Supplement for

SODA4: a mesoscale ocean/sea ice reanalysis 1980-2024

Gennady A. Chepurin, James A. Carton, Luyu Sun, and Stephen G. Penny

Text: S1
Tables: S1-S6
Figures: S1-S18

Introduction

This document contains six tables. **Table S1** contains the depths of the 75 levels. **Tables S2-S5** contain lists of variables contained in the original grid and regridded 5dy and monthly ocean and sea ice files, as well as the original grid transport files and the temperature and salinity increment files. **Table S6** lists the original grid 5dy files that are missing from the archive and unrecoverable.

This document contains 17 figures. **Fig. S1** shows the difference between the GEBCO 2014 grid used in this reanalysis and the latest release of the GEBCO topography (GEBCO 2024) showing that the depth differences are small. **Fig. S2** shows the basin definitions. **Figs S3-S6** provide time series describing the SODA4 discharge data set which is discussed in **Text S1**. **Fig. S7** shows time series of temperature and salinity observations for each basin. **Fig. S8** shows the impact of changes in the choice of advective scheme on 3000m temperature**.** **Figs S9** and **S10** show the decade means of temperature and salinity analysis increments. **Fig. S11** shows the standard deviation of monthly 0-300m temperature increments. **Fig. S12** shows the standard deviation of monthly 0-300m temperature anomaly from monthly climatology for SODA4 and EN4.2.2 in the Southern Hemisphere. **Fig. S13** compares the regression of 0-300m average temperature on the NINO3.4 SST index time series for EN4.2.2 and SODA4. **Fig. S13** provides a Hovmoller plot of EN4.2.2 and SODA4 0-300m temperature anomaly from their climatological monthly cycle with time along 0°N. **Figs S15** through **S17** present transport time series for three major straits: Bering Strait, Fram Strait, and Drake Strait, whose time mean values are given in **Table 2**. Finally, **Fig. S18** presents the Atlantic meridional overturning transport time series across 26.5°N

Text S1
The construction of the SODA4 global monthly gridded continental discharge data set builds on the compilation by *Dai* (2017) and Dai and Trenberth (2020) of gauge-based discharge estimates for the 925 largest rivers, which in turn is based on the Global Runoff Data Centre river discharge data (www.bafg.de/GRDC/EN/, *Wilkinson et al*, 2014).

The Dai and Trenberth data set cannot be used without modification because of the presence of missing data which increases dramatically with time (**Fig. S3**). Explanations for this problem include loss of trained personnel, reduction in funding for maintenance, and a lack of interest by some countries in making hydrologic information publicly available (*Chandanpurkar et al., 2017*). Examples of missing data include: the gauging station record on the Orinoco River at Puente Angostura which ends January 1989, the gauging station on the Congo River at Kinshasa which ends March 2010 and the gauging station record on the Amazon River at Obidos which ends January 1998. In the Arctic the gauging station on the Yenisey River at Igarka ends November 2015, the gauging station on the Ob River at Salekhard ends November 2015, and the gauging station on the Lena River at Stolb ends November 2002.

To address this problem we first attempt to find alternate sources of monthly discharge for major river systems. **Fig. S4** shows the Dai and Trenberth record of monthly discharge for the Ob and Yenisei rivers into the Arctic Ocean in comparison to the SODA4 discharge updated based on data from arcticgreatrivers.org (*Shiklomanov et al., 2020*). We also add monthly discharge from ice sheets such as Greenland which are entirely missing from Dai and Trenberth (**Fig. S5**). Many gaps remained after this initial exercise. The remaining gaps we fill with climatological monthly values, inflating the implied discharge rates to account for the ~30-40% of the discharge that Dai and Trenberth (2002) estimate is from ungauged sources. After this the discharge rates are adjusted basin-by-basin to balance the basin-integrated time-mean water lost through net evaporation (estimated from the ERA5 climatology). The final SODA4 monthly time series of global discharge is shown in **Fig. S6**. Its time mean, 1.14×10^6 m^3/s, falls at the lower end of the span of recently published estimates: 1.1-1.4×10^6 m^3/s (*Syed et al., 2010; GRDC, 2014; Wilkinson et al., 2014; Müller-Schmied*, et al. *2014*).

Table S1 Depths of model levels for thermodynamic variables (T,S) and for horizontal velocity (u,v)

Level	Z* Depth (m)		Level	Z* Depth (m)		Level	Z* Depth (m)	
#	T,S,w	u,v	#	T,S,w	u,v	#	T,S,w	u,v
1	0.54	1.08	26	121.49	128.06	51	1453.4	1515.4
2	1.68	2.28	27	135.31	142.57	52	1581.2	1646.9
3	2.94	3.60	28	150.59	158.60	53	1716.2	1785.5
4	4.33	5.06	29	167.45	176.30	54	1858.3	1931.0
5	5.87	6.68	30	186.08	195.85	55	2007.0	2082.9
6	7.57	8.46	31	206.64	217.42	56	2161.9	2240.9
7	9.45	10.43	32	229.33	241.23	57	2322.6	2404.3
8	11.52	12.61	33	254.37	267.51	58	2488.5	2572.7
9	13.82	15.02	34	281.99	296.48	58	2659.2	2745.6
10	16.35	17.68	35	312.45	328.42	60	2834.1	2922.5
11	19.15	20.62	36	346.02	363.61	61	3012.6	3102.8
12	22.25	23.87	37	382.99	402.37	62	3194.5	3286.1
13	25.67	27.46	38	423.69	445.01	63	3379.1	3472.1
14	29.45	31.43	39	468.45	491.89	64	3566.1	3660.2
15	33.62	35.82	40	517.62	543.36	65	3755.3	3850.3
16	38.24	40.66	41	571.59	599.82	66	3946.2	4042.0
17	43.34	46.02	42	630.73	661.64	67	4138.6	4235.1
18	48.98	51.94	43	695.42	729.21	68	4332.2	4429.3
19	55.21	58.48	44	766.07	802.92	69	4526.9	4624.5
20	62.09	65.70	45	843.03	883.13	70	4722.5	4820.5
21	69.69	73.69	46	926.65	970.17	71	4918.8	5017.2
22	78.10	82.51	47	1017.24	1064.32	72	5115.8	5214.4
23	87.38	92.25	48	1115.07	1165.82	73	5313.3	5412.1
24	97.64	103.02	49	1220.31	1274.80	74	5511.2	5610.2
25	108.97	114.92	50	1333.08	1391.35	75	5709.4	5808.7

Table S2 Ocean state variables at 5dy or monthly intervals available on both the (3600x2700x75L) original grid and the 0.1°x0.1° (3600x1700) REGRIDED lon-lat grid. Velocity components on the original grid are defined relative to the numerical grid.

