Supporting Information for "Decadal variations and trends of the global ocean carbon sink"

Peter Landschützer,[1,2] Nicolas Gruber,[1] and Dorothee C.E. Bakker,[3]

Contents of this file

1. Text S1 to S5

P. Landschützer, Max Planck Institute for Meteorology, 20146 Hamburg, Germany. (Peter.Landschuetzer@mpimet.mpg.de)

N. Gruber, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland. (nicolas.gruber@env.ethz.ch)

D.C.E. Bakker, Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK. (d.bakker@uea.ac.uk)

[1]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland

[2]Max Planck Institute for Meteorology, Hamburg, Germany

[3]Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

2. Figures S1 to S9

S1. Residual analysis

As shown in previous studies the SOM-FFN product is capable of reconstructing the gridded SOCATv2 observations with a negligible small global mean bias of 0.16 μatm and a root mean squared error (RMSE) of about 15 μatm for the entire time period from 1982 through 2011. This conclusion applies also to each 5-year sub period, during which the data availability differs hugely (Figure S1). We find for each period and each ocean basin separately, that the bias remains small ranging from a minimum of roughly -2 μatm in the Indian Ocean throughout our analysis period to a maximum of about 4 μatm in the Atlantic Ocean from 1997 through 2001, i.e., a period with fewer observations. In general though, we find only little indication that the bias magnitude depends on the amount of available data. More importantly though, there is no systematic trend in the bias, which gives us confidence in our trend estimates. In contrast, the RMSE actually increases with increasing number of observations, which we interpret to be caused by the progressively larger inclusion of data stemming from highly variable regions, such as coastal margins and the high latitudes.

The SOCAT-based residual assessment is confirmed when the SOM-FFN pCO_2 estimates are compared with independent observations from the LDEOv2013 database, i.e., by comparing our estimates to data from this database that are not included in SOCATv2 (Figure S2). Globally, the bias for each 5-year period stays close to zero, with the exception of the first period when data are very sparse. But this comparison also shows more substantial sub-basin differences, with 5-year average biases exceeding $\pm 10\mu$atm in the Pacific Ocean from 1982 through 1986 and similar biases in the Indian Ocean from 2002 through 2006. We attribute these periods of

large differences to low data availability. Encouragingly, there is still no significant trend in the global $p\mathrm{CO}_2$ bias or any sub-basin bias, giving us confidence in our findings.

We further check for potential spatial biases in the residuals by plotting the temporal mean bias for the periods 1982-1991, 1992-2011 and 2002-2011, again for the SOCATv2 dataset (Figure S3) and for the data from the LDEOv2013 dataset that were not used to create our product (Figure S4). We find that the largest neural network-observation mismatch occurs in the high latitudes and equatorial Pacific, i.e. the high variability regions of the global ocean, confirming our previous results. However, there is no indication of a systematic bias in the spatial domain, i.e., no large ocean region is either continuously biased low or high.

Overall we conclude that our method is capable of reproducing the available observations within a small error margin, based on the analysis of dependent as well as independent observations.

S2. Comparison with timeseries stations

As a second test to assess the ability of SOM-FFN to reconstruct decadal $p\mathrm{CO}_2$ variability and trends, we compare our neural network derived surface ocean $p\mathrm{CO}_2$ estimates with the observed $p\mathrm{CO}_2$ from two of the longest running timeseries stations, i.e., with data that are not included in SOCATv2. We therefore extracted the combined record of the BATS - Bermuda Atlantic Timeseries Station and Bermuda Hydrostation S located 31.66°N, 64.16°W in the Atlantic Ocean (Figure S5) and the HOT- Hawaiian Ocean Timeseries station ALOHA located 22.75°N, 158.00°W in the Pacific Ocean (Figure S6). Both timeseries stations offer more than 20 years of roughly monthly measured surface ocean carbon system parameters. For the analysis, we compare the timeseries station data with the average of the four closest SOM-FFN pixels around the timeseries location, weighted by their distance. The length of both timeseries allows us to in-

vestigate several variability signals individually, namely the seasonal cycle, decadal trends and interannual to decadal variability.

The comparison with the Bermuda record reveals that the SOM-FFN estimate captures the observed $p\mathrm{CO}_2$ variations rather well, despite the fact that the SOM-FFN method was trained on a rather limited data set in this region, i.e., covering mostly just the post 2006 period (see blue dots) (Figure S5a). Also the mean seasonal cycle is well represented, even though there is a tendency for the seasonal amplitude to be underestimated (Figure S5a). We suspect that this is an consequence of our method relying on data covering an entire biogeochemical region in order to create its final product, thus there is likely a tendency to underestimate the true variability. Also the SOM-FFN based long-term $p\mathrm{CO}_2$ trend, calculated by removing the mean seasonal cycle from the $p\mathrm{CO}_2$ timeseries, appears to be somewhat underestimated relative to the observations (Figure S5c). This is particularly visible in the first decade, where the least amount of data were available to train the neural network. Given that a key result of our global SOM-FFN analysis is the stalling/reduction of the global carbon sink during the 1980s and 1990, the fact that the SOM-FFN method underestimates the $p\mathrm{CO}_2$ increase in the first decades implies that if anything, our method underestimates the carbon sink reduction over this period, at least in the biogeochemical region of the subtropical North Atlantic. A similar tendendy for underestimation of the variability is also seen on interannual timescales, where the $p\mathrm{CO}_2$ variability from the SOM-FFN method calculated by removing the slope of the linear trend shows much smaller amplitude than the observations. This is particularly the case for the first 20 years (see Figure S5d). In addition, the timing of the reconstructed interannual variability signal does not agree well with the observations, but the difference is not significant given the

substantial amount of high frequency noise in the observations, likely stemming from eddies and local weather events, which are not considered in the SOM-FFN method.

Figure S6a-d shows the same analysis for the HOT station ALOHA. Similar to BATS, the neural network based seasonal cycle underestimates the seasonal cycle at HOT, whereas we find better agreement regarding the timing of minimum and maximum pCO$_2$. Again, the HOT data reveal a steeper pCO$_2$ trend indicating that the neural network underestimates the carbon sink reduction in the first 20 years of the analysis period in the subtropics of the northern hemisphere. Figure S6d further suggests a larger, ENSO related, variability signal around the year 2000, indicating that the neural network based pCO$_2$ estimates might locally underestimate natural climate variability signals.

