

Considering the importance of SST described in the Introduction and taking a long-term view, the identification of dataset discrepancies and limitations should also be considered as a valuable result. The two interpolated gridded datasets have significant discrepancies in some areas, especially prior to 1940, that affect their trends. This study should only be considered as a small step towards an improved description of SST and ocean climate variability in the NWA, as a partial basis for increased understanding and eventually predictive capabilities. Further comparisons, evaluations and interpretative analyses of various datasets are needed, in addition to continued and desirably expanded monitoring, and process and modelling studies.

Acknowledgements

This work was motivated by DFO's Aquatic Climate Change Adaptation Services Program (ACCASP). We are grateful to the many individuals and organizations who have contributed to the collection, quality control, analysis and dissemination of the datasets used in this study. We would especially like to thank Roger Pettipas for providing most of the AZMP datasets and assisting with the analyses of various datasets, Eugene Colbourne for providing the updated data for Station 27, and them and other colleagues at DFO's BIO, Northwest Atlantic Fisheries Centre, St. Andrews Biological Station and Integrated Data Management Services for maintaining the important AZMP time series. We also thank Dave Hebert and Ingrid Peterson for their comments on an early version of the manuscript, Brian Petrie and Igor Yashayaev for discussions. DFO Colleagues, Chantelle Layton and Lindsay Beazley , internally reviewed this report, and provided helpful comments.

References

Bakker, P., et al. (2016). Fate of the Atlantic Meridional Overturning Circulation: Strong decline under continued warming and Greenland melting. *Geophysical Research Letters, 43*, 12,252 – 12,260. doi:10.1002/2016GL070457.

Belkin, I. (2009). Rapid warming of large marine ecosystems. *Progress in Oceanography, 81,* 207-213. doi:10.1016/j.pocean.2009.04.011.

Bellucci, A., Mariotti, A., & Gualdi, S. (2017). The role of forcings in the twentieth-century North Atlantic multidecadal variability: The 1940–75 North Atlantic cooling case study. *Journal of Climate, 30,* 7317-7337, doi:10.1175/JCLI-D-16-0301.1.

Booth, B.B.B., Dunstone, N.J., Halloran, P.R., Andrews, T., & Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature, 484*, 228-232. doi:10.1038/nature10946.

Brickman, D., Hebert, D., & Wang, Z. (2018). Mechanism for the recent ocean warming events on the Scotian Shelf of eastern Canada. *Continental Shelf Research, 156,* 11-22. doi:10.1016/j.csr.2018.01.001.

Brock, R.J., Kenchington, E., & Martinez-Arroto, A. (Eds.) (2012). *Scientific guidelines for designing resilient Marine Protected Area networks in a changing climate*. *Annex 1* Commission for Environmental Cooperation: Montreal, Canada. 95 p. Retrieved from http://www3.cec.org/islandora/en/item/10820-scientific-guidelines-designing-resilient-marine-protected-area-networks-in-changing-en.pdf.

Brown, P. T., Li, W., Jiang, J.H., & Su, H. (2016). Spread in the magnitude of climate model interdecadal global temperature variability traced to disagreements over high-latitude oceans. *Geophysical Research Letters, 43*, 12,543–12,549. doi:10.1002/2016GL071442.

Buckley, M.W., & Marshall, J. (2016). Observations, inferences, and mechanisms of Atlantic Meridional Overturning Circulation variability: A review. *Reviews of Geophysics, 54 (1)*, 5-63. doi:10.1002/2015RG000493.

Carton, J. A., & Giese, B.S. (2008). A reanalysis of ocean climate using simple ocean data assimilation (SODA). *Monthly Weather Review, 136,* 2999–3017. doi:10.1175/2007MWR1978.1.

Chen, X., & Tung, K.-K. (2018). Global-mean surface temperature variability: space-time perspective from rotated EOFs. *Climate Dynamics, 51,* 1719-1732. doi:10.1007/s00382-017-3979-0.

Chen, K., Gawarkiewicz, G., Kwon, Y.-O., & Zhang, W. G. (2015). The role of atmospheric forcing versus ocean advection during the extreme warming of the Northeast U.S. continental shelf in 2012. *Journal of Geophysical Research Oceans, 120*, 4324–4339. doi:10.1002/2014JC010547.

Chen, X., Wallace, J.M., & Tung, K.-K. (2017). Pair-wise rotated EOFs of global SST. *Journal of Climate, 30*, 5473-5489. doi:10.1175/JCLI-D-16-0786.1.

Chepurin, G.A., Carton, J.A., & Leulitte, E. (2014). Sea level in ocean reanalyses and tide gauges. *Journal of Geophysical Research: Oceans, 119*, 147-155. doi:10.1002/2013JC009365.

Claret, M., Galbraith, E.D., Palter, J.B., Bianchi, D., Fennel, K., Gilbert, D., & Dunne, J.P. (2018). Rapid coastal deoxygenation due to ocean circulation shift in the northwest Atlantic. *Nature Climate Change, 8,* 868-872. doi:10.1038/s41558-018-0263-1.

Colbourne, E., Holden, J., Senciall, D., Bailey, W., Snook, S., & Higdon, J. (2016). Physical oceanographic conditions on the Newfoundland and Labrador Shelf during 2015. Fisheries and Oceans Canada, *Canadian Science Advisory Secretariat Research Document, 2016/079*. v + 40 p. Retrieved from http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2016/2016_079-eng.pdf

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., …, Wehner, M. (2013). Long-term climate change: projections, commitments and irreversibility. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* pp 1029-1136. Cambridge, UK: Cambridge University Press.

Danabasoglu , G., Yeager, S.G., Kwon, Y.-O., Tribbia, J.J., Phillips, A.S., & Hurrell, J.W. (2012). Variability of the Atlantic Overturning Circulation in CCSM4, *Journal of Climate, 25,* 5153-5172. doi:10.1175/JCLI-D-11-00463.1.

de Jong, M.F., & de Steur, L. (2016). Strong winter cooling over the Irminger Sea in winter 2014-2015, exceptional deep convection, and the emergence of anomalously low SST. *Geophysical Research Letters, 43,* 7106-7113, doi:10.1002/2016GL069596.

Delworth, T.L., Zeng, F., Zhang, L., Zhang, R., Vecchi, G.A., & Yang, X. (2017). The central role of ocean dynamics in connecting the North Atlantic Oscillation to the extratropical component of the Atlantic Multidecadal Oscillation. *Journal of Climate, 30*, 3789-3805, doi:10.1175/JCLI-D-16-0358.1.

Deser, C., Alexander, M.A., Xie, S.-P., & Phillips, A.S. (2010). Sea surface temperature variability: patterns and mechanisms. *Annual Reviews of Marine Science, 2,* 115-143. doi:10.1146/annurev-marine-120408-151453.

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? *Journal of Climate*, *25,* 8373-8379. doi:10.1175/JCLI-D-12-00490.1.

Drinkwater, K.F., Beaugrand, G., Kaeriyama, M., Kim, S., Ottersen, G., Perry, R.I., Pörtner, H.-O., Polovina, J., & Takasuka, A. (2010). On the processes linking climate to ecosystem changes. *Journal of Marine Systems, 79,* 374-388. doi:10.1016/j.jmarsys.2008.12.014.

Drinkwater, K.F., Miles, M., Medhaug, I., Otterå, O.H., Kristiansen, T. Sundby, S., & Gao, Y. (2014). The Atlantic Multidecadal Oscillation: its manifestations and impacts with special emphasis on the Atlantic region north of 60°N. *Journal of Marine Systems, 133*, 117-130. doi:10.1016/j.jmarsys.2013.11.00.

Duchez, A., et al. (2016). Drivers of exceptionally cold North Atlantic Ocean temperature and their link to the 2015 European heat wave. *Environmental Research Letters, 11*, doi:10.1088/1748-9326/11/7/074004.

Enfield, D.B., Mestas-Nunez, A.M., and Trimble, P.J. (2001). The Atlantic Multi-decadal Oscillation and its relationship to rainfall and river flows in the continental US. *Geophysical Research Letters, 28*, 2077–2080. doi:10.1029/20000GL012745.

Folland, C.K., Boucher, O., Colman, & Parker, D.E. (2018). Causes of irregularities in trends of global mean surface temperature since the late 19th century. *Science Advances, 4 (6)*, 16pp, doi:10.1126/sciadv.aao5297.

Foukal, N.P., & Lozier, S. (2016). No inter-gyre pathway for sea-surface temperature anomalies in the North Atlantic. *Nature Communications, 7*, 11333. doi:10.1038/ncomms11333.

Frankignoul, C., Gastineau, G. & Kwon, Y.O. (2017). Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic Multidecadal Oscillation and the Pacific Decadal Oscillation. *Journal of Climate, 30,* 9871-9895. doi:10.1175/JCLI-D-17-0009.1.

Freeland, E., et al. (2017). ICOADS Release 3.0: a major update to the historical marine climate record. *International Journal of Climatology, 37*, 2211–2232. doi:10.1002/joc.4775.

Friedland, K.D., & Hare, J.A. (2007). Long-term trends and regime shifts in sea surface temperature on the continental shelf of the northeast United States. *Continental Shelf Research, 27*, 2313–2328. doi:10.1016/j.csr.2007.06.001.