Variable	Also regridded	Long name	Units
temp	yes	Potential temperature	°C
salt	yes	Salinity	PSU
u	yes	Zonal velocity	m s-1
v	yes	Meridional velocity	m s-1
wt	yes	Vertical velocity	m s-1
prho		Potential density	kg m-3
ssh	yes	Sea Surface Height (effective sea level (eta_t + patm/(rho0*g)) on T cells)	m
mlt	yes	Mixed layer depth determined by temperature (ΔT=0.2K)	m
mlp	yes	Mixed layer depth determined by potential density ($\lvert\Delta S\rvert$=0.01PSU)	m
mls	yes	Mixed layer depth determined by salinity ($\Delta\rho$=0.0003 kg/m^3)	m
taux		Zonal wind stress	N/m^2
tauy		Meridional wind stress	N/m^2
net_heating		surface ocean heat flux coming through coupler and mass tr	W/m^2
evap_heat		Latent heat flux into ocean (<0 cools ocean)	W/m^2
hflux_pme		Heat flux (relative to 0C) from PME transfer of water across ocean surface	W/m^2
hflux_total		Surface heat flux from coupler plus restore (omits mass transfer heating	W/m^2
lw_heat		Longwave flux into ocean (<0 cools ocean)	W/m^2
sens_heat		Sensible heat into ocean (<0 cools ocean)	W/m^2
pme_river		mass flux of precip-evap+river via sbc (liquid, frozen, evaporation)	kg m^-2 s^-1
river		mass flux of river (runoff + calving) entering ocean	kg m^-2 s^-1

Table S3 5-day and monthly sea ice variables available on both the (3600x2700) original grid and on the uniform 0.1°x0.1° (3600x1700) lon-lat grid. Several variables are specified at multiple levels.

Variable	Also regridded	Long name	Units
bheat	yes	ocean to ice heat flux	W/m^2
bmelt	yes	bottom surface melting energy flux	W/m^2
cn	yes	ice concentration	0-1
e2melt	yes	heat needed to melt ice	J/m^2
hfrazil	yes	energy flux of frazil formation	W/m^2
hi	yes	ice thickness	m-ice
hs	yes	snow thickness	m-snow
mi	yes	ice mass	kg/m^2
tmelt	yes	upper surface melting energy flux	W/m^2
ui	yes	ice velocity - x component	m/s
vj	yes	ice velocity - y component	m/s
ix_trans		x-direction ice transport	kg/s
iy_trans		y-direction ice transport	kg/s
sh		sensible heat flux	W/m^2
lh		latent heat flux	W/m^2
sw		short wave heat flux	W/m^2
lw		long wave heat flux over ice	W/m^2
saltflx		ice to ocean salt flux	kg/(m^2*s)
qfres		Restoring heat flux	W/m^2

Table S4 Ocean transport variables at 10-day intervals on the (3600x2700x75L) original grid (the vertically summed transports are available at one level).

Variable	Long name	Units
tx_trans	T-cell i-mass transport	10^9 kg/s
ty_trans	T-cell j-mass transport	10^9 kg/s
tx_int	T-cell i-mass transport vertically summed	10^9 kg/s
ty_int	T-cell j-mass transport vertically summed	10^9 kg/s

Table S5 1°x1° gridded analysis increments ($\mathbf{K}[\omega^o - \mathbf{H}\omega^f]$) produced by the ocean data assimilation at 10 dy and monthly intervals on the 1°x1°75L grid.

Variable	Long name	Units
tcor	temperature increment	deg-C/10dy
scor	salinity increment	PSU/10dy

Table S6 Missing 5-dy and 10-dy average original grid output files.

Ocean	Sea ice	Transport
01/01/1980- 12/31/1982	06/15/1986	01/01/1981- 12/31/1981
06/15/1986		12/23/1993
12/20/1994		2/10/2000
12/25/1994		1/3/2009
12/30/1994		1/13/2009
07/22/2016		7/25/2016
07/27/2016		9/27/2021

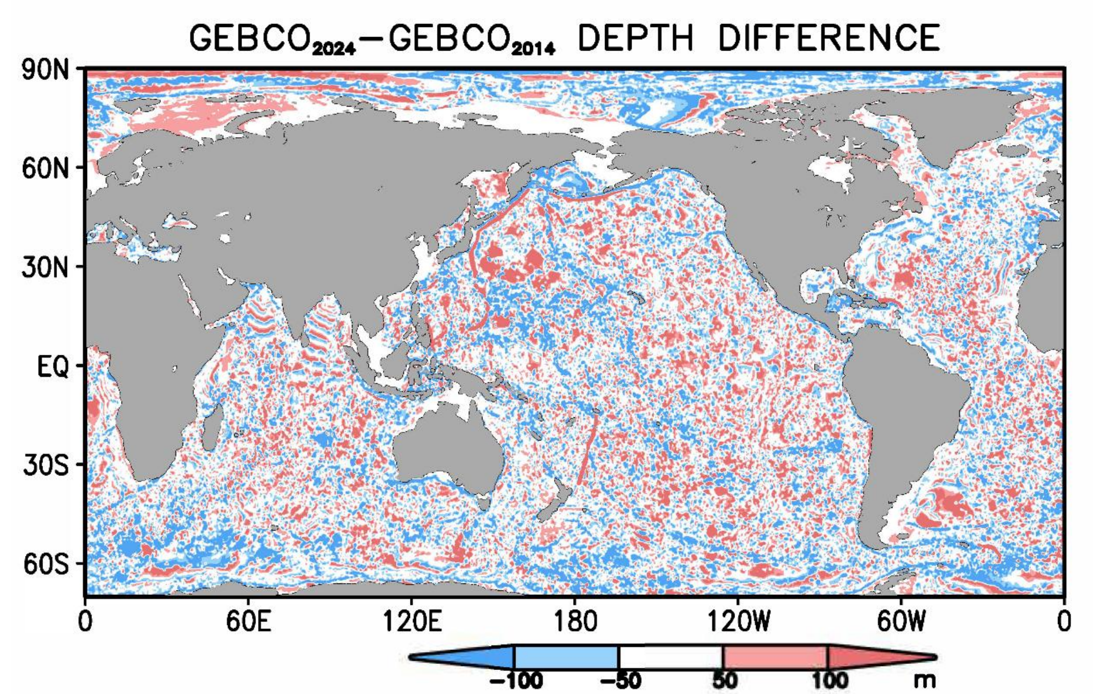


Fig. S1 Differences between the GEBCO 2014 topography used in SODA4 and the 2024 release of the GEBCO topography (www.gebco.net). Most differences have small horizontal scales.