Overall, the evaluation of our SOM-FFN estimates with independent data from the subtropical timeseries stations reveals a remarkably good agreement, but with the SOM-FFN estimates having the tendency to understimate the surface ocean pCO$_2$ variability in the subtropics. This further highlights the importance of long running and continuous timeseries data. .

S3. Drivers of air-sea CO$_2$ flux variability

In order to determine the relative contribution of the two major terms in equation (1) to the reconstructed air-sea CO$_2$ flux variability and trends, we re-calculated the global and basin-wide integrated carbon fluxes assuming either constant ΔpCO$_2$ $= pCO_2 - pCO_2^{atm}$ or constant gas transfer/sea-ice fraction ($k_w \cdot S \cdot (1 - f_{ice})$ - global average) over time. In both cases, we used the global mean values over the entire period as the constant value. Figure S7 reveals that the identified trends almost entirely stem from the variability in ΔpCO$_2$ and only to a lesser degree from changes in wind and the gas transfer, which is in agreement with previous studies. With longer term variations in ΔpCO$_2$ almost entirely being driven by sea surface pCO$_2$ (with the

important exception of the long-term trend), this suggests that most of the air-sea flux variations must be caused by oceanic processes affecting temperature, DIC, and Alk.

S4. EOF analyses of the air-sea flux of CO_2

We further tested for potential influences of the gas transfer coefficient on our Δ EOF analysis by repeating the global EOF analysis for the detrended air-sea flux density fields. Figure S8 illustrates the spatial amplitude of the first and second EOF. The pattern of Figure S8 are almost identical to Figure 4b-c in the main text, confirming our previous finding that the air-sea CO_2 flux variations are primarily a consequence of changes in oceanic pCO_2. Somewhat larger differences emerge for EOF2, where the dominance of the tropical Pacific signal in the EOF2 of ΔpCO_2 is muted in EOF2 of the air-sea flux. In particular, a stronger EOF2 related signal emerges in the Southern Ocean, owing to the substantial differences in the mean windspeed between the two regions.

S5. Sensitivity of the integrated flux to data availability and drivers

As a first point of reference, we compare the annual mean integrated flux from this study to the results published in, who reported and discussed the globally integrated ocean carbon sink from 1998 through 2011 (Figure S9). Both reconstructions differ little and agree within their respective uncertainties, even though there were considerably differences in the way the driver data were used (see main text for details). In particular, both reconstructions suggest a rapid strengthening of the ocean carbon sink within the most recent decade.

Next, we tested whether the substantial amount of additional data from the SOCATv3 database impact the decadal variability signal. This is particularly relevant as this newest version adds much new data in the latest part of the record and also closes some earlier gaps (see Figure S9b for the change in temporal coverage). It turns out that the SOCATv3 based recon-

structions agree very well with those based on SOCATv2, with the exception, perhaps, of the years 2010 and 2011, when the SOCATv3-based reconstruction suggest a somewhat smaller ocean carbon sink. We interpret this to be a consequence of the data availability post 2011 in the SOCATv3 case, which helps to constrain the trend in the last part of the record, i.e., a period that the SOM-FFN method is sensitive to (see also below).

In order to further test the sensitivity of the method to data availability, we run a number of validation cases, where data only from alternating 5-year intervals are used for training. Thus in run "Validation5-9", only data from 1985 though 1989, 1995 through 1999, and 2005 though 2009 are used for training. Analogously, in run "Validation0-4" only data from 1980 through 1984, 1990 through 1994, 2000 through 2004, and 2010 through 2014 are selected. These 5-year periods are highlighted in Figure S9 as shaded areas in the background. For this sensitivity runs, we employed data from SOCATv3.

These sensitivity experiments reveals no significant effect to this strong data thinning. An exception is run "Validation5-9" in the early 1980s, where the reconstructed flux is outside the uncertainty interval of the results presented in this manuscript. As discussed in the main text, this is a result of the SOM-FFN method having problems to extrapolate in time, owing to its having to rely on atmospheric CO_2 to reconstruct a trend. Thus, if no training data exist for an entire period, the influence function associated with atmospheric CO_2 is ill constrained. If this period happens to be either at the beginning or the end, this lack of constraints can lead to rather wrong trends in surface ocean pCO_2, i.e., wrong trends in the air-sea CO_2 flux. This problem is worse in the 1980s, when the data constraints are overall weak anyway. Despite this sensitivity to data availability in the 1980s, the reconstructed reduction of the sink in the 1990s

and the following strengthening of the carbon sink are robust against the data density from 1990 onward.

Finally, we tested if our results are sensitive to our switching from fully varying driver data for the post 1998 period to only climatologically varying driver data for the first period (1982 to 1997) (the affected driver data are only chlorophyll and mixed layer depth). To this end, we ran a sensitivity case, where we reconstructed $p\mathrm{CO}_2$ without chlorophyll and mixed layer depth as drivers, i.e., with only SST, sea-surface salinity and atmospheric CO_2 as drivers. We find that while the observational mismatch suffers from removing the climatologies, the integrated flux does not change significantly (not shown). This is arguably an extreme sensitivity case, but it bounds the magnitude of the potential impact of our using climatologies for chlorophyll and mixed layer depth rather than their actual timeseries.

In conclusion, these analyses and tests demonstrate that the SOM-FFN based $p\mathrm{CO}_2$ estimates are quite robust against changes in data density and driver availability, giving us good confidence that the reconstructed variations are not caused by an artefact of the method or of the data.

Figure S1. Residuals between the SOM-FFN pCO_2 estimates and the SOCATv2 gridded observations for the global ocean and the 4 major ocean basins (seperated at 44°S). Panel (a) shows the bias and (b) the root-mean-squared error (RMSE) of the residuals divided into 5-year periods. The number of gridded observations is illustrated in panel (c)

Figure S2. As Figure S1 but for residuals calculated between the neural network output and LDEOv2013 gridded observations that are not included in SOCATv2. Note the difference in the y-axis between Figures S1 and S2.

Figure S3. Spatial distribution of the residuals between the neural network estimates and the SOCATv2 gridded observations for 3 different timeperiods, namely 1982 through 1991, 1992 through 2001 and 2002 through 2011.

Figure S4. As Figure S3 but for residuals calculated between the neural network output and LDEOv2013 gridded observations that are not included in SOCATv2.