Galbraith, P.S., & Larouche, P. (2013). Trends and variability in eastern Canada sea-surface temperatures. Ch.1 In: Aspects of climate change in the Northwest Atlantic off Canada [Loder, J.W., G. Han, P.S. Galbraith, J. Chassé and A. van der Baaren (Eds.)]. Can. Tech. Rep. Fish. Aquat. Sci. 3045: x + 192 p. Retrieved from http://www.dfo-mpo.gc.ca/Library/350208.pdf

Galbraith, P.S., Larouche, P., Chassé, J., & Petrie, B. (2012). Sea surface temperature in relation to air temperature in the Gulf of St. Lawrence: interdecadal variability and long-term trends. *Deep-Sea Research II, 77-80,* 10-20. doi:10.1016/j.dsr2.2012.04.001.

Giese, B. S., & Ray, S. (2011). El Niño variability in simple ocean data assimilation (SODA), 1871–2008. *Journal of Geophysical Research, 116,* C02024. doi:10.1029/2010JC006695.

GISTEMP Team (2016). *GISS Surface Temperature Analysis (GISTEMP)*. NASA Goddard Institute for Space Studies. Dataset accessed 2016-11-04 at https://data.giss.nasa.gov/gistemp/

Hakkinen, S., Rhines, P.B., & Worthen, D.L. (2011). Atmospheric blocking and Atlantic Multidecadal ocean variability. *Science, 334,* 655-659. doi:10.1126/science.1205683.

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews in Geophysics*, *48 (4)*, RG4004. doi:10.1029/2010RG000345.

Hartmann, D.L., et al. (2013). Observations: atmosphere and surface. In T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.) *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* pp. 159-254. Cambridge, UK: Cambridge University Press.

Hausfather, Z., Cowtan, K., Clarke, D.C., Jacobs, P., Richardson, M., & Rohde, R. (2017). Assessing recent warming using instrumentally homogeneous sea surface temperature records. *Science Advances*, *3 (1)*, 13pp. doi:10.1126/sciadv.1601207 (published on-line 4 January 2017).

Hebert, D., Pettipas, R., Brickman, D., & Dever, M. (2016). Meteorological, sea ice and physical oceanographic conditions on the Scotian Shelf and in the Gulf of Maine during 2015. Fisheries and Oceans Canada, *Canadian Science Advisory Secretariat Research Document, 2016/083*. v + 49 p. Retrieved from http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2016/2016_083-eng.pdf

Holgate, S.J., et al. (2013). New data systems and products at the Permanent Service for Mean Sea Level. *Journal of Coastal Research, 29* (3), 493-504. doi:10.2112/JCOASTRES-D-12-00175.1.

Huang, B., et al. (2015). Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4): Part I. Upgrades and intercomparisons. *Journal of Climate, 28*, 911-930, doi:10.1175/JCLI-D-14-00006.1.

Huang, B., Thorne, P., Smith, T., Liu, W., Lawrimore, J., Banzon, V., Zhang, Peterson, H., & Menne, M. (2016). Further exploring and quantifying uncertainties for Extended Reconstructed Sea Surface Temperature (ERSST) version 4 (v4). *Journal of Climate, 29*, 3119-3142. doi:10.1175/JCLI-D-15-0430.1.

Hurrell, J. W., & Deser, C. (2010). North Atlantic climate variability: the role of the North Atlantic Oscillation. *Journal of Marine Systems, 79,* 231-244. doi:10.1016/j.jmarsys.2009.11.002.

IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri & L.A. Meyer (eds.), IPCC, Geneva, Switzerland, 151pp.

Ishii, M., Shouji, A., Sugimoto, S., & Matsumoto, T. (2005). Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using COADS and the KOBE collection. *International Journal of Climatology, 25,* 865-869. doi:10.1002/joc.1169.

Josey, S.A., Hirschi, J.J.-M., Sinha, B., Duchez, A., Grist, J.P., & Marsh, R. (2018). The recent Atlantic cold anomaly: causes, consequences, and related phenomena. *Annual Review of Marine Science 2018, 10,* 475–501. doi:10.1146/annurev-marine-121916-063102.

Joyce, T.M., & Zhang, R. (2010). On the path of the Gulf Stream and the Atlantic Meridional Overturning Circulation. *Journal of Climate, 23*, 3146-3154. doi:10.1175/2010JCLI3310.1.

Karl, T.R., et al. (2015). Possible artifacts of data biases in the recent global surface warming hiatus. *Science 348* (6242), 1469-1472. doi:10.1126/science.aaa5632.

Kavanaugh, M. T., Rheuban, J. E., Luis, K. M. A., & Doney, S. C. (2017). Thirty-three years of ocean benthic warming along the U.S. Northeast Continental Shelf and Slope: Patterns, drivers, and ecological consequences. *Journal of Geophysical Research: Oceans, 122*, 9399–9414. https://doi.org/10.1002/2017JC012953.

Kennedy, J. J. (2014). A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Reviews of Geophysics, 52*, 1-32. doi:10.1002/2013RG000434.

Kennedy J.J., Rayner, N.A., Smith, R.O., Parker, D.E., & Saunby, M. (2011a). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research, 116*, D14103. doi:10.1029/2010JD015218

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, & M. Saunby (2011b) Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *Journal of Geophysical Research, 116*, D14104, doi:10.1029/2010JD015220.

Knight, J.R., Allan, R.J., Folland, C., Vellinga, M., & Mann, M.E. (2005). A signature of persistent natural thermohaline cycles in observed climate. *Geophysical Research Letters, 32*, L20708. doi:10.1029/2005GL024233.

Knutti, R., & Sedláček, J. (2013). Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change, 3*, 369-373. doi:10.1038/nclimate1716.

Larouche, P., & Galbraith, P.S. (2016). Canadian coastal seas and Great Lakes sea surface temperature climatology and recent trends. *Canadian Journal of Remote Sensing, 42 (3),* 243-258. doi:10.1080/07038992.2016.1166041.

Li, F., Orsolini. Y.J., Wang, H., Gao, Y., & He, S. (2018). Atlantic Multidecadal Oscillation modulates the impacts of Arctic sea ice decline. *Geophysical Research Letters, 45*, 2497-2506, doi:10.1002/2017GL076210.

Lian, T., Shen, Z., Ying, J., Tang, Y., Li, J., & Ling, Z. (2018). Investigating the uncertainty in global SST trends due to internal variations using an improved trend estimator. *Journal of Geophysical Research Oceans, 123,* 1877-1895. doi:10.1002/2017JC013410.

Liu, W., Huang, B., Thorne, P.W., Banzon, V.F., Zhang, H.-M. Freeman, E., Lawrimore, J., …, Woodruff, S.D. (2015). Extended Reconstructed Sea Surface Temperature version 4 (ERSST.v4): Part II. Parametric and structural uncertainty estimations. *Journal of Climate, 28*, 931-951, doi:10.1175/JCLI-D-14-00007.1.

Loder, J.W., & Greenberg, D.A. (1986). Predicted positions of tidal fronts in the Gulf of Maine. *Continental Shelf Research, 6 (3)*, 397-414.

Loder, J.W., & Wang, Z. (2015). Trends and variability of sea surface temperature in the Northwest Atlantic from three historical gridded datasets. *Atmosphere-Ocean, 53 (5),* 510-528. doi:10.1080/07055900.2015.1071237.

Loder, J.W., Hannah, C.G., Petrie, B.D., & Gonzalez, E.A. (2003). Hydrographic and transport variability on the Halifax section. *Journal of Geophysical Research, 108 (C11)*, 8003, 1-18. doi:10.1029/2001/JC001267.

Loder, J.W., Petrie, B., & Gawarkiewicz, G. (1998). The coastal ocean off northeastern North America: a large-scale view. In A.R. Robinson and K.H. Brink (Eds.) *The Global Coastal Ocean: Regional Studies and Synthesis. The Sea, Volume 11,* pp 105-133. New York, USA: John Wiley & Sons, Inc.

Loder, J.W, van der Baaren, A., & Yashayaev, I. (2015). Climate comparisons and change projections for the Northwest Atlantic from six CMIP5 models. *Atmosphere-Ocean, 53 (5)*, 529-555. doi:10.1080/07055900.2015.1087836.

Loder, J.W., Wang, Z., van der Baaren, A., & Pettipas, R. (2013). Trends and variability of sea surface temperature (SST) in the North Atlantic from the HadISST, ERSST and COBE datasets. *Canadian Technical Report of Hydrography and Ocean Sciences, 292,* viii + 36 p. Retrieved from http://www.dfo-mpo.gc.ca/Library/350066.pdf

Lozier, M., et al. (2017). Overturning in the Subpolar North Atlantic Program: a new international ocean observing system. *Bulletin of the American Meteorological Society, 98 (4)*, 737-752. doi:10.1175/BAMS-D-16-0057.1.

Mann, K.H., & Lazier, J.R.N. (2006). *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans.* Oxford, UK: Blackwell Publishing Limited, 496 pp. doi:10.1002/9781118687901.

Mann, M.E., Steinman, B.A., & Miller, S.K. (2014). On forced temperature changes, internal variability, and the AMO. *Geophysical Research Letters, 41*, 3211–3219. doi:10.1002/2014GL059233.