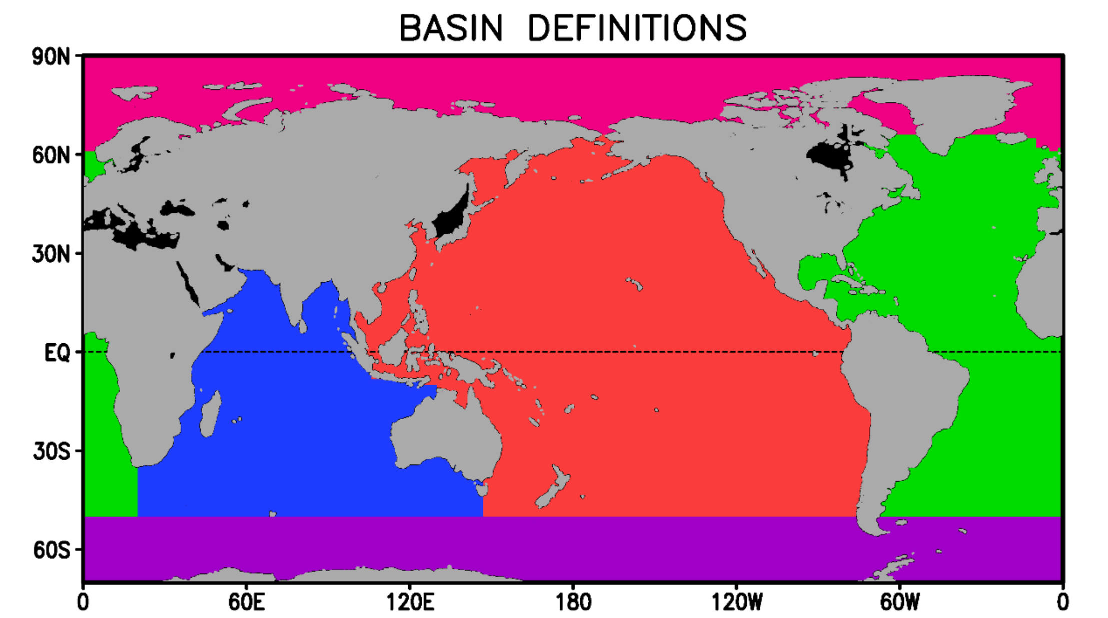
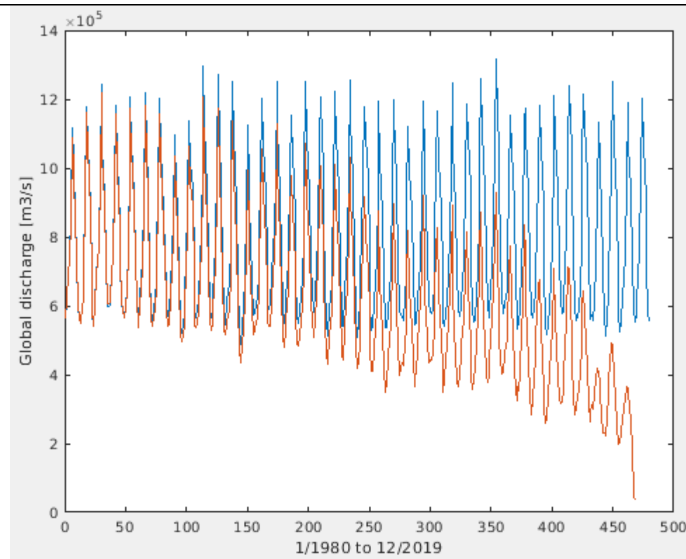
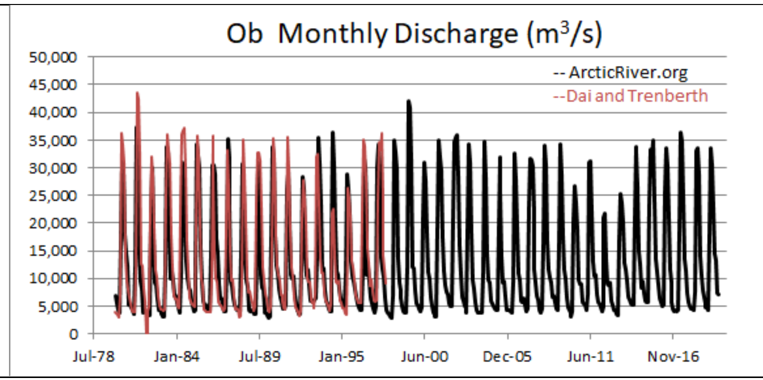


Fig. S2 Basin definitions for the Pacific, Atlantic, Indian, Arctic, and Southern Oceans. These definitions are used when producing basin-average quantities.

Fig. S3. Global sum of monthly discharge during 1980-2019. (Orange) Dai (2017). Discharge begins to decline in the 1980s due to loss of data. (blue) SODA4. Units: 10^5 m^3/s.	



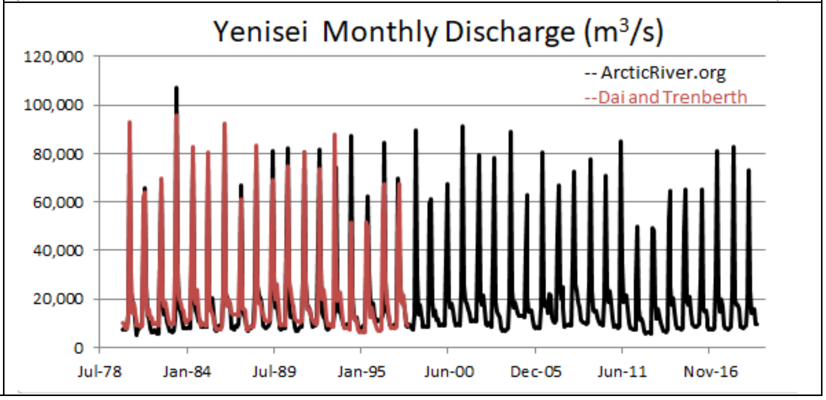


Fig. S4 Eurasian rivers Ob and Yenisei that discharge into the Arctic Ocean. (red) *Dai (2017)*, (black) SODA4. Units: m^3/s