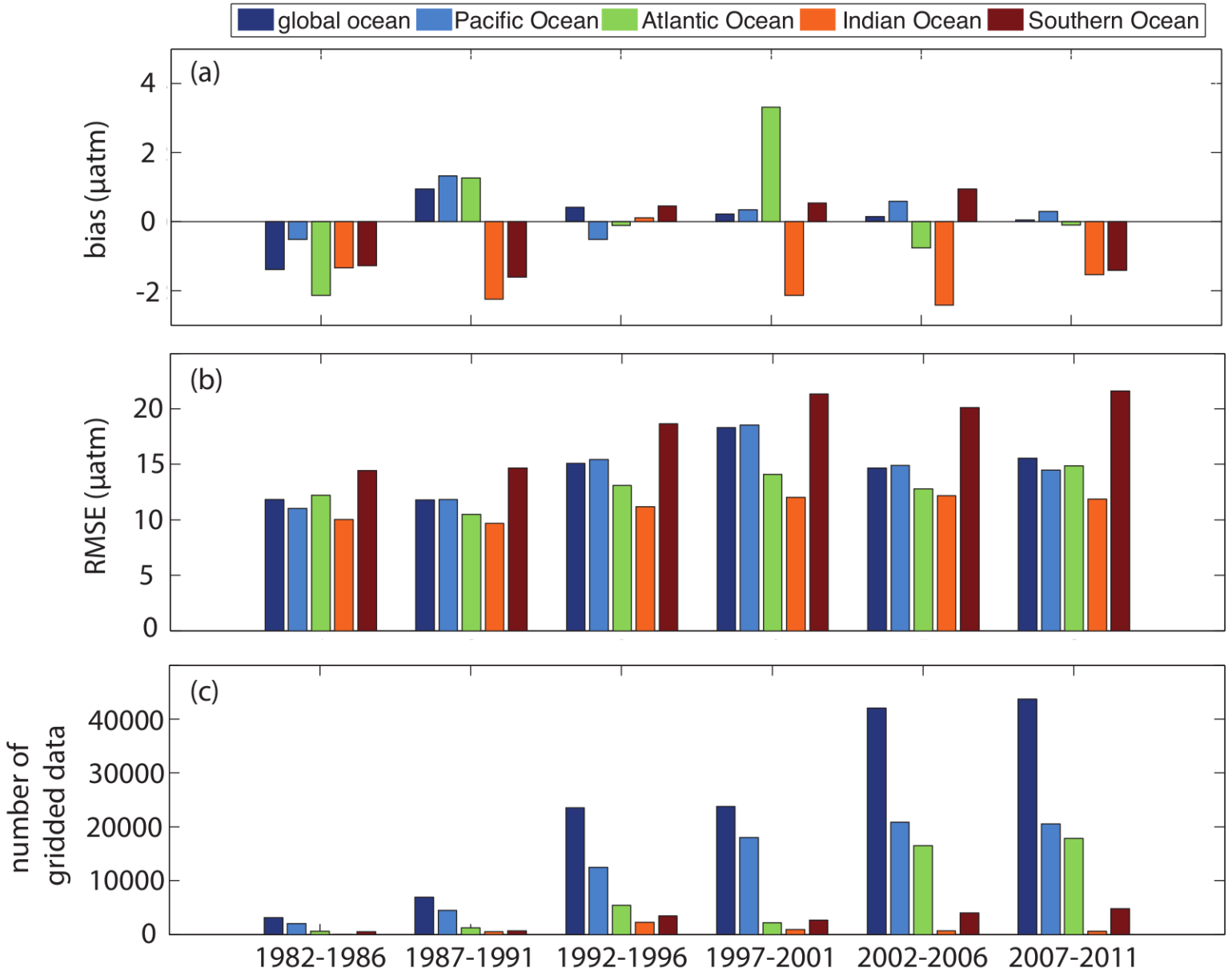
Figure S5. Comparison between the neural network estimates (black) and the combined Bermuda Atlantic Timeseries Station and Hydrostation S records (red). (a) shows the monthly timeseries, (b) the mean seasonal cycle, (c) the linear trend and (d) the interannual variability.

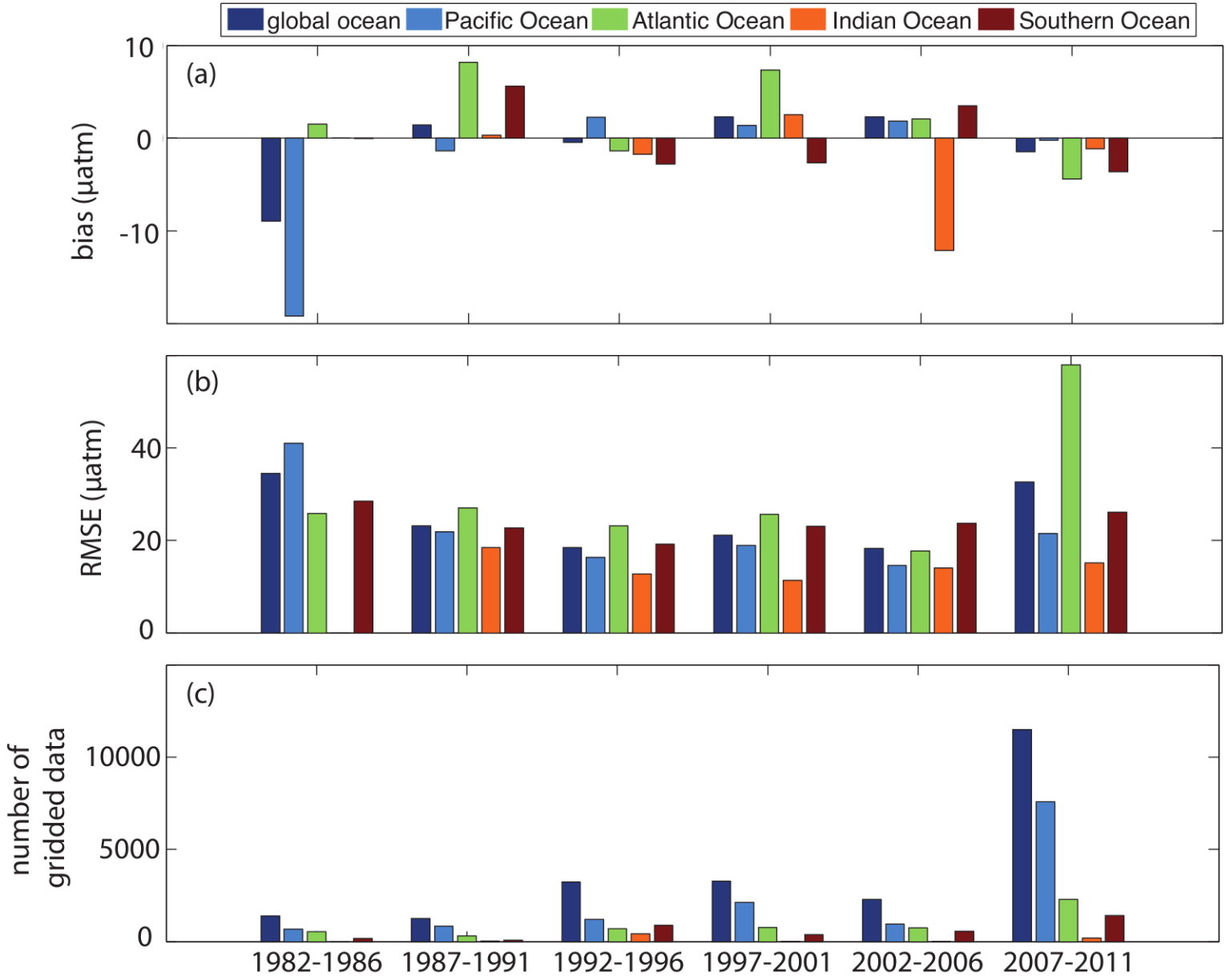
Figure S6. As Figure S5 but for the comparison between the neural network estimates (black) and the Hawaiian Ocean Timeseries Station (red)