McCarthy, G.D., Joyce, T.M, & Josey, S.A. (2018). Gulf Stream variability in the context of quasi-decadal and multi-decadal Atlantic climate variability. *Geophysical Research Letters, 45,* 11,257 – 11,264. doi: 10.1029/2018GL079336.

Mills, K.E., Pershing, A.J., Brown, C.J., Chen, Y., Chiang, F.-S., Holland, D.S., Lehuta, S., …, Wahle R.A. (2013). Fisheries management in a changing climate: Lessons from the 2012 ocean heat wave in the Northwest Atlantic. *Oceanography, 26 (2),* 191–195. doi:10.5670/oceanog.2013.27.

Pershing, A.J., Alexander, M.A., Hernandez, C.M., Kerr, L.A., Le Bris, A., Mills, ,K.E., Nye, J.A., …, Thomas, A.C. (2015). Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. *Science 350 (6262)*, 809-812. doi:10.1126/science.aac9819.

Petrie, B. (2004). The Halifax Section: A brief history. *Atlantic Zone Monitoring Program Bulletin, 4*, 26–29. Retrieved from http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/azmp-pmza/publications-eng.html

Petrie, B. (2007). Does the North Atlantic Oscillation affect hydrographic properties on the Canadian Atlantic continental shelf? *Atmosphere-Ocean, 45*, 141–151. doi:10.3137/ao.450302.

Petrie, B., & Drinkwater, K.D. (1993). Temperature and salinity variability on the Scotian Shelf and Gulf of Maine 1945-1990. *Journal of Geophysical Research,* 98, 20079-20089.

Petrie, B., Topliss, B.J., & Wright, D.G. (1987). Coastal upwelling and eddy development off Nova Scotia. *Journal of Geophysical Research, 29 (C12)*, 12,979-12,991.

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,* 475-480. doi: 10.1038/nclimate2554.

Rayner, N.A., Brohan, P. Parker, D.E., Folland, C.K., Kennedy, J.J., Vanicek, M., Ansell, T.J., & Tett, S.F.B. (2006). Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset. *Journal of Climate, 19*, 446- 449. doi:10.1175/JCLI3637.1.

Rayner, N.A., Kaplan, A., Kent, E.C., Reynolds, R.W., Brohan, P., Casey, K.S., …,,Brandon, T. (2010). Evaluating climate variability and change from modern and historical SST observations. In J. Hall, D.E. Harrison, & D. Stammer (Eds.), *Proceedings of the OceanObs'09 Conference, Vol. 2.2 – Community White Papers (Part 2)* (pp. 819-829). European Space Agency WPP-306. Retrieved from http://www.oceanobs09.net/proceedings/

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

Rhein, M., Kieke, D., & Steinfeldt, R. (2015). Advection of North Atlantic Deep Water from the Labrador Sea to the southern hemisphere. *Journal of Geophysical Research Oceans, 120*, 2471–2487. doi:10.1002/2014JC010605.

Ruprich-Robert, Y., Msadek, R., Castruccio, F., Yeager, S., Delworth, T., & Danabasoglu, G. (2017) Assessing the climate impacts of the observed Atlantic Multidecadal Variability using the GFDL CM2.1 and NCAR CESM1 global coupled models. *Journal of Climate, 30,* 2785-2810, doi:10.1175/JCLI-D-16-0127.1.

Saba, V. S., Griffies, S.M., Anderson, W.G., Winton, M., Alexander, M.A., Delworth, T.L., Hare, J.A., …, Zhang, R. (2016). Enhanced warming of the Northwest Atlantic Ocean under climate change. *Journal of Geophysical Research Oceans, 121*, 118-132. doi:10.1002/2015JC011346.

Shan, S., Sheng, J., Thompson, K.R., & Greenberg, D.A. (2011). Simulating the three-dimensional circulation and hydrography of Halifax Harbour using a multi-nested coastal ocean circulation model. *Ocean Dynamics, 61* (7), 951-976. doi:10.1007/s10236-011-0398-3.

Shearman, R.K., & Lentz, S.J. (2010). Long-term sea surface temperature variability along the U.S. east coast. *Journal of Physical Oceanography, 40*, 1004–1017. doi:10.1175/2009JPO4300.1.

Sheridan, S.C., & Lee, C.C. (2018). Temporal trends in absolute and relative extreme temperatures across North America. *Journal of Geophysical Research Atmospheres, 123 (2)*, 11,889-11,898. doi:10.1029/2018JD029150.

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., et al. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophysical Research Letters, 45*, 1527–1533. doi:10.1002/2017GL076350.

Smith, T.M., & Reynolds, R.W. (2003). Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997). *Journal of Climate, 16,* 1495–1510.

Smith, T.M., & Reynolds, R.W. (2004). Improved extended reconstruction of SST (1854–1997). *Journal of Climate, 17,* 2466-2477.

Smith, T.M., Reynolds, R.W., Peterson, T.C., & Lawrimore, L. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006*). Journal of Climate, 21,* 2283-2296. doi:10.1175/2007JCLI2100.1.

Soon, W., Connolly, R., & Connolly, M. (2015). Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century. *Earth-Science Reviews, 150*, 409-452, doi:10.1016/j.earscirev.2015.08.010.

Steffen, W., Leinfelder, R., Zalasiewicz, J., Waters, C.N., Williams, M., Summerhayes, C., Barnosky, A.D., …, Schellnhuber, H.J. (2016). Stratigraphic and Earth system approaches to defining the Anthropocene. *Earth's Future, 4,* 324–345. doi:10.1002/2016EF000379.

Stocker , T.S., Qin, D., Plattner, D.-K., Allen, K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley P.M. (Eds.) (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press, 1535 pp, doi:10.1017/CBO9781107415324.

Terray, L. (2012). Evidence for multiple drivers of North Atlantic multi-decadal climate variability. *Geophysical Research Letters 39 (19)*, L19712. doi:10.1029/2012GL053046.

Thibodeau, B., Not, C., Zhu, J., Schmittner, A., Noone, D., Tabor, C., Zhang, J., & Liu, Z. (2018). Last century warming over the Canadian Atlantic shelves linked to weak Atlantic Meridional Overturning Circulation. *Geophysical Research Letters, 45*, 12,376-12,385. doi:10.1029/2018GL080083.

Thomas, A.C, Pershing, A.J., Friedland, K.D., Nye, J.A., Mills, K.A., Alexander, M.A., Record, N.R., Weatherbee, R., & Henderson, M.E. (2017). Seasonal trends and phenology shifts in sea surface temperature on the North American northeastern continental shelf. *Elementa - Science of the Anthropocene, 5,* 48, doi:10.1525/elementa.240.

Thompson, K.R., Loucks, R.H., & Trites, R.W. (1988). Sea surface temperature variability in the shelf-slope region of the northwest Atlantic. *Atmosphere-Ocean, 26 (2)*, 282-299.

Ting, M., Kushnir, Y., & Li, C. (2014). North Atlantic multi-decadal SST oscillation: External forcing versus internal variability. *Journal of Marine Systems, 133,* 27-38. doi:10.1016/j.jmarsys.2013.07.006.

Vannitsem, S., & Ghil, M. (2017). Evidence of coupling in ocean-atmosphere dynamics over the North Atlantic. *Geophysical Research Letters, 44*, 2016-2026. doi: 10.1002/2016GL072229.

Vincent, L.A., Wang, X.L., Milewska, E.J., Wan, H., Yang, F., & Swail, V. (2012). A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis. *Journal of Geophysical Research, 117*, D18110. doi:10.1029/2012JD017859.

Wang, Z., Lu, Y., Dupont, F., Loder, J.W., Hannah, C., & Wright, D.G. (2015). Variability of sea surface height and circulation in the North Atlantic: forcing mechanisms and linkages. *Progress in Oceanography, 132*, 273-286. doi:10.1016/j.pocean.2013.11.004.

Watelet, S., Beckers, J.-M., & Barth, A. (2017). Reconstruction of the Gulf Stream from 1940 to the present and correlation with the North Atlantic Oscillation. *Journal of Physical Oceanography, 47*, 2741-2754. doi:10.1175/JPO-D-17-0064.1.

Woodruff, S.D., et al. (2011). ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *International Journal of Climatolology, 31,* 951-967. doi:10.1002/joc.2103

Wu, L., et al. (2012). Enhanced warming over the global subtropical western boundary currents. *Nature Climate Change*, *2*, 161-166. doi:10.1038/nclimate1353.

Yang, H., Lohmann, G., Wei, W., Dima, M., Ionita, M., & Liu, J. (2016). Intensification and poleward shift of subtropical western boundary currents in a warming climate. *Journal of Geophysical Research Oceans, 121,* 4928–4945. doi:10.1002/2015JC011513.

Yashayaev, I., & Loder, J.W. (2016). Recurrent replenishment of Labrador Sea Water and associated decadal-scale variability. *Journal of Geophysical Research Oceans, 121*, 8095-8114. doi:10.1002/2016JC012046.

Yasunaka, A.S., & Hanawa, K. (2011). Intercomparison of historical sea surface temperature datasets. *International Journal of Climatology, 31*, 1056-1073. doi:10.1002/joc.2104.