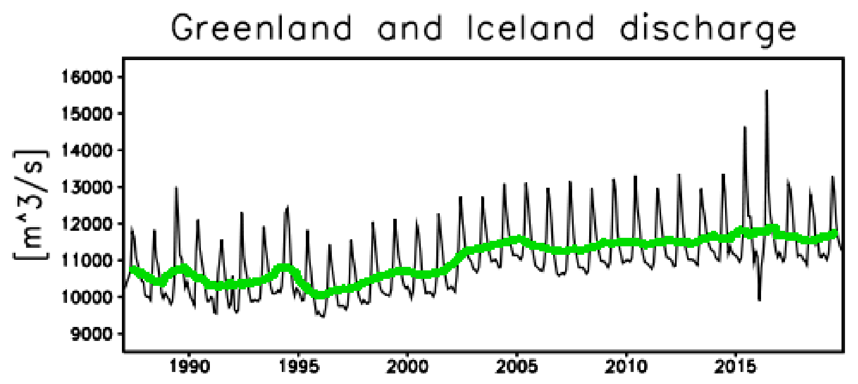


Fig.S5 (black) Sum of monthly Greenland plus Iceland discharge. (green) annual running average. Units: m^3/s.

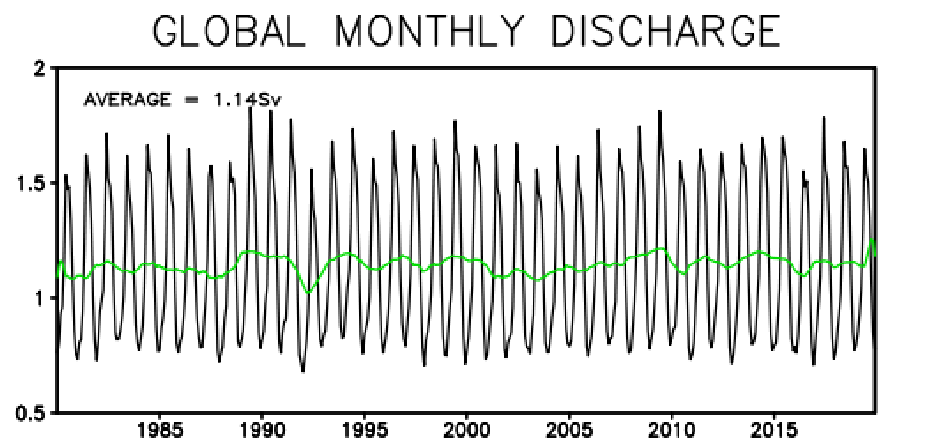


Fig. S6 SODA4 global monthly discharge. Similar to **Fig. S3**, but corrected for ungauged discharge. (green) Annual running mean. Units: 10^6 m^3/s.

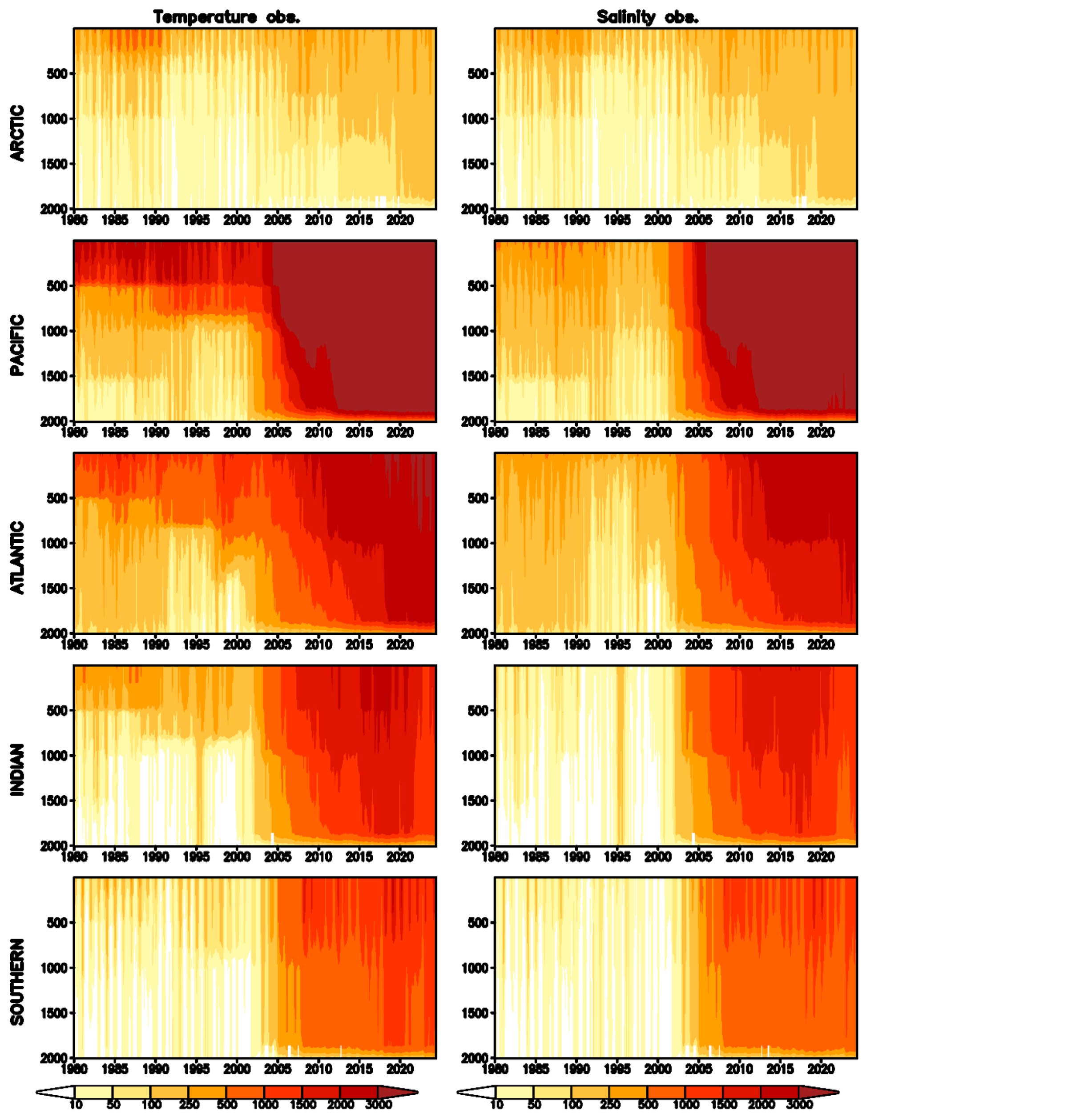


Fig. S7 (left) temperature and (right) salinity profiles with depth expressed as numbers per (1°x1°x10dy) cell per year. Note the dramatic increase in observational coverage following the deployment of the Argo system in the early 2000s.