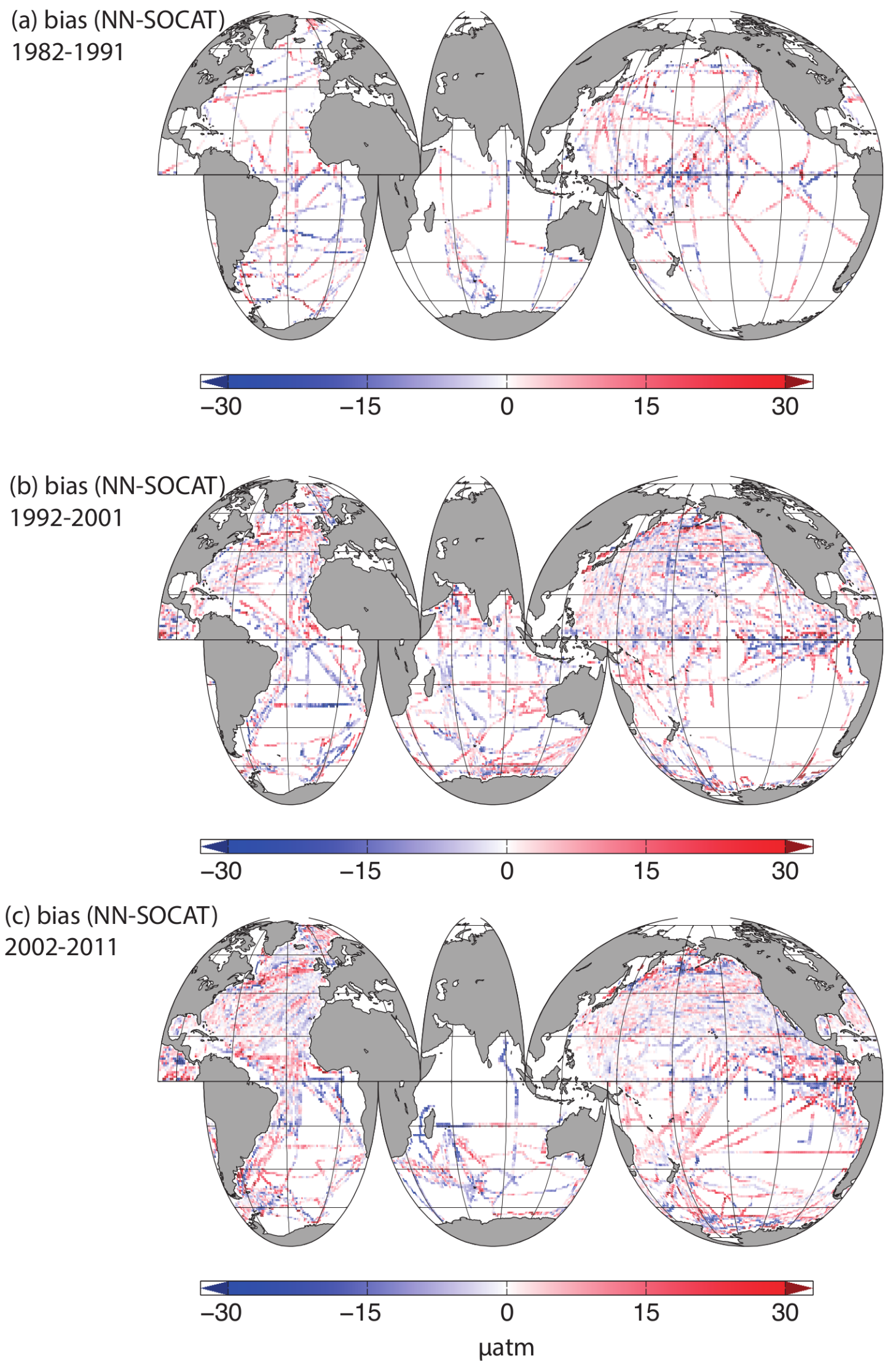
Figure S7. Annual mean integrated air-sea CO_2 flux from 1982 through 2011 based on the ETH30yr output and the contributions of each basin (separated at 44°S) for case (a) with constant ΔpCO_2 and case (b) with constant gas exchange rate for comparison to Figure 1 in the main text.