Zhang, R. (2017). On the persistence and coherence of subpolar sea surface temperature and salinity anomalies associated with the Atlantic multidecadal variability. *Geophysical Research Letters, 44*, 7865–7875. doi:10.1002/2017GL074342.

Zhang, R., et al.(2013). Have aerosols caused the observed Atlantic Multidecadal Variability? *Journal of Atmospheric Sciences, 70*, 1135-1144. doi:10.1175/JAS-D-12-0331.1.

Appendix A: Glossary of Abbreviations and Acronyms

AHCCD: Adjusted and Homogenized Canadian Climate Data of ECCC

AMO (or AMV): Atlantic Multi-decadal Oscillation (or Variability)

AMOC: Atlantic Meridional Overturning Circulation

AT: Air temperature

AZMP: Atlantic Zone Monitoring Program of DFO

Be-Br: Bermuda minus Bravo

BIO: Bedford Institute of Oceanography

COBE: Centennial in situ Observation Based Estimates dataset

DFO: Fisheries and Oceans Canada

EB: Emerald Basin on the Scotian Shelf

ECCC: Environment and Climate Change Canada

EOF: Empirical orthogonal function

ERSST (or ER): Extended Reconstruction of SST dataset

FC-NB: Flemish Cap and Newfoundland Basin

GM-SS: Gulf of Maine and Scotian Slope

GMST: Global mean (annual, land-sea) surface temperature

GoM: Gulf of Maine

GSN: Gulf Stream North

GTS: Global Telecommunications System

HadISST1 (or HD): Hadley Centre Interpolated Sea Ice and Sea Surface Temperature dataset

HadSST2, HadSST3: versions 2 and 3 of the Hadley Centre SST dataset (non-interpolated)

HH: Halifax Harbour

ICOADS: International Comprehensive Ocean-Atmosphere Data Set

IB: Irminger Basin

LS: Labrador Sea

LW2015: Loder and Wang (2015)

MAB: Middle Atlantic Bight

MAB-NE and MAB-SW: Northeast (NE) and southeast (SE) parts of the MAB

ML-WNA: Mid-latitude Western North Atlantic

NA: North Atlantic

NAO: North Atlantic Oscillation

NE-NA: Northeast North Atlantic

NWA: Northwest Atlantic

NW-WNA: Northwest Western North Atlantic

P5: Prince 5 off St. Andrews, New Brunswick

PC: Principal component time series for an EOF mode

Ψ: Depth-integrated transport streamfunction

R1, R2, R3: Response modes (e.g., EOF-R1, PC-R1) in an EOF analysis

S27: Station 27 off St. John's, Newfoundland

SarS: Sargasso Sea

SI: Supporting Information

SODA: Simple Ocean Data Assimilation (SODA)

SSH: Sea surface height

SST: Sea surface temperature

SL2010: Shearman and Lentz (2010)

US NASA: US National Aeronautics and Space Agency

US NOAA: US National Oceanic and Atmospheric Administration

WNA: Western North Atlantic

WSC: Wind stress curl

Appendix B: Intercomparison of Air and Ocean Temperatures in the Scotian Shelf Region

A long time series (back to 1926) of monthly coastal SST for HH is available from DFO's AZMP (http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/azmp-pmza/climat/sst-tsm/costal-cotieres-eng.asp?id=Halifax; also see Hebert et al., 2016). The series originated with twice-daily near-surface bucket sampling (on week workdays) at DFO's former Halifax Fisheries Laboratory (e.g., Hachey, 1939) which continued until August 1998 (Brian Petrie, BIO, personal communication, 2017). In July 1993, DFO's BIO started the ongoing deployment of moored temperature recorders at 2m below the surface off the nearby Maritime Museum of the Atlantic. The five years of concurrent data allow the computation of a multi-year average monthly offset (bucket minus recorder) between the series. The offsets range from -0.2°C in December to 1.9°C in July, apparently reflecting the seasonal near-surface stratification cycle. These offsets have been applied to the recorder data since September 1998 to obtain an "adjusted" HH monthly SST time series consisting of the bucket observations prior to that month and the adjusted recorder monthly means since then (provided by Roger Pettipas of BIO, personal communication, 2017). Seasonal- and annual-mean time series have been computed for the same seasons in the same way as for the other DFO sites, with seasonal means computed for both the cases of data from at least two months in a season, and data from any month.

These adjusted HH series (for the case of data from any month) are displayed in Fig. B1 together with the corresponding EB series (same as in Fig. 7 except that data from at least two months were required there). Also shown are time series of annual and seasonal air temperature (AT) for Halifax (1872-2015) and Sable Island (1898-2015; on the outer Scotian Shelf), computed from ECCC's AHCCD. The trend values for each time series are shown in Table B1, for their entire records and for the periods of the relevant SST series.

Differences and similarities between the EB and HH time series are apparent (Fig. B1), in addition to the HH temperatures being notably colder than their EB counterparts in summer and fall as expected with wind-driven coastal upwelling (e.g., Petrie et al., 1987). Correlations between the EB and HH series in particular seasons are only weak-to-moderate, with coefficients ranging from 0.33 in fall to 0.60 in spring. Seasonal trends during the 1946-2015 overlap period (Table B1) have some qualitative similarities (such as non-significant negative trends in winter and significant positive trends of 0.13-0.21°C per decade in summer) as well as quantitative differences. It appears that an additional process is contributing to the variability at HH (such as variable upwelling and/or embayment-shelf exchange) compared to the broader-scale offshore variability affecting the EB and HadISST1 time series (note the coastal resolution issue that affects the ERSST time series at EB, as discussed in Section 3c).

The Halifax and Sable Island AT series (Fig. B1) show the expected continental-versus-maritime seasonal differences in mean values. In contrast to the SSTs, they are moderately-to-

highly correlated during their 1898-2015 overlap period with coefficients ranging from 0.71 in summer to 0.86 in winter. Their trends (Table B1) show overall similarity but also notable differences. They are positive with magnitudes ~0.1ºC per decade in all seasons during the overlap period (1898-2015), with little seasonal variation at Sable Island but a winter-to-summer increase by a factor of 2 at Halifax. The AT trends' seasonal variation is different during the 1947-2015 period of the EB SST observations, with the spring and summer trend magnitudes at both sites increased to ~0.2ºC per decade and no significant trend in winter at both sites and in fall at Halifax. This pattern is similar to that in both the EB and HH SST, with the consistency of the Sable Island AT and EB SST trends in particular providing support for the reliability of the strong seasonality in the EB SST trends (since 1947).

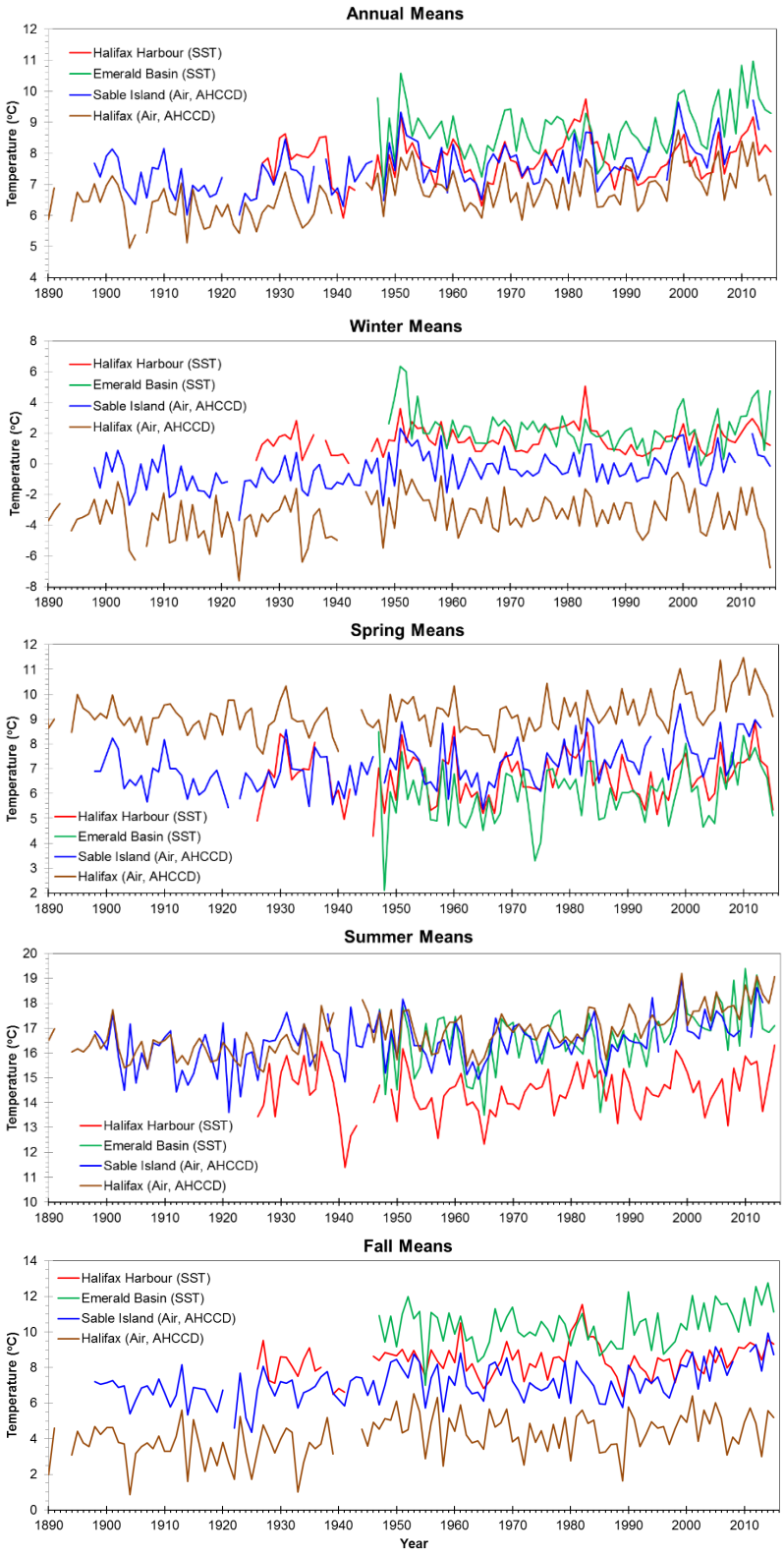

Figure B1. Annual and seasonal SST for Halifax Harbour and Emerald Basin from DFO AZMP, and air temperature for Sable Island and Halifax from ECCC's AHCCD. Seasonal SST means are retained for cases with data from any month.