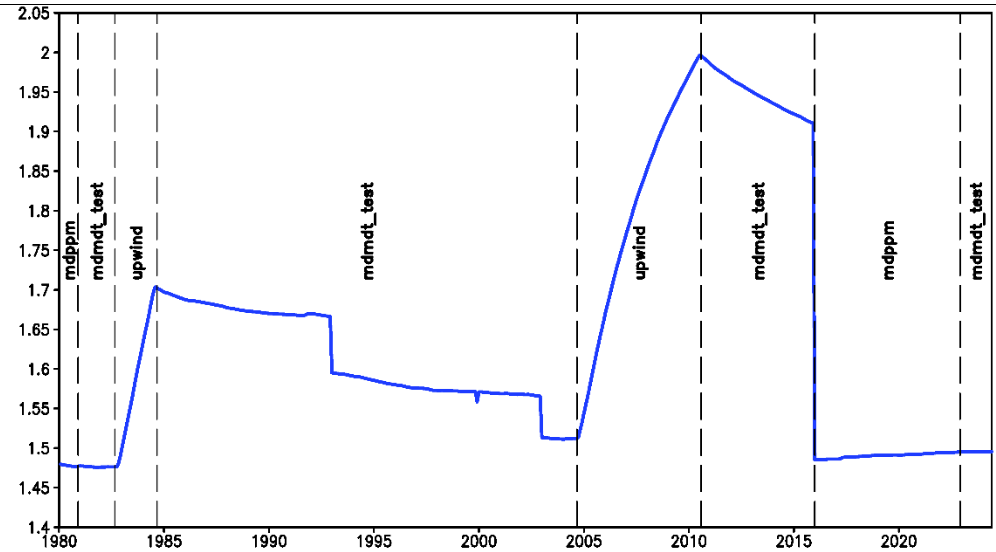


Fig. S8 Global average potential temperature at 3000m depth with time 1980-2024. Three advective schemes were tried: two monotonicity-preserving schemes of Daru and Tenaud (2003) and a third order upwind advective scheme. Jumps in 3000m global average temperature coincide with occasions when the advective scheme was changed to improve numerical stability. For comparison, the time mean global average potential temperature at 3000m depth is 1.72°C according to the World Ocean Atlas 2023.

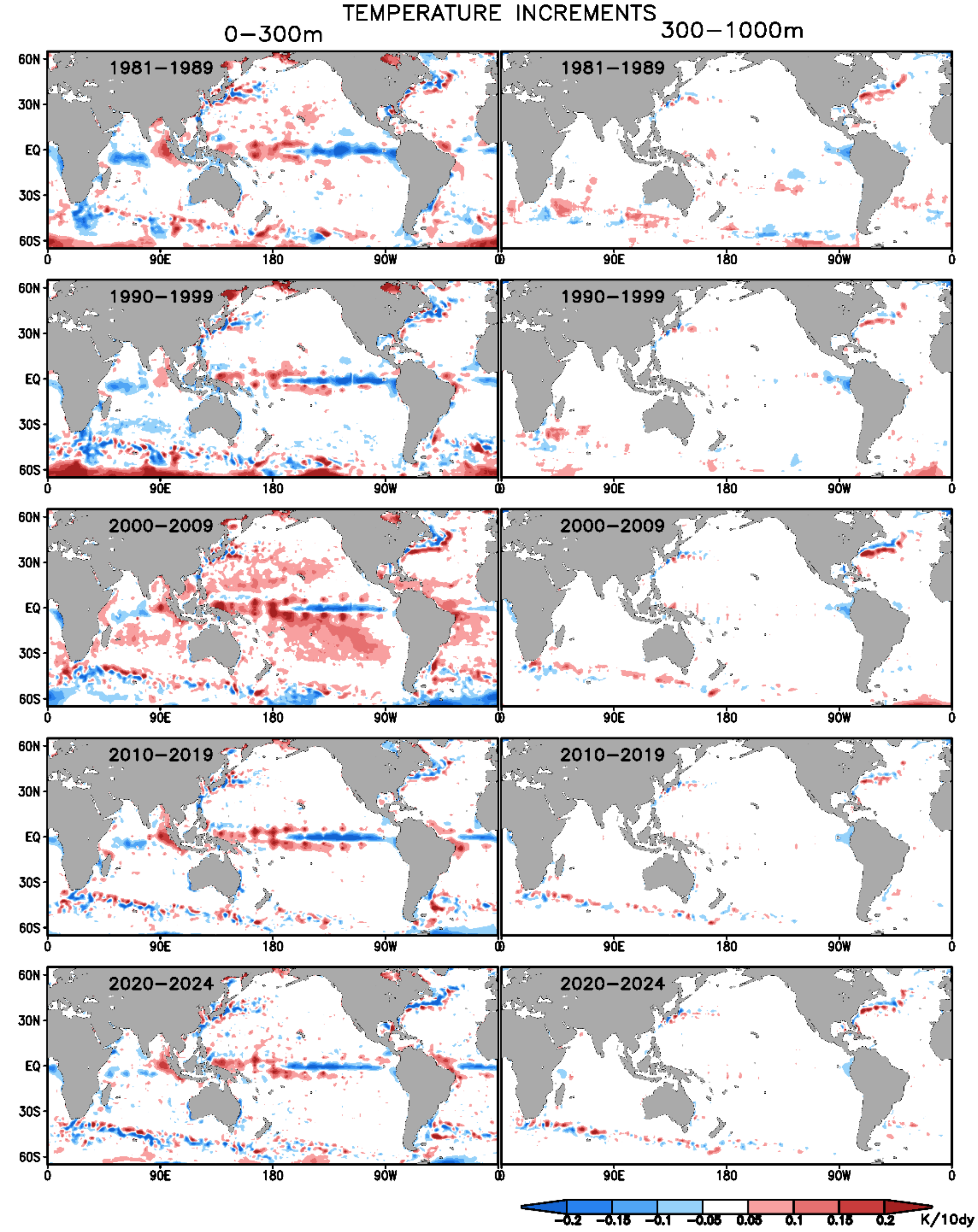


Fig. S9 Mean (left) 0-300m and (right) 300-1000m temperature analysis increments averaged by decade.

