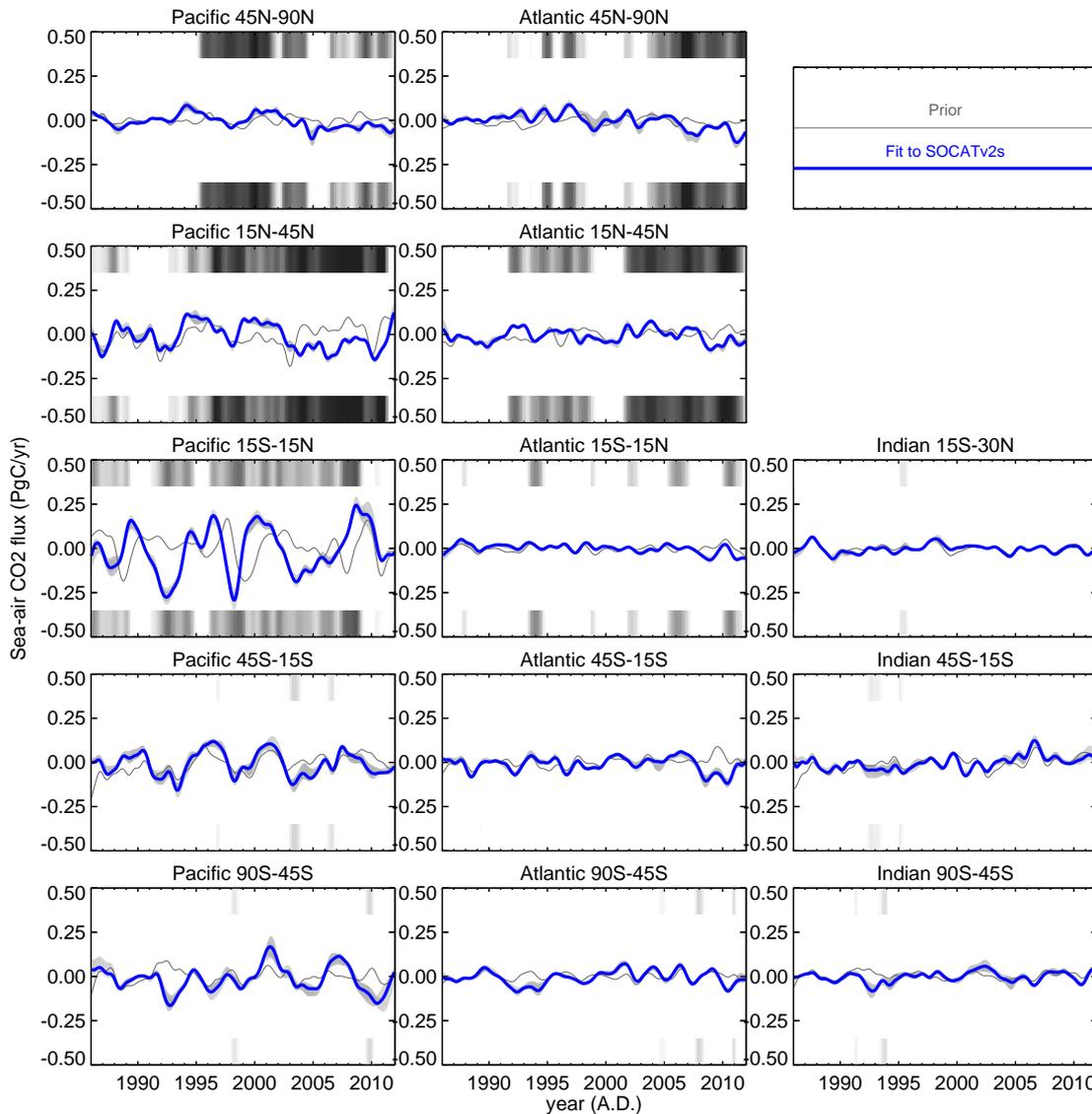


## Supplementary material

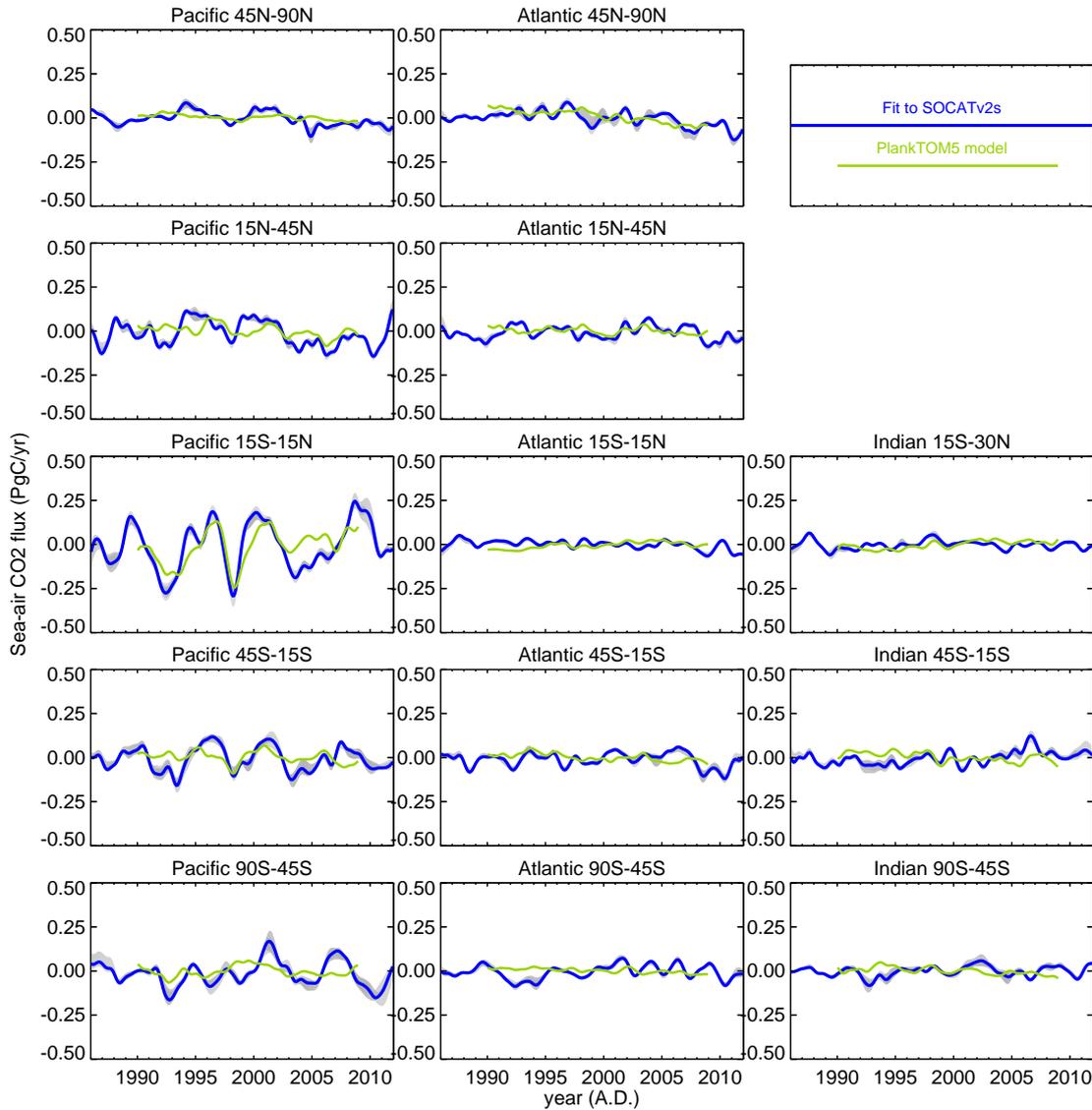
# Interannual sea–air CO<sub>2</sub> flux variability from an observation-driven ocean mixed-layer scheme