Table B1. Trends (in ºC per decade) in annual and seasonal AT and SST time series from the Scotian Shelf region, in the vicinity of the Halifax Section (e.g., Loder et al., 2003; Petrie, 2004). The AT series for Sable Island and Halifax from ECCC AHCCD, and the SST series for EB and HH from DFO AZMP, are shown in Fig. B1. The ERSST (ER) and HadISST1 (HD) SST series for EB are shown in Fig. 7. Trends are provided for the different periods of the AHCCD and AZMP series. Those significant at the 95% confidence level are in **boldface**, and those significant at the 90% level are in *italics*.

Site	Period	Annual	Winter	Spring	Summer	Fall
Sable Island AT	1898-2015	**0.11**	**0.12**	**0.12**	**0.10**	**0.13**
"	1947-2015	**0.12**	0.03	**0.19**	**0.16**	**0.14**
Halifax AT	1872-2015	**0.11**	**0.09**	**0.07**	**0.13**	**0.13**
"	1898-2015	**0.11**	**0.07**	**0.09**	**0.16**	**0.11**
"	1926-2015	**0.11**	0.02	**0.15**	**0.20**	*0.08*
"	1947-2015	**0.08**	-0.09	**0.19**	**0.25**	-0.01
DFO EB SST	1947-2015	**0.12**	-0.09	*0.14*	**0.24**	**0.15**
ER EB SST	1947-2015	**0.15**	**0.13**	**0.23**	**0.21**	**0.13**
HD EB SST	1947-2015	**0.23**	**0.19**	**0.25**	**0.37**	**0.23**
DFO HH SST	1947-2015	0.05	-0.02	0.04	**0.13**	0.04
"	1926-2015	*0.05*	0.04	0.03	0.06	**0.08**
ER EB SST	1926-2015	**0.10**	**0.07**	**0.13**	**0.14**	**0.06**
HD EB SST	1926-2015	**0.18**	**0.18**	**0.17**	**0.18**	**0.19**

On the other hand, it is important to note that the lack of a winter trend in EB SST (and in Sable Island AT) is influenced by a period with record or near-record warm winters in the early 1950s. The winter EB SST series based on data from at least two months (Fig. 7) have only two relatively warm years (1951 and 1954), however, the winter means for 1950 and 1952 based on data from only one month (Fig. B1) also have comparably high values in this period. The winter ATs at Sable Island (Fig. B1) were at near-record values during the early 1950s (comparable to or higher than those since 2000), supporting the indication of no significant (or a negative) trend in winter air and surface ocean temperatures in the vicinity of Sable Island and Emerald Basin since the 1940s. Similarly, relatively warm winter AT values at Halifax in the early 1950s probably contributed to the relatively warm HH SST values during these years, and hence to the absence of statistically-significant winter trends in these records during 1947-2015.

The annual and seasonal gridded SST time series for the ML-WNA region (Figs. 2 and 3) also indicate relatively warm conditions during the early 1950s with largest anomalies in winter, indicating that this is a widespread feature. The winter trends in ERSST and HadISST1 during 1950-2015 (0.09ºC per decade; Table 1) are consistent on this ML-WNA space scale, but differ

locally at EB for this period and also deviate from the trends based on the *in situ* observations of SST at EB and AT at Halifax and Sable Island (Table B1).

Appendix C: Intercomparison of ERSST and HadISST1 for Major Oceanographic Regions

Another view of the comparison of the annual and seasonal time series from ERSST and HadISST1 averaged over the three major regions is shown in Fig. C1 where the differences between these series (calculated as ERSST minus HadISST1) are displayed for each season and region, and also for the annual means. Pronounced differences in the variability of the SST differences among the regions are apparent.

In the NE-NA, there is a small (magnitude typically ~0.2ºC) negative difference in summer during most of the record, with a steady decrease in magnitude since 1950. Prior to then, the summer difference fluctuated on a decadal time scale including a few years around 1940 when it had a positive value of ~0.5ºC similar to those in the other seasons. Otherwise, the NE-NA differences in the other seasons generally fluctuated decadally between 0 and 0.5ºC prior to 1940, and have been small and steady since 1945. Since this region has less spatial structure than the others, the differences in this region have the highest potential of being indicative of broad-scale systematic methodological biases between the two datasets.

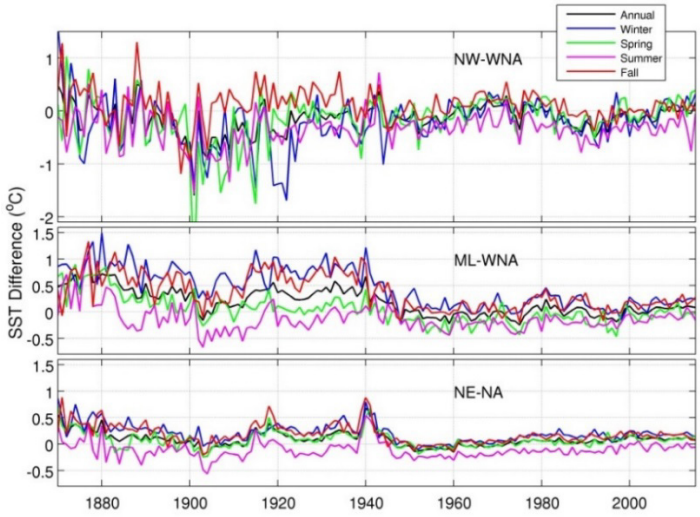

Figure C1. Time series of HadISST1 minus ERSST differences for the annual and four seasonal means in the three major regions. The same SST scale is used in each panel.

In the ML-NWA, the differences are larger but still with magnitudes typically less than 0.4ºC since the mid 1940s, during which time the spring and summer differences tended to be negative and the fall and winter ones positive. Prior to then, the differences generally had larger magnitudes, with positive values in fall and winter fluctuating on a decadal time scale with magnitude in the 0-1ºC range, generally negative values in summer with magnitude between 0 and 0.5ºC, and variable sign in spring fluctuating around 0 except for positive values near 0.5ºC prior to 1890. It appears that there was a methodological difference (e.g., sampling method or coverage) affecting at least one of the datasets in the ML-NWA prior to 1940. The comparison of

the ERSST and HadISST1 time series with the ICOADS data for the GM-SS area in Section 3b (e.g., Fig. 6b) indicates that low HadISST1 estimates in fall and winter were a contributor to this seasonally-varying difference.

In the NW-NWA, there is some similarity to the other regions with reduced differences since the late 1940s (compared to earlier), but also notable contrasts such as the larger magnitudes and seasonal fluctuations among the differences between 1895 and 1925. Also, the interannual-to-decadal variability in the NW-WNA differences is larger than in the other regions throughout the record, and prior to 1950 in particular. The data sparsity in the LS area prior to 1940, discussed in Section 3b (e.g. Fig. 6a), is likely a contributor to these differences; in particular, to those in winter and spring between 1895 and 1925 associated with the apparent flawed interpolation in ERSST.

Appendix D: Correlations of PC-Rs for 1900-2015 across Datasets and Seasons

In Table D1, we present the correlation coefficients between the PC-R time series from the ERSST and HadISST1 EOF analyses for 1900-2015, for each of the three leading PC-Rs and each of the annual and seasonal SST datasets (see Section 5b for discussion of the results of these analyses, and Fig. 10c,d for time series plots of the PC-Rs).

Table D1. Correlations between the corresponding PC-Rs in ERSST and HD, for each of the annual, winter, spring, summer and fall means, from the 1900-2015 EOF analyses with prior de-trending.