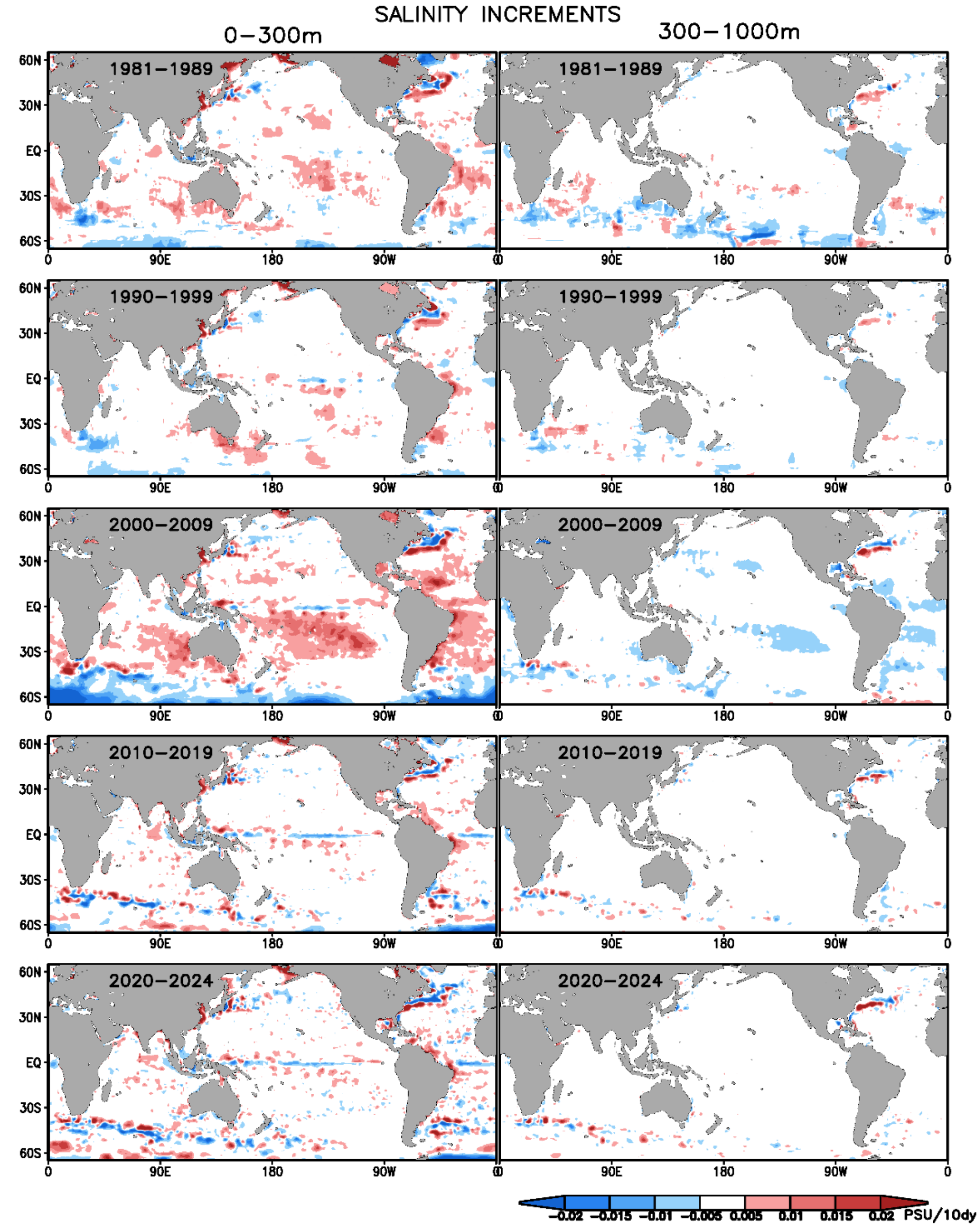


Fig. S10 Mean (left) 0-300m and (right) 300-1000m salinity analysis increments averaged by decade.

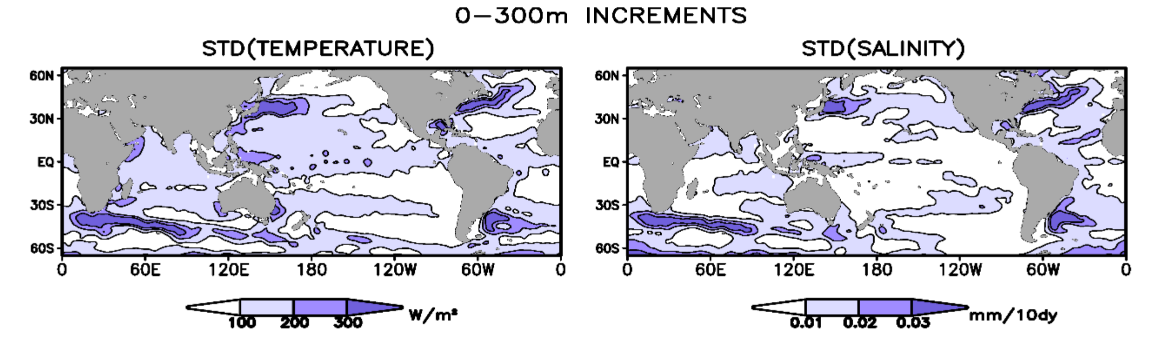


Fig. S11 Standard deviation of (left) temperature and (right) salinity increments integrated vertically 0-300m, expressed in units of Wm^{-2} and mm/10dy.

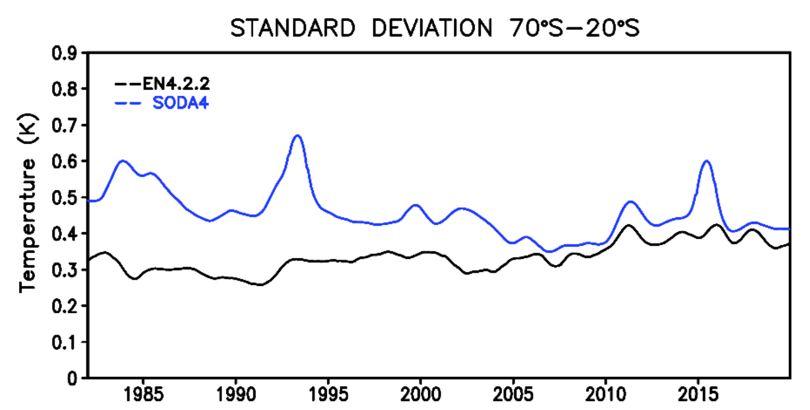


Fig. S12 Standard deviation of monthly 0-300m temperature anomaly from monthly climatology plotted versus time in the latitude band 70°S-20°S. The time series have been annually smoothed. Peaks in SODA4 variability coincide with major El Nino events.