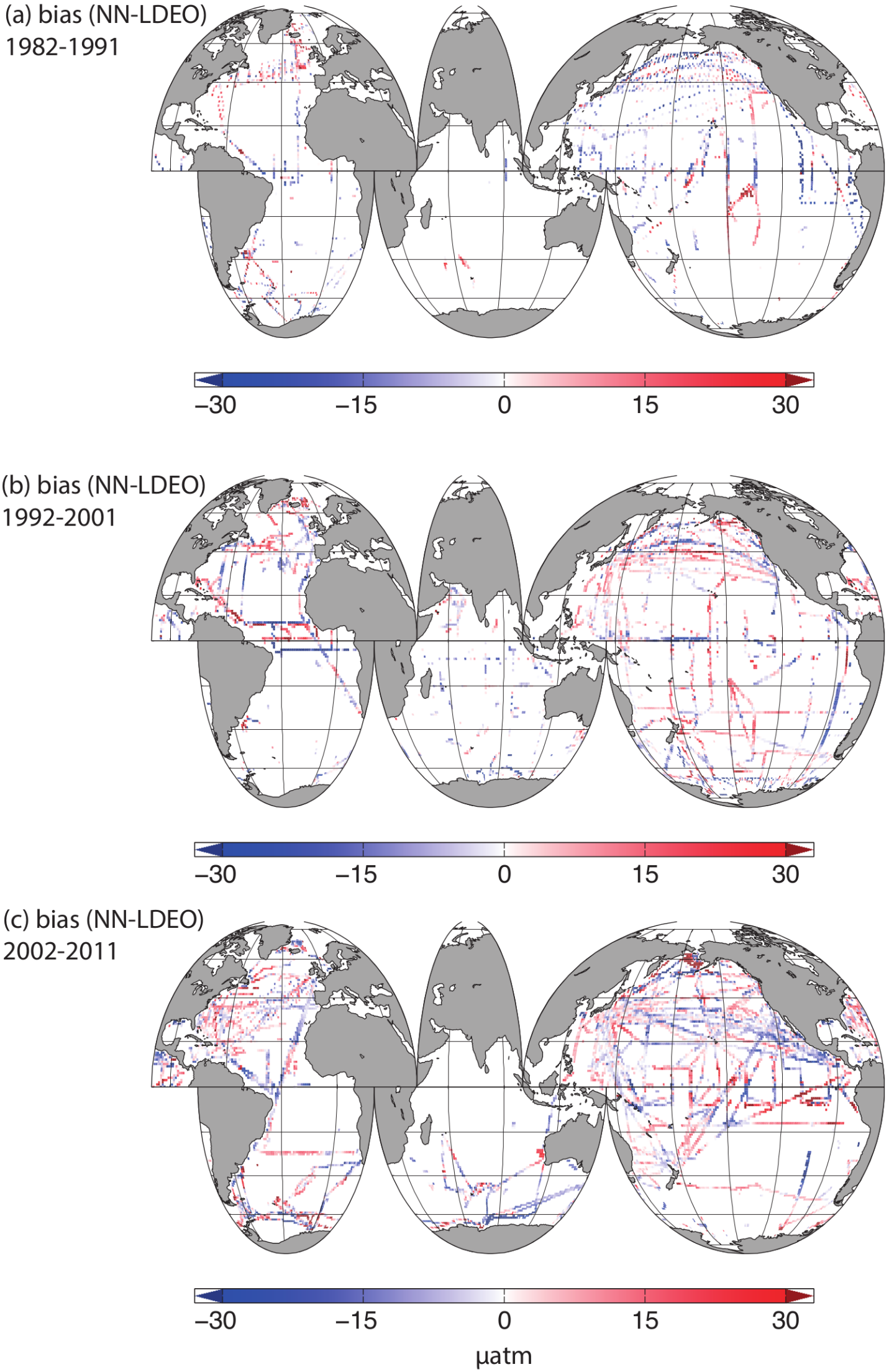
Figure S8. Map of the spatial amplitude of the deseasonalized first and second global EOF of the air-sea flux density.

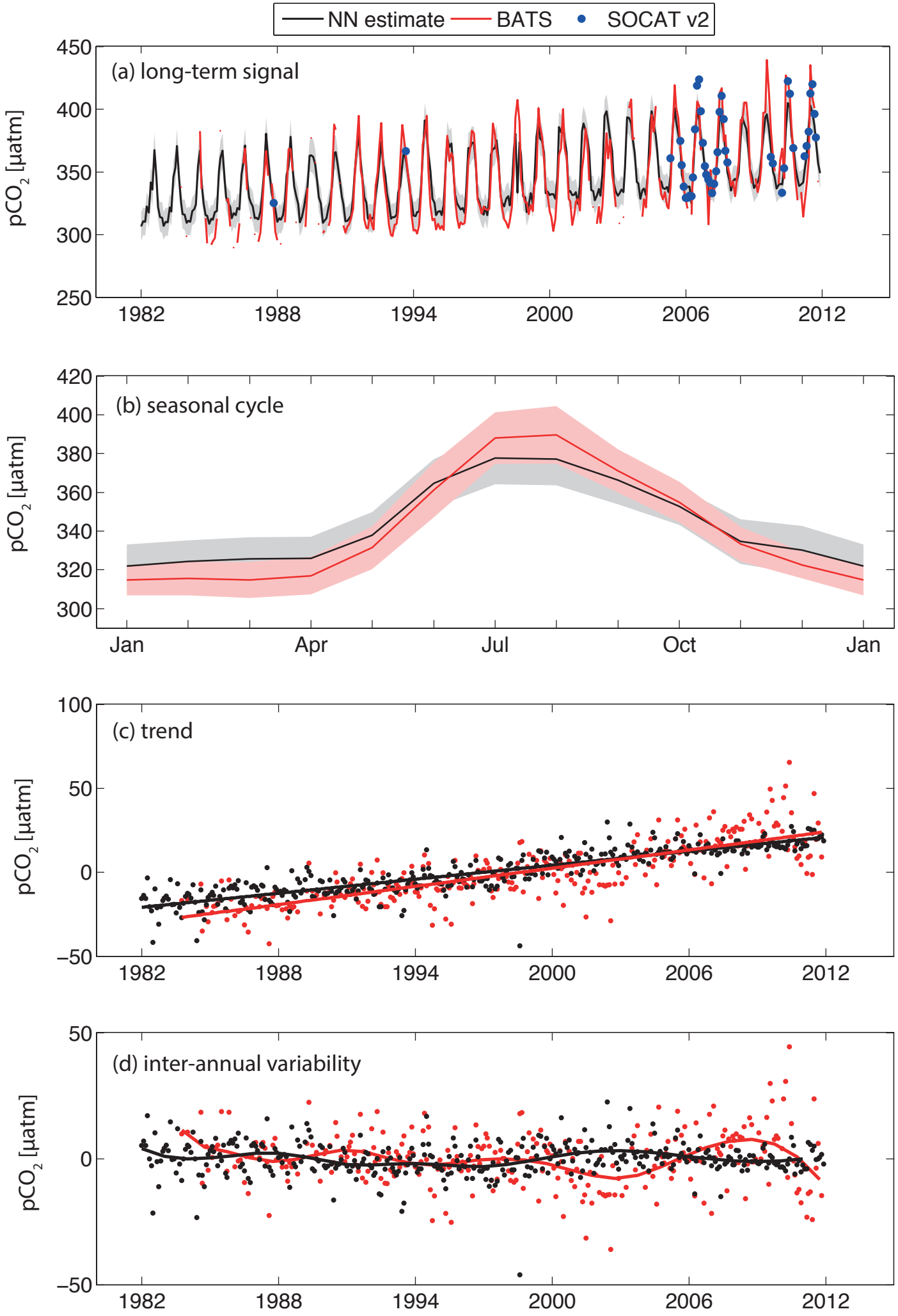
Figure S9. (a) SOM-FFN method sensitivity runs in comparison. The black line with uncertainty shading illustrates the annual mean integrated carbon sink from this study. The green line shows the same but with additional observations from the SOCATv3 database. The red line illustrates previously published results. The light blue line illustrates the results without the use of climatological data. The dark blue and puple lines represent runs where data were only selected at alternating 5 year periods (highlighted by the shaded background in the same color). (b) The number of gridded observations within the SOCATv2 and SOCATv3 datasets.