	PC-R1	PC-R2	PC-R3
Annual	**0.90**	**0.38**	**0.34**
Winter	**0.67**	**0.37**	*0.16*
Spring	**0.68**	**0.24**	0.11
Summer	**0.92**	**0.94**	**0.88**
Fall	**0.91**	**0.91**	**0.85**

Overall, the correlations between the HadISST1 and ERSST PC-Rs are highest for PC-R1, and in summer and fall for all three PC-R1s. The high correlation for PC-R1 is as might be expected for the predominant mode in terms of variance explained in all of the datasets, and the highest correlations in summer and fall are understandable since these are the seasons in which this mode explains the highest percentage of variance. The most surprising feature of the correlations for PC-R2 and PC-R3 is the large difference between the relatively low correlations in winter-spring and the relatively high ones in summer-fall, for both modes. The most likely reason for this is the difference in data coverage between these different "bi-seasons" (or half years) and a consequent increased sensitivity of the winter-spring modal structure to the different interpolation techniques used in the two datasets. In the case of PC-R2 which has some correlation with the winter NAO, it is somewhat surprising that the correlations are so low in winter-spring, but this is not inconsistent with this mode accounting for only 11-19% of the variance. It is curious that PC-R2 accounts for 19% of the variance in spring in HadISST1, while this season has the lowest PC-R2 correlation between the datasets; but, this is not inconsistent with spring being the season in which PC-R2 accounts for the least (11%) variance in ERSST.

In Table D2, we present the correlation coefficients between the annual and each seasonal PC-R, and among the seasonal PC-Rs, from each of the ERSST and HadISST1 EOF analyses. The relatively high inter-seasonal correlations among the PC-R1s (Table D2a) in both HadISST1 and ERSST are consistent with this mode being the predominant one and associated with the well-established AMO variation. The highest PC-R1 correlations between the annual and the spring, summer and fall means are consistent with previous findings and expectations. The

generally lower correlations between non-adjacent seasons are as expected, due to either or both of actual thermal inertia and temporal smoothing associated with the interpolations.

Table D2. Correlations between the annual and seasonal PC-Rs, and across the seasonal PC-Rs, for each PC-R in the 1900-2015 EOF analyses of each of ERSST and HadISST1 with prior de-trending. The correlations for HadISST1 are shown in the upper right portion of the table, those for ERSST in the lower left portion, and the unity diagonal applies to both. All correlations are significant at the 99% level.

(a)

PC-R1	Annual	Winter	Spring	Summer	Fall
Annual	**1**	**0.80**	**0.93**	**0.89**	**0.89**
Winter	**0.77**	**1**	**0.78**	**0.55**	**0.61**
Spring	**0.92**	**0.74**	**1**	**0.80**	**0.74**
Summer	**0.95**	**0.63**	**0.83**	**1**	**0.75**
Fall	**0.91**	**0.60**	**0.80**	**0.85**	**1**

(b)

PC-R2	Annual	Winter	Spring	Summer	Fall
Annual	**1**	**0.72**	**0.76**	**0.34**	**0.28**
Winter	**0.71**	**1**	**0.52**	*0.17*	**0.24**
Spring	**0.83**	**0.60**	**1**	**0.19**	0.05
Summer	**0.65**	**0.29**	**0.47**	**1**	**0.29**
Fall	**0.67**	**0.40**	**0.43**	**0.27**	**1**

(c)

PC-R3	Annual	Winter	Spring	Summer	Fall
Annual	**1**	**0.60**	**0.77**	**0.51**	0.12
Winter	**0.72**	**1**	**0.53**	**0.29**	0.08
Spring	**0.81**	**0.68**	**1**	0.12	-0.03
Summer	**0.81**	**0.31**	**0.53**	**1**	**0.24**
Fall	**0.57**	**0.25**	**0.27**	**0.39**	**1**

The correlations for PC-R2 (Table D2b) and PC-R3 (Table D2c) are generally lower than those for PC-R1, and the inter-seasonal correlations for the ERSST modes are generally lower than those for the HadISST1 modes (in spite of the greater spatial smoothing in the ERSST dataset; e.g., Fig. 9). The highest correlations of the annual PC-R2s and PC-R3s in both datasets are with their spring counterparts, consistent with spring being one of the seasons with the highest percentage variance explained in most of these (four) cases. As with the PC-R1s, there is a tendency for lower correlations for non-adjacent seasons as might be expected. For PC-R2 which has been identified with winter NAO forcing, it is noteworthy that the highest inter-seasonal correlations in both datasets are between winter and spring (0.52 in ERSST and 0.6 in HadISST1). For PC-R3, the highest inter-seasonal correlations are also between winter and

spring (0.68 in HadISST1 and 0.53 in ERSST), indicating that this PC-R may have a winter origin as well.

Supporting Information

Averaged Annual and Seasonal SST for Four Sub-periods

Figure SI-1 shows the averaged SST field during four different sub-periods, for each of the annual and seasonal time series from each of ERSST and HadSST1. The well-known seasonal variation of coolest temperatures (or greatest equatorward extent of cold subpolar water) occurring in winter, and warmest temperatures (or greatest poleward extent of warm water) in summer are apparent, associated with seasonal surface heating in spring and summer. The subpolar-subtropical transition zone extending from the MAB to Northeast Atlantic is apparent.

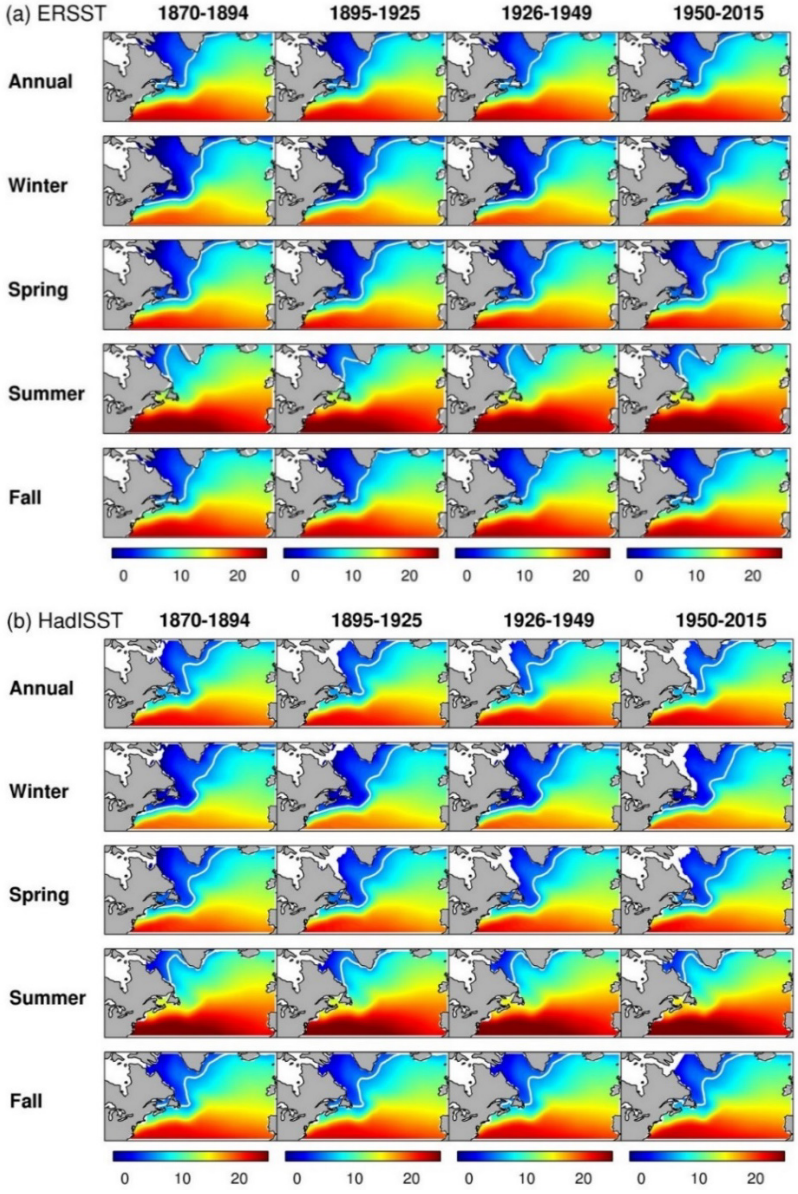

Figure SI-1. Maps of annual and seasonal SST from (a) ERSST and (b) HadISST1 for four sub-periods: 1870-1894, 1895-1925, 1926-1949 and 1950-2015. These fields were used to compute the difference maps in Fig. 4.

There is general similarity in all of the SST distributions; however, with the wide range of the colour bar needed to cover the 25°C latitudinal change, differences between the datasets and between the periods are hard to see. Closer examination reveals differences across the seasons, sub-periods and datasets which are more apparent in the difference plots (ERSST minus HadISST1) shown in Fig. 4 and discussed in Sections 3a and 3b.