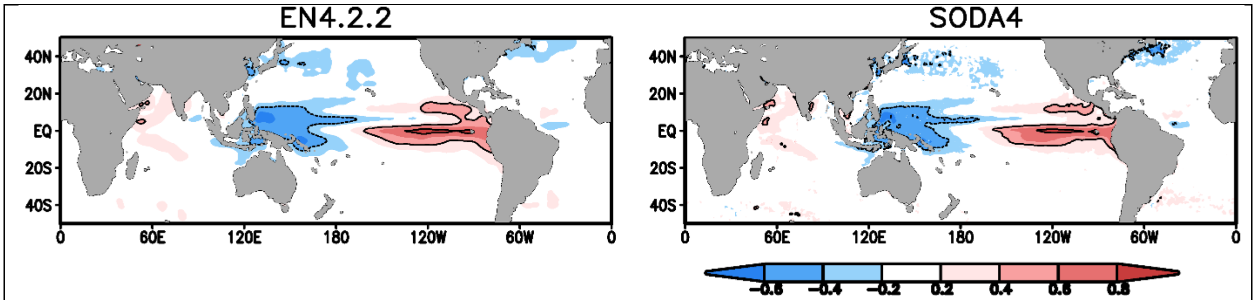


Fig. S13 Vertically averaged monthly (black) EN4.2.2 and (blue) SODA4 temperature by basin. (left) 0-300m, (right) 300-1000m

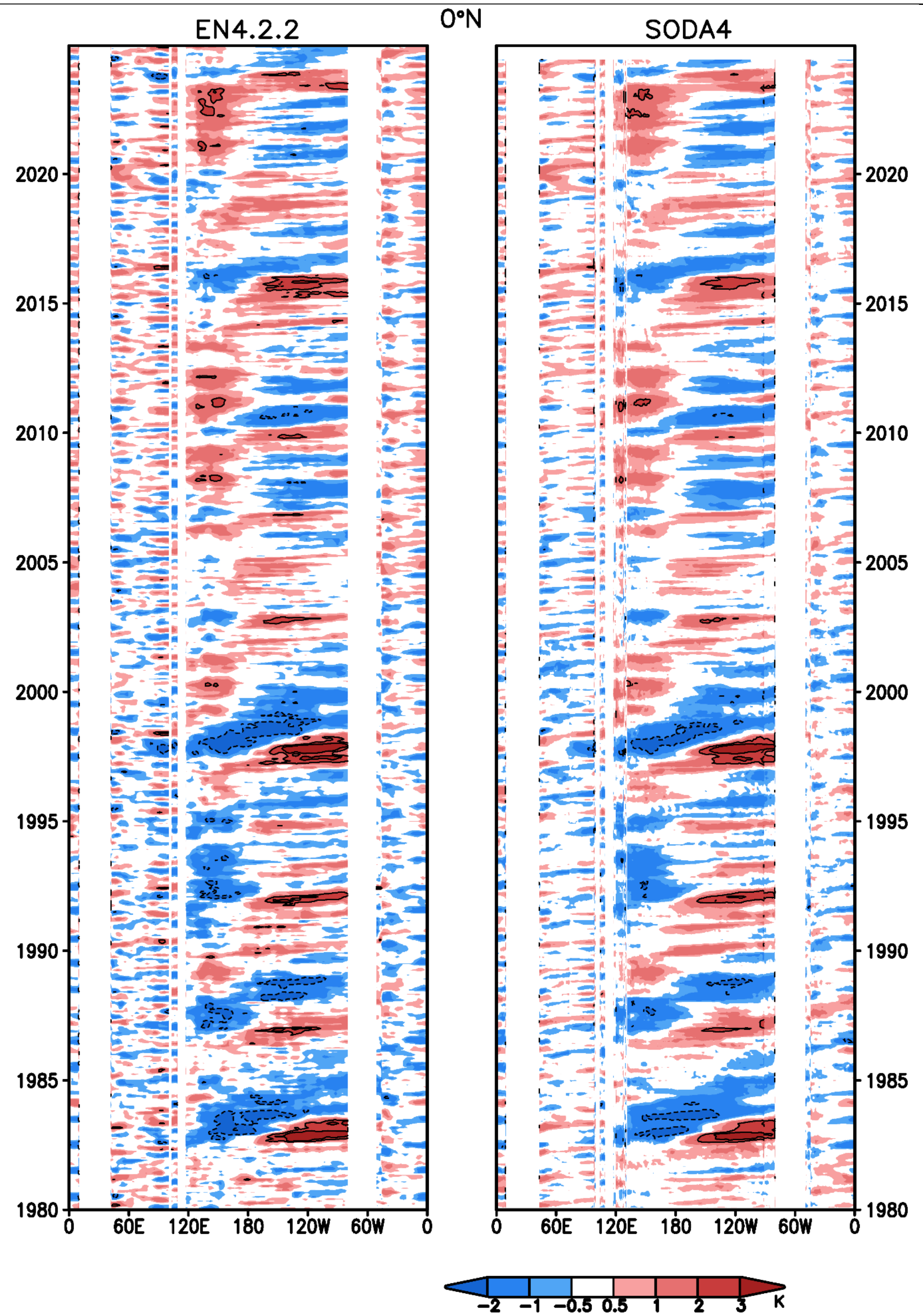


Fig. S14 Hovemoller diagram of monthly 0-300m average temperature anomaly from monthly climatology with time and longitude along 0°N. (left) EN4.2.2 (right) SODA4.

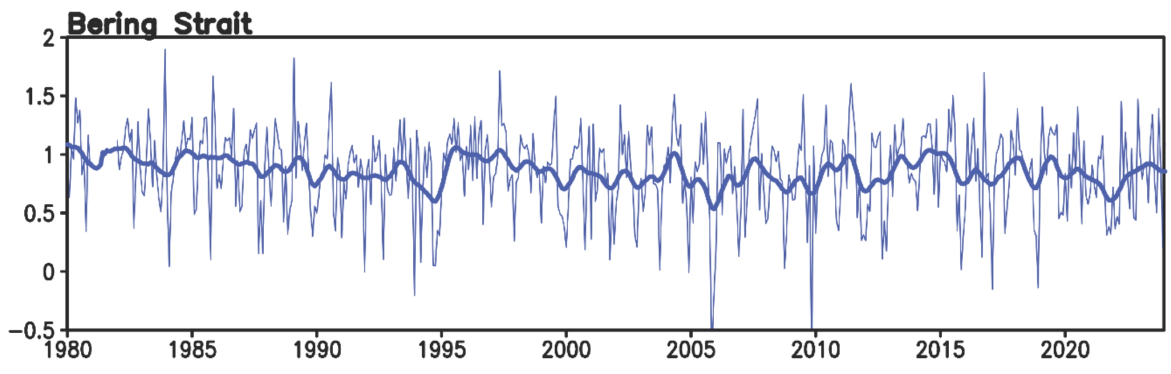


Fig. S15 Total volume transport ($10^6 m^3/s$) through Bering Strait (65.9°N) with time. (thin) 10dy average (solid) annual average. Time mean transport (1980-2024) is 0.86±0.13 $10^6 m^3/s$.