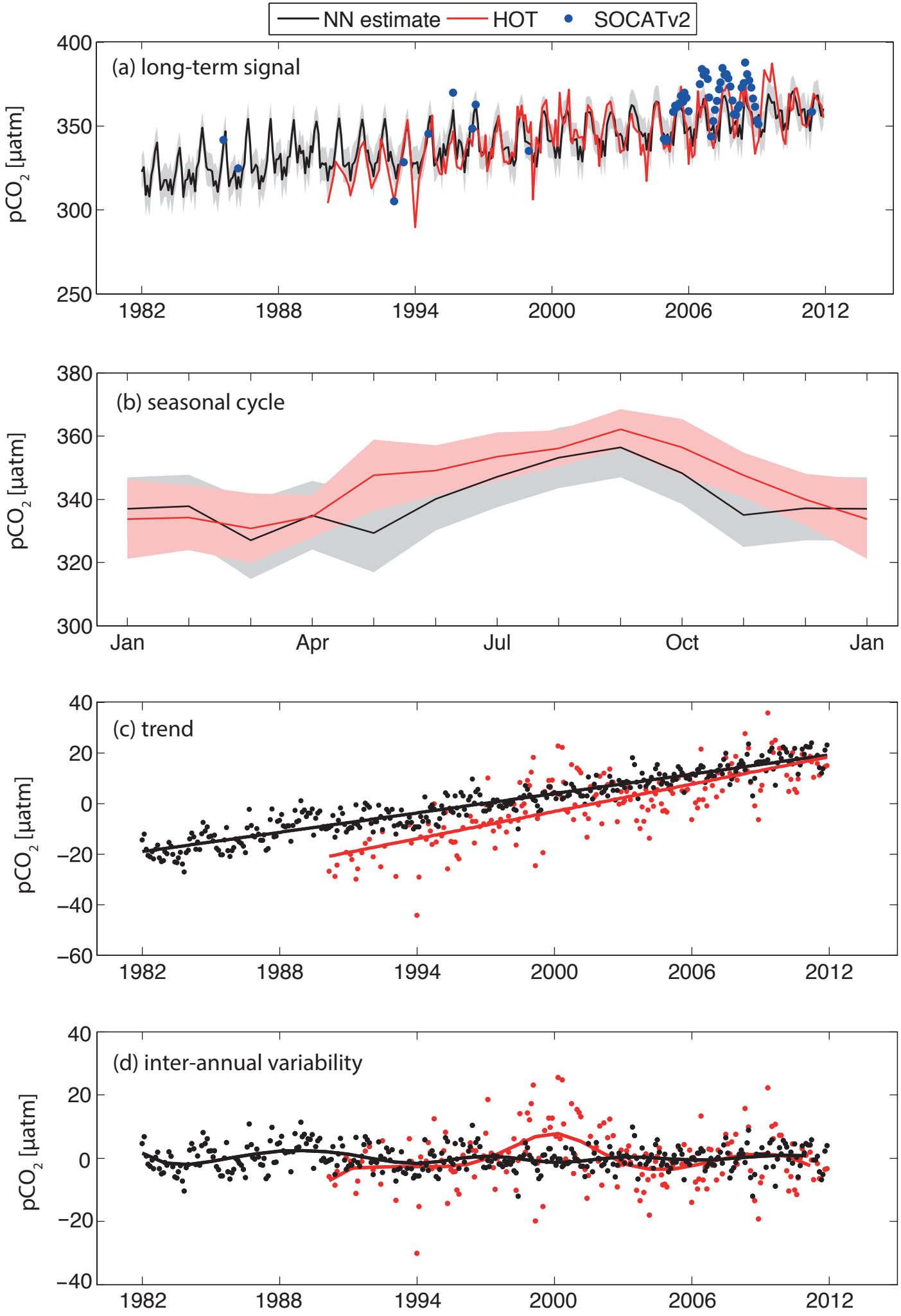

global ocean
Pacific Ocean
Atlantic Ocean
Indian Ocean
Southern Ocean
(a)
bias (μatm)
4
2
0
-2
(b)
RMSE (μatm)
20
15
10
5
0
(c)
number of gridded data
40000
30000
20000
10000
0
1982-1986
1987-1991
1992-1996
1997-2001
2002-2006
2007-2011