Intercomparisons of Seasonal SST Time Series from Four Datasets in Key Areas with Persistent ERSST-HadISST1 Differences

Computations

Six areas with persistent differences between ERSST and HadISST1 are apparent in Fig. 4 and discussed in Section 3b. The areas are denoted as the Irminger Sea (IB), Labrador Sea (LS), Flemish Cap-Newfoundland Basin (FC-NB), Gulf of Maine – Scotian Slope (GM-SS), Gulf Stream North (GSN) and Sargasso Sea (SarS). The latitudes and longitudes of the approximate center positions of these areas are listed in Table SI-1. They were chosen to be even-integer numbers so that they correspond to both the center point of a 2° x 2° grid box in ERSST and the corresponding grid point in HadISST1. Time series of the annual means from ERSST and HadISST1 for each of the six areas during 1870-2015, together with those from the SODA 2.2.4 reanalysis product up to 2010 (Chepurin et al., 2014), are shown in Fig. 5. Time series of the seasonal means from ERSST, HadISST1, ICOADS 3.0 (up to 2014) and SODA 2.2.4 for the LS and GM-SS are shown in Fig. 6, for the IB and FC-NB in Fig. SI-2, and for the NGS and SarS in Fig. SI-3. Correlations among the annual and seasonal means from the four gridded datasets are presented in Table SI-2.

Table SI-1. Center positions chosen for the six key ERSST-HadSST1 (ER-HD) difference areas, together with the center positions used for the 25-grid-point averaging of the SODA dataset (same except for the IB where the SODA position is displaced to the southeast by 0.5° of each of latitude and longitude). Also shown are the latitude and longitude boundaries of the 4° x 4° COADS box for which spatially-averaged (over 4 grid boxes) seasonal time series were extracted.

Area	ER & HD Positions	COADS3.0 Grid Box	SODA Center Positions
IB	60°N, 40°W	58-62°N, 38-42°W	59.5°N, 39.5°W
LS	58°N, 52°W	56-60°N, 50-54°W	58°N, 52°W
FC-NB	46°N, 46°W	44-48°N, 44-48°W	46°N, 46°W
GM-SS	42°N, 66°W	40-44°N, 64-68°W	42°N, 66°W
GSN	40°N, 56°W	38-43°N, 54-58°W	40°N, 56°W
SarS	34°N, 60°W	32-36°N, 58-62°W	34°N, 60°W

The ICOADS seasonal means were taken as the average of the appropriate monthly medians in the four 2° x 2° grid boxes surrounding each grid position (see Table SI-1 for the

boundaries of these grid boxes). The averages were retained if any (of the potential 12) medians were available from ICOADS, thereby providing an indication of minimal data availability (but not necessarily a robust estimate of seasonal-mean SST in the area). The HadISST2 dataset of monthly statistics in 5º x 5º grid boxes (Rayner et al., 2006) was also investigated and generally indicated similar data availability/sparsity, but this is not reported here since these boxes were generally offset from the key areas' center positions resulting in additional SST biases. The SODA annual and seasonal means were computed as the averages over the appropriate months and 25 grid points (0.25º latitude x 0.4º longitude resolution) centred on each area of discrepancy (Table SI-1), with the exception of a slight displacement to the southwest for the IB area (in order to provide better agreement of its SST means with those of the interpolated datasets in this SST gradient area). Considering the 1º x1º and the smoothed 2º x 2º resolutions of HadISST1 and ERSST, respectively, these different spatial resolutions were taken as a compromise between retaining representative observations from ICOADS and approximate consistency in resolution.

Annual Means

The annual time series in all six areas (Fig. 5) are discussed in Section 3b.1. There are persistent differences of ~1ºC or greater between ERSST and HadISST1 over all or most of the records in the IB, LS, FC-NB and GM-SS, and between the 1950s and 1990s in the NGS, and a small persistent difference in the SarS throughout the record. The differences correspond to relatively high ERSST values in the IB, FC-NB, GM-SS and SarS, and relatively low ERSST values in the LS and NGS. Correlations between the ERSST and HadISST1 annual means are in the 0.55-0.83 range (Table SI-2).

The SODA annual SSTs are generally between or in the vicinity of the ERSST and HadISST1 ones in the IB, LS, FC-NB and SarS, but outside the latter's range for substantial portions of the records in the GM-SS and NGS. They are in better agreement with ERSST throughout the record for the IB and during the 1960s to 1980s in the FC-NB, but in better agreement with HadISST1 through most of the LS record (especially between 1890 and 1940) and most of the time between the 1880s and 1950 in the FC-NB. Correlations of the SODA annual SSTs with HadISST1 are in the 0.64-0.85 range, except only 0.54 for the FC-NB (Table SI-2). Those with ERSST are in the 0.60-0.76 range, except only 0.35 for the NGS and 0.50 for the FC-NB.

Seasonal Means: LS, IB and FC-NB

Comparison of the seasonal time series for the LS from the four gridded datasets (Fig. 6a) is discussed in detail in Section 3b.2. It appears that data sparsity in winter prior to the mid 1940s, and to a lesser degree in the other seasons, together with the interpolation and/or smoothing methodology in ERSST, are the primary contributors to the low ERSST values in this area during this period.

The annual interpolated series (Fig. 5) for IB and FC-NB have notable similarity in ERSST being higher than HadISST1 throughout the study period, and by 1-2°C most of the time. The seasonal series for both (Fig. SI-2) indicate that there are contributions in all seasons to this difference. In the IB, a situation somewhat similar to LS in relation to a data-sparsity origin, but notably different in that the differences are of the opposite sign (Fig. 5), occurs prior to 1980. The ERSST seasonal means are consistently higher than the HadISST ones, and more closely follow the ICOADS ones (prior to the 1990s), and the ERSST-HadISST1 differences are largest in winter which has virtually no observations prior to the 1930s. ERSST appears to be more reliable than HadISST1 since 1940 in this area, but the more reliable dataset prior to then (and especially in winter) and the source of the discrepancies are unclear beyond a possible interpolation issue in this area where there is a substantial spatial gradient in SST (Fig. SI-1). The SODA seasonal means track both the ICOADS and ERSST means well since 1950, and lie between the ERSST and HadISST1 ones in winter between the 1880s and 1920s when there were no observations (but this is partly due to our offset of the center position for SODA).

The FC-NB area (Fig. SI-2b) is similar to the IB in that the ERSST means are consistently larger than the HadISST1 ones and follow the ICOADS means more closely since 1910 in all seasons, except for summer in FC-NB when they all track closely (r ~0.9 for the interpolated datasets, and r ~0.7 between each and ICOADS). Further, the enhanced interannual variability in ICOADS in spring and winter prior to 1910 suggests a data sparsity issue in this area, too. On the other hand, the areas differ in that the ICOADS means for FC-NB tend to be larger than those from both interpolated datasets in winter

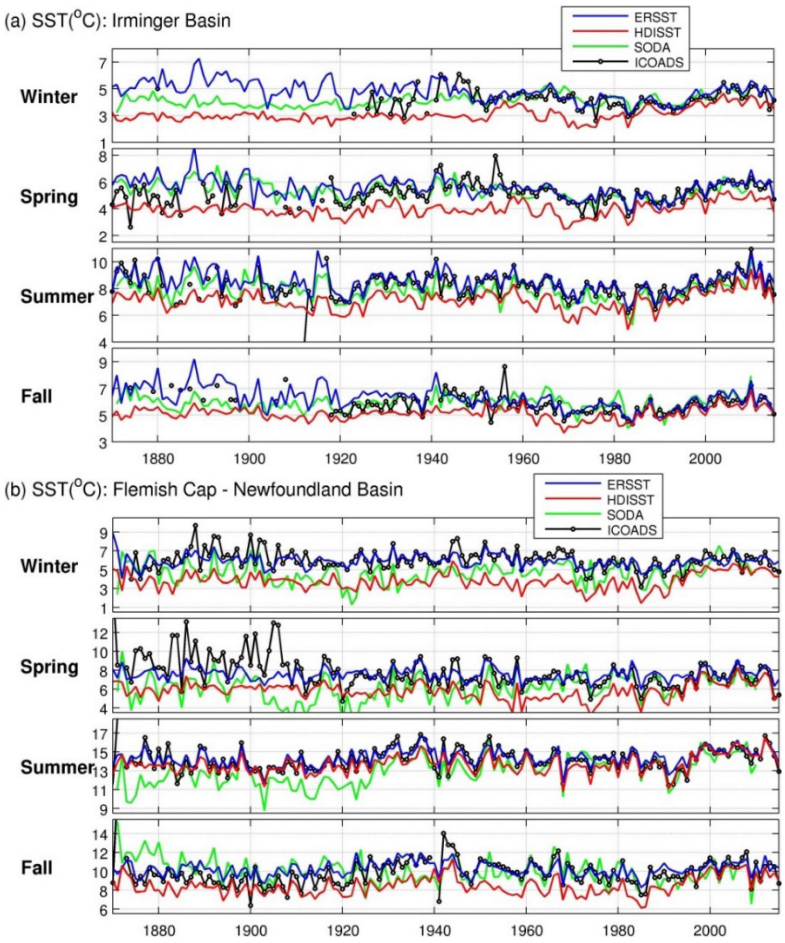

Figure SI-2. Time series of seasonal SST from ERSST, HadISST1, SODA and ICOADS3 for the (a) IB and (b) FC-NB.

and spring before 1910, and the SODA means for the FC-NB in spring and summer prior to 1930 have substantial discrepancies from the others. Discrepancies in this complex offshore region where the NA's upper-ocean western boundary currents compete are not surprising.