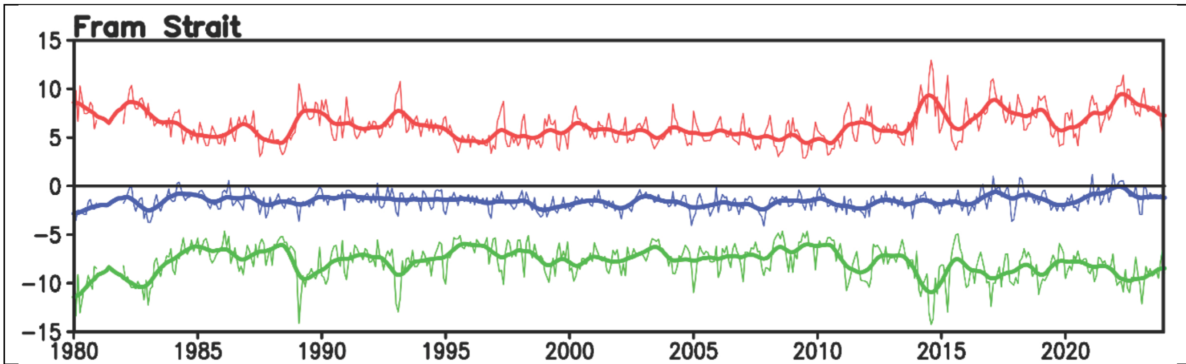


Fig. S16 (red) Northward, (green) southward, and (blue) total volume transports ($10^6 m^3/s$) through Fram Strait (78.0°N) with time. (thin) 10dy average (solid) annual average. Time mean northward and southward transports (1980-2024) are: 6.3 $10^6 m^3/s$ and -7.8$10^6 m^3/s$.

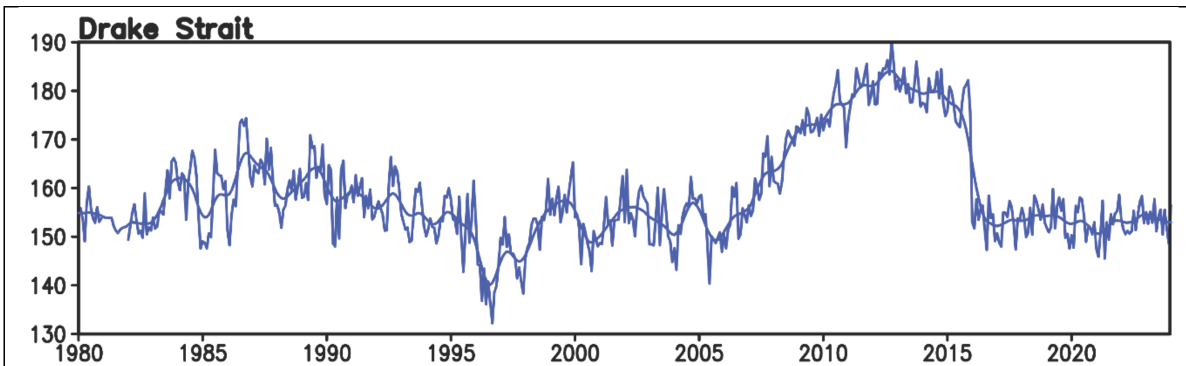


Fig. S17 Total volume transport ($10^6 m^3/s$) through Drake Strait (62.4°W) with time. (thin) 10dy average (solid) annual average. Time mean transport (1980-2024) is 159±9.2 $10^6 m^3/s$. The increase in transport in 2005 coincides with a switch to the upwind advective scheme (see **Fig. S8**).

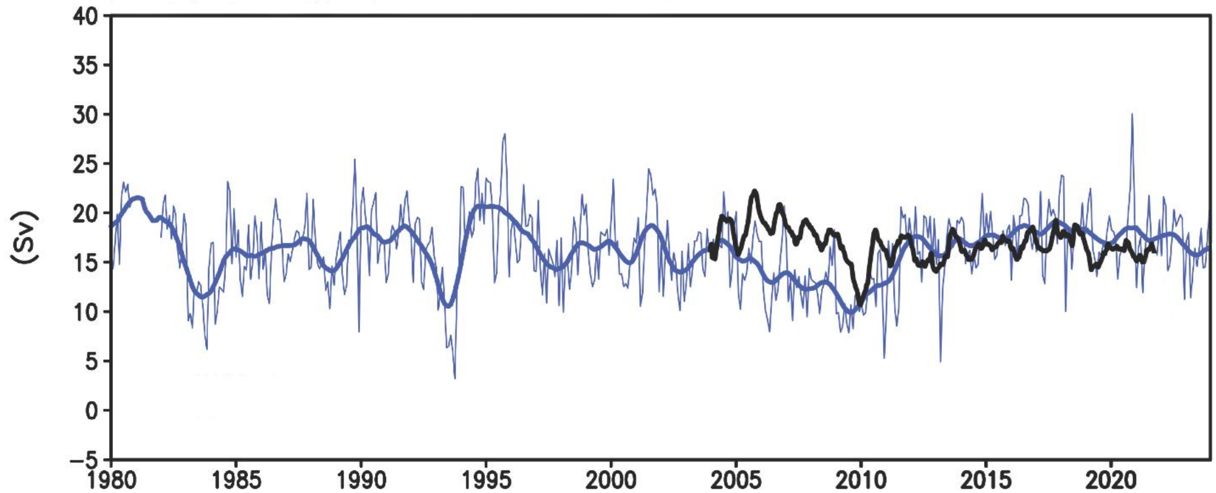


Fig. S18 Atlantic meridional overturning circulation across 26.5°N (10^6 m^3/s). (blue) SODA4, (black) RAPID time series. The time mean SODA4 overturning transport is: 16.5±0.37 x10^6 m^3/s.