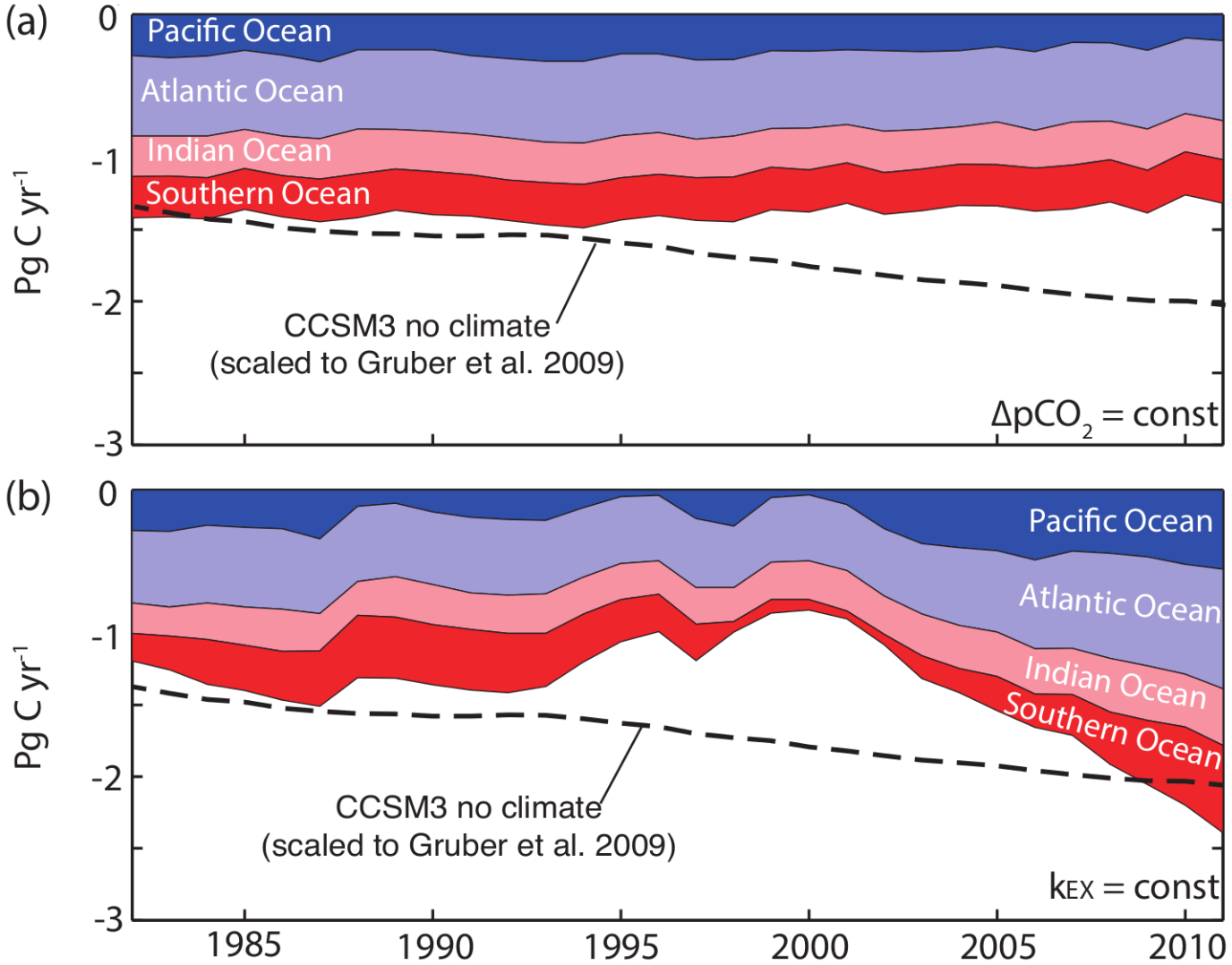

global ocean
Pacific Ocean
Atlantic Ocean
Indian Ocean
Southern Ocean
(a)
bias (µatm)
10
0
-10
(b)
RMSE (µatm)
40
20
0
(c)
number of gridded data
10000
5000
0
1982-1986
1987-1991
1992-1996
1997-2001
2002-2006
2007-2011


(a) bias (NN-SOCAT)
1982-1991
−30
−15
0
15
30
(b) bias (NN-SOCAT)
1992-2001
−30
−15
0
15
30
(c) bias (NN-SOCAT)
2002-2011
−30
−15
0
15
30
µatm


(a) bias (NN-LDEO)
1982-1991
−30
−15
0
15
30
(b) bias (NN-LDEO)
1992-2001
−30
−15
0
15
30
(c) bias (NN-LDEO)
2002-2011
−30
−15
0
15
30
µatm