Seasonal Means: SS-GM, GSN and SarS

Comparison of the seasonal time series for the SS-GM (Fig. 6b) is discussed in detail in Section 3b.2. There is generally good agreement among the ERSST, HadISST1 and ICOADS seasonal means in all seasons since 1940, but persistent difference prior to then. The latter together with enhanced interannual variability in the ICOADS means in the early decades (especially prior to 1900) points to data sparsity and undersampling issues. It appears that the HadISST1 means prior to 1940 are underestimates, and the ICOADS data prior to 1900 point to a possibility of the same for the ERSST means (except in fall) but by a lesser amount.

The annual time series (Fig. 5) indicate that ERSST is higher than HadISST1 in the SarS by up to ~1°C prior to 1950 and by up to ~0.5°C since 1980. In contrast, in the GSN, ERSST is

lower than HadISST1 through most of the record, and especially between the mid 1940s and mid 1990s with differences up to 1ºC.

The seasonal means from all four datasets follow each other relatively well in all seasons in the SarS since 1940, with the slightly higher ERSST annual means since 1980 and prior to 1950 receiving small contributions from all seasons except summer (Fig. SI-3). There is an indication that the fall and winter HadISST1 means follow the ICOADS ones more closely (than the ERSST ones) prior to 1940, but there is increased interannual variability in the ICOADS series prior to 1920 and little ICOADS data prior to 1880 (but note that this is for only a small area of the actual Sargasso Sea).

In the GSN, the ERSST, HadISST1 and ICOADS summer and fall means follow each other closely since 1920, but the winter and spring HadISST1 means are generally higher than the ERSST ones and follow the corresponding ICOADS means more closely. Prior to this the ERSST-HadISST1 differences are more variable with season and year. The ICOADS means have greater interannual and interseasonal variability prior to 1905, and substantial deviations from the interpolated SSTs which probably reflect data sparsity in the presence of the strong spatial variability in this region (e.g., Gulf Stream meanders and eddies). It is noteworthy than none of these SST time series (in either the NGS or GM-SS) show the anomalously low temperatures that have been suggested to have occurred in the bottom water over the MAB slope in the spring of 1882 as a probable contributor to the major tilefish kill at that time (Marsh et al., 1999), but this could be because we have not examined that specific area.

The SODA seasonal means in the SarS show good approximate tracking of both interpolated SST series in all seasons since 1940, and of the fall and winter HadISST1 series between 1920 and 1940, but poorer agreement before then. On the other hand, in the NGS, the SODA series show significant deviations from the interpolated series throughout the records in fall, winter and spring, although generally not as large as in the GM-SS in spring and summer (Fig. 7b). It is beyond our scope to pursue further the origin of these discrepancies, although they (together with those for the FC-NB) point to a deficiency in the SODA reanalysis for the region north of the Gulf Stream in the WNA. This seems consistent with the poorer agreement found by Chepurin et al. (2014) for sea level along the east coast of North America than elsewhere in the global ocean.

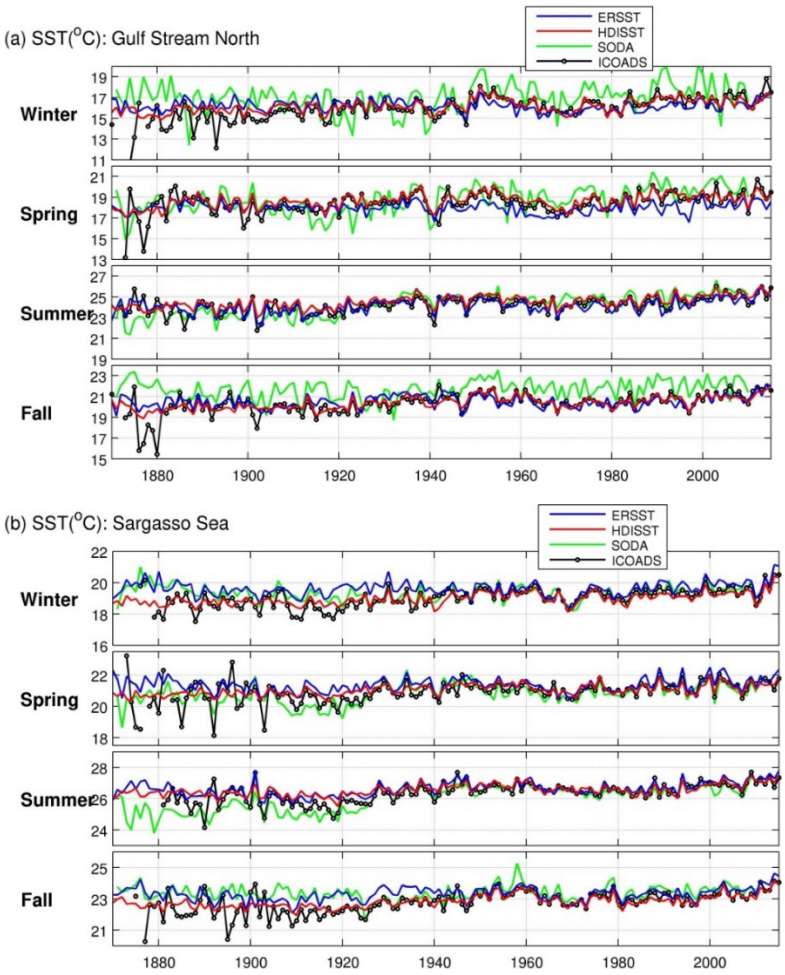

Figure SI-3. Time series of seasonal SST from ERSST, HadISST1, SODA and ICOADS3 for the (a) GSN and (b) SarS.

Correlations between the SODA annual means and those from the interpolated datasets are again only moderate, with higher values for HadISST1 except for ERSST in the IB and a notably low value (0.35) for ERSST in the NGS (Table D1).

Table SI-2. Correlation coefficients (r) between each of ERSST (ER) and HadISST1 (HD), and (i) the corresponding ICOADS time series for each season during 1870-1914, and (ii) the corresponding SODA time series during 1871-2010 for each of the annual and four seasonal means, in the six areas with persistent ER-HD differences. The correlations between ER and HD during the SODA period are also shown, as well as the number (N) of years with ICOADS means for each season and area. The parenthesized correlations for ICOADS in the Annual rows are the averages of the four seasonal ones. The various time series are displayed in Figs. 5, 6, SI-2 and SI-3. Correlations significant at the 95% confidence level are in **boldface**, and those significant at the 90% level are in *italics*.

Area	Period	ER-HD	ER-CD		HD-CD	ER-SD	HD-SD
		r	r	N	r	r	r
Irminger Basin	Annual	**0.55**	(0.63)		**0.76**	**0.68**	(0.53)
	Winter	*0.15*	**0.58**	87	*0.16*	**0.43**	**0.38**
	Spring	**0.53**	**0.52**	127	**0.70**	**0.44**	**0.49**
	Summer	**0.72**	**0.74**	131	**0.76**	**0.73**	**0.74**
	Fall	**0.57**	**0.66**	107	**0.48**	**0.50**	**0.52**
Labrador Sea	Annual	**0.56**	(0.50)		**0.60**	**0.69**	(0.49)
	Winter	0.00	0.07	71	**0.26**	**0.40**	**0.35**
	Spring	**0.35**	**0.53**	125	**0.19**	**0.47**	**0.32**
	Summer	**0.71**	**0.78**	125	**0.61**	**0.62**	**0.67**
	Fall	**0.67**	**0.60**	92	**0.58**	**0.62**	**0.60**
Flemish Cap – Newfound-land Basin	Annual	**0.72**	(0.63)		**0.50**	**0.54**	(0.62)
	Winter	**0.65**	**0.61**	142	**0.29**	**0.51**	**0.56**
	Spring	**0.64**	**0.46**	145	**0.47**	**0.39**	**0.52**
	Summer	**0.87**	**0.84**	145	**0.69**	**0.66**	**0.79**
	Fall	**0.75**	**0.62**	144	**0.30**	**0.30**	**0.59**
Gulf of Maine – Scotian Slope	Annual	**0.83**	(0.59)		**0.72**	**0.85**	(0.57)
	Winter	**0.67**	**0.60**	142	**0.51**	**0.63**	**0.59**
	Spring	**0.86**	**0.51**	144	**0.69**	**0.74**	**0.36**
	Summer	**0.72**	**0.53**	143	**0.63**	**0.77**	**0.58**
	Fall	**0.70**	**0.73**	144	**0.36**	**0.44**	**0.76**
North Gulf Stream	Annual	**0.64**	(0.56)		**0.35**	**0.64**	(0.72)
	Winter	**0.56**	**0.42**	142	0.10	**0.36**	**0.73**
	Spring	**0.58**	**0.46**	145	**0.30**	**0.54**	**0.60**
	Summer	**0.79**	**0.79**	144	**0.60**	**0.74**	**0.82**
	Fall	**0.64**	**0.55**	144	**0.22**	**0.33**	**0.72**
Sargasso Sea	Annual	**0.71**	(0.64)		**0.72**	**0.81**	(0.68)
	Winter	**0.62**	**0.53**	141	**0.63**	**0.47**	**0.69**
	Spring	**0.67**	**0.55**	141	**0.66**	**0.64**	**0.51**
	Summer	**0.79**	**0.84**	136	**0.65**	**0.71**	**0.77**
	Fall	**0.67**	**0.60**	141	**0.51**	**0.53**	**0.75**