### S1 Regional disaggregation of results

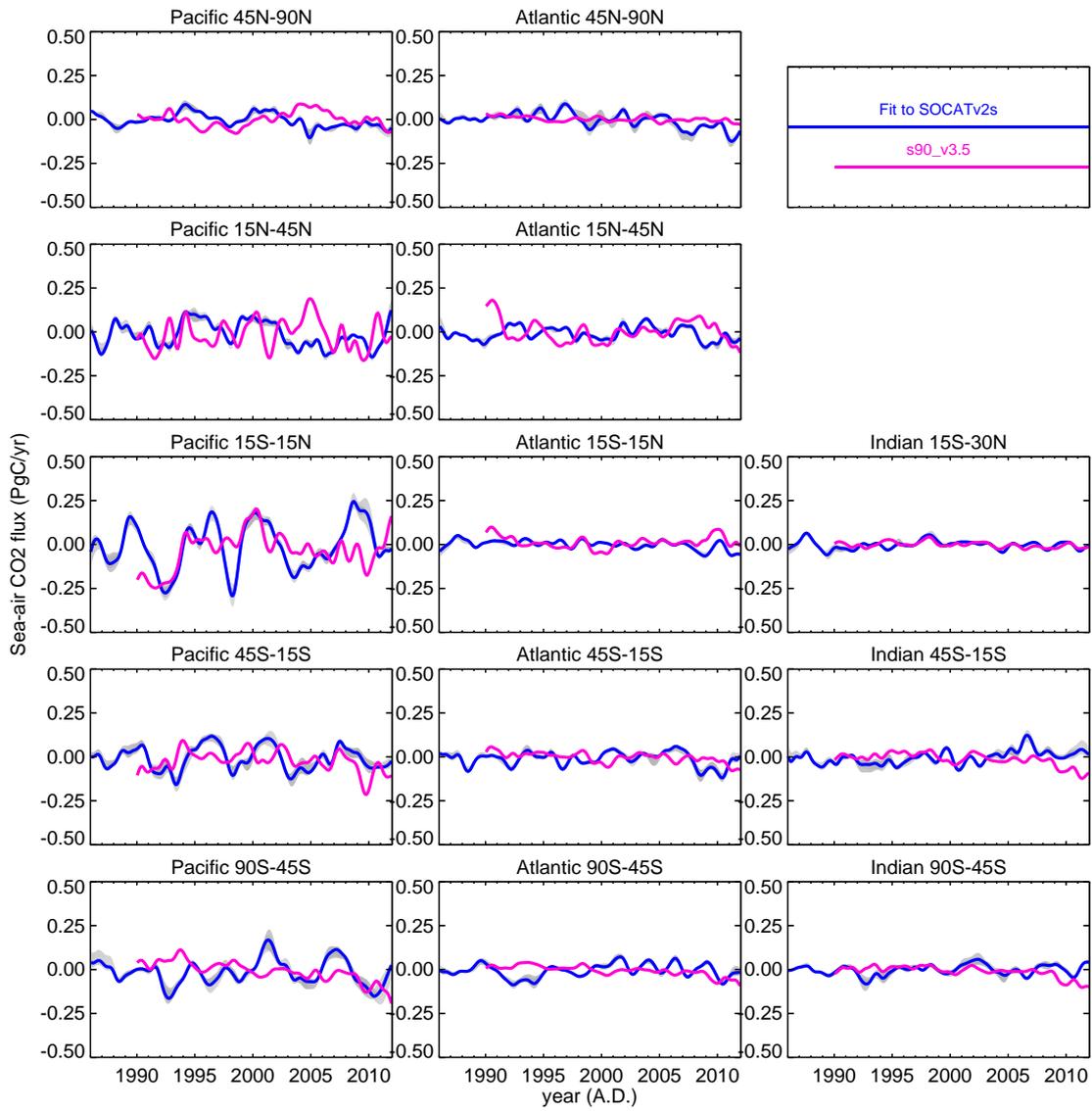
The SOCAT-based sea–air CO<sub>2</sub> flux estimates are shown in finer spatial disaggregation (Fig. S1). The comparison to the ocean model and to the atmospheric inversion is shown for the same regions (Fig. S2 and S3).



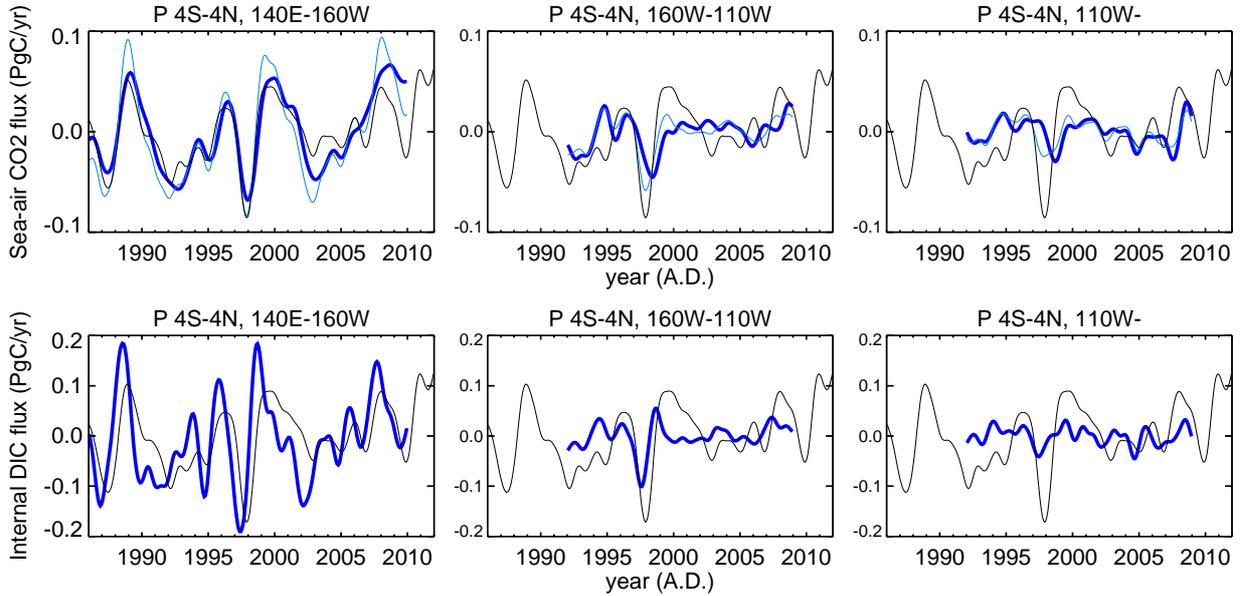
**Fig. S1.** Interannual anomalies of the sea–air CO<sub>2</sub> exchange estimated from SOCAT data (blue), integrated over 5 latitude bands split into ocean basins (including open ocean, coastal areas and marginal seas; panels in geographical arrangement). Time series have been deseasonalized and filtered for interannual variations, and the 1993–2008 mean has been subtracted (positive values indicate an increased ocean CO<sub>2</sub> source, negative values an increased CO<sub>2</sub> sink). The light-grey band around the standard case comprises a set of sensitivity cases where the a-priori uncertainty of all degrees of freedom, the a-priori uncertainty of the interannual degrees of freedom, or the a-priori spatial correlation lengths have been changed. The dark grey line gives the a-priori state of the diagnostic scheme responding to changes in temperature and other driving variables. The shading at the top and bottom of each panel illustrates the Reduction of Uncertainty (RoU), with darker shading indicating better data constraints (white: RoU  $\leq$  0.2, black: RoU  $\geq$  0.8).



**Fig. S2.** Regional disaggregation of Fig. 8: Interannual variability of SOCAT-based sea–air CO<sub>2</sub> flux estimates in comparison with the ocean process model NEMOv2.3 with PlankTOM5 by Buitenhuis et al. (2010) (using simulated  $p^{\text{CO}_2}$ , but gas exchange as in run **SFC**).



**Fig. S3.** Regional disaggregation of Fig. 9: Interannual variability of SOCAT-based sea–air CO<sub>2</sub> flux estimates in comparison with the results of an atmospheric transport inversion (Jena inversion s90\_v3.5).

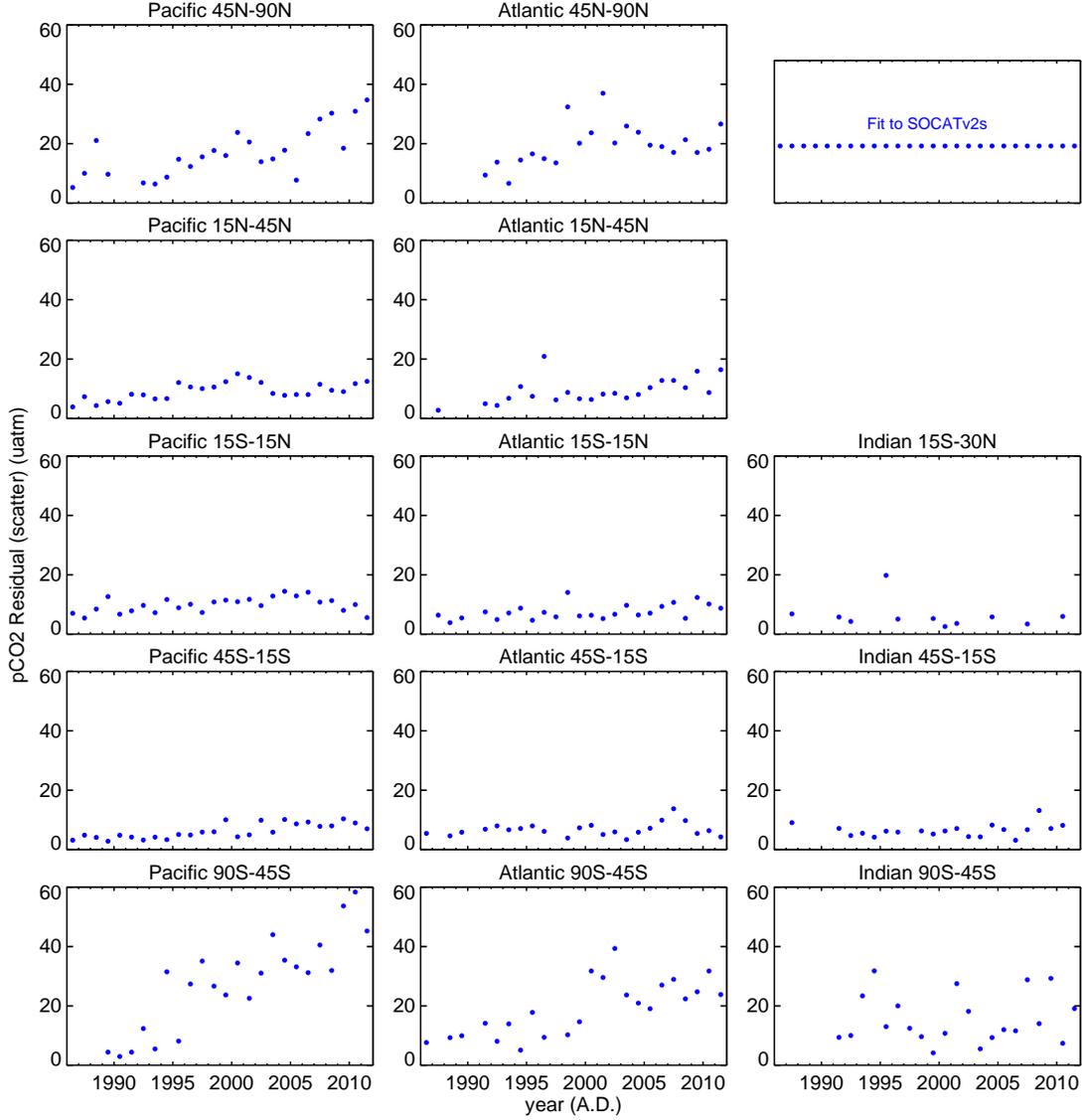


**Fig. S4.** Interannual anomalies of carbon fluxes in the equatorial Pacific, shown for three parts: Western (*left*), Central (*middle*), and Eastern (*right*). *Top*: Sea–air CO<sub>2</sub> flux anomalies ( $f_{\text{ma}}^{\text{CO}_2}$ , blue), together with its contribution OIS caused by ocean–interior sources/sinks (thin light-blue); *bottom*: anomalies in the ocean–internal DIC sources and sinks ( $f_{\text{int}}^{\text{DIC}}$ , blue, larger vertical range). Lines have been plotted only for years where the “reduction of uncertainty” generally exceeds 0.25. In each panel, the MEI index by Wolter and Timlin (1993) (interannually filtered as the fluxes, reversed sign, in arbitrary units) is overplotted as a thin black line.

## S2 Illustration of time lags

Sect. 3.4 and Fig. 6 of the main article discuss time lags between anomalies in different parts of the equatorial Pacific, or between ocean–internal and sea–air carbon fluxes. Using the identical MEI line as temporal reference between panels, Fig. S4 illustrates again several statements:

- Sea–air CO<sub>2</sub> flux anomalies ( $f_{\text{ma}}^{\text{CO}_2}$ , upper panels, dark blue) in