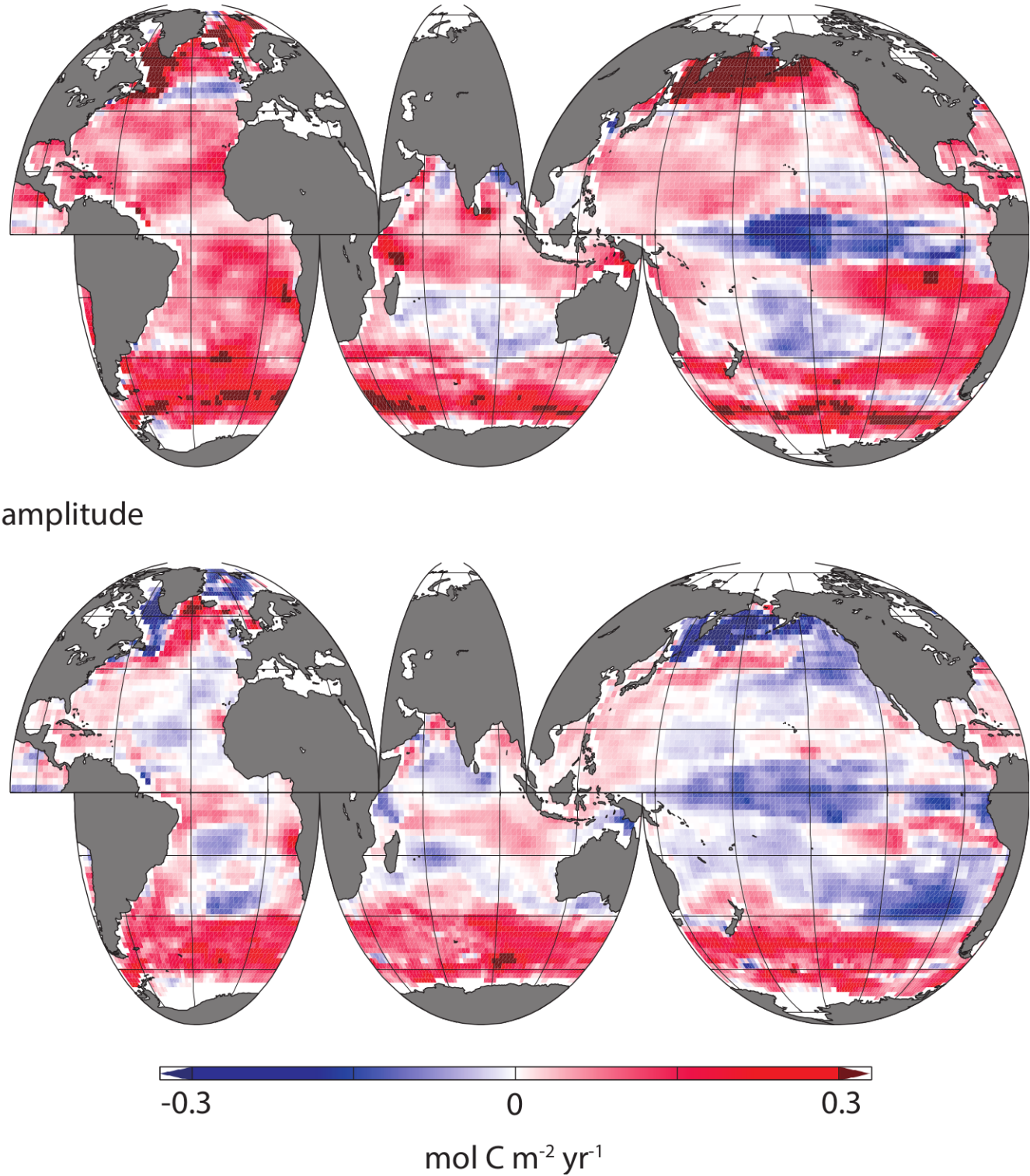

NN estimate
BATS
SOCAT v2
(a) long-term signal
pCO2 [µatm]
450
400
350
300
250
1982
1988
1994
2000
2006
2012
(b) seasonal cycle
pCO2 [µatm]
420
400
380
360
340
320
300
Jan
Apr
Jul
Oct
Jan
(c) trend
pCO2 [µatm]
100
50
0
−50
1982
1988
1994
2000
2006
2012
(d) inter-annual variability
pCO2 [µatm]
50
0
−50
1982
1988
1994
2000
2006
2012

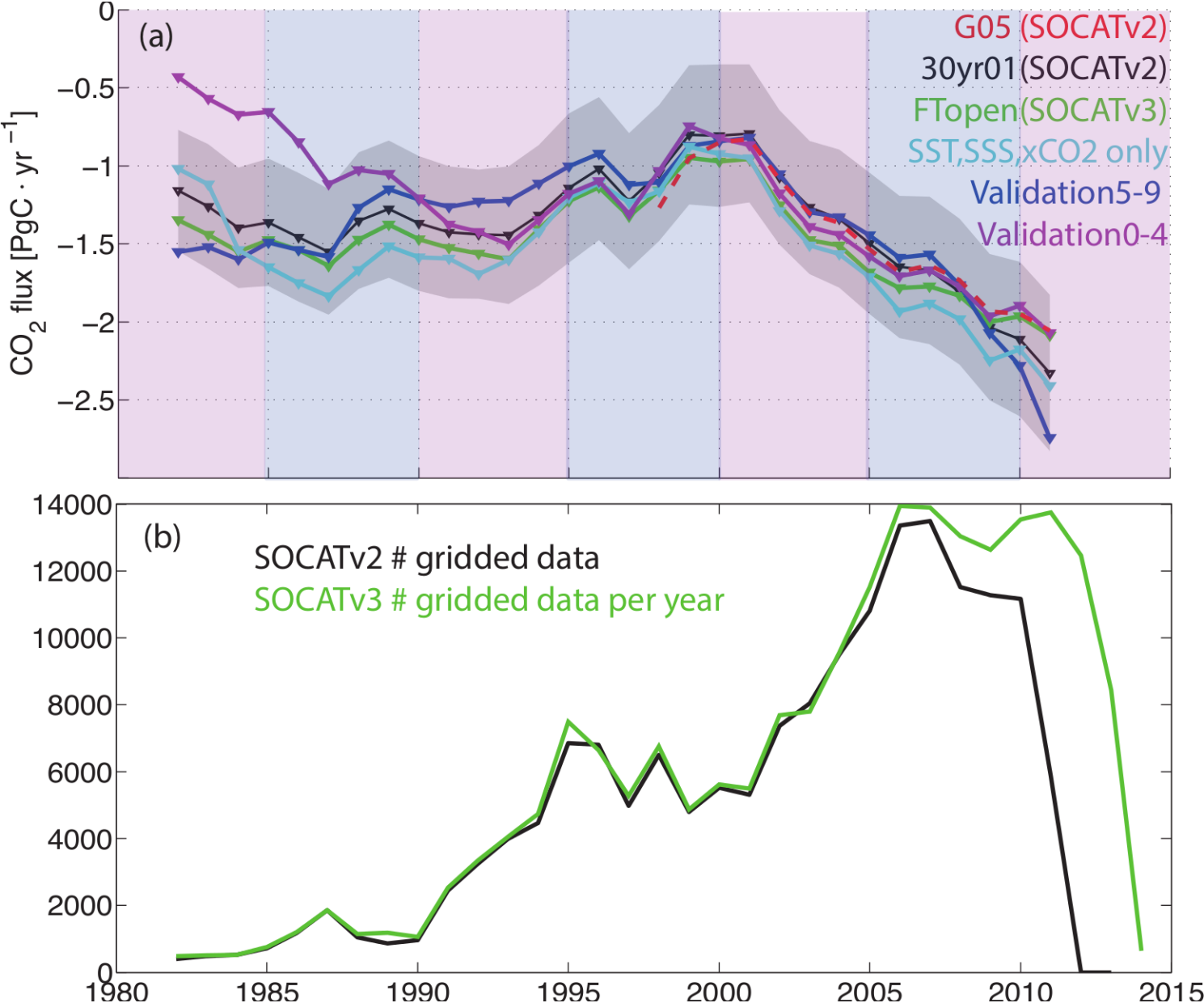

NN estimate
HOT
SOCATv2
(a) long-term signal
pCO2 [µatm]
400
350
300
250
1982
1988
1994
2000
2006
2012
(b) seasonal cycle
380
360
340
320
300
Jan
Apr
Jul
Oct
Jan
(c) trend
40
20
0
−20
−40
−60
(d) inter-annual variability


(a)
0
-1
-2
-3
Pg C yr-1
Pacific Ocean
Atlantic Ocean
Indian Ocean
Southern Ocean
CCSM3 no climate
(scaled to Gruber et al. 2009)
ΔpCO2 = const
(b)
0
-1
-2
-3
Pg C yr-1
Pacific Ocean
Atlantic Ocean
Indian Ocean
Southern Ocean
CCSM3 no climate
(scaled to Gruber et al. 2009)
kEX = const
1985
1990
1995
2000
2005
2010

(a) EOF1 amplitude

(b) EOF2 amplitude




(a)
G05 (SOCATv2)
30yr01(SOCATv2)
FTopen(SOCATv3)
SST,SSS,xCO2 only
Validation5-9
Validation0-4
CO_2 flux [PgC · yr^{-1}]
0
−0.5
−1
−1.5
−2
−2.5
(b)
SOCATv2 # gridded data
SOCATv3 # gridded data per year
14000
12000
10000
8000
6000
4000
2000
0
1980
1985
1990
1995
2000
2005
2010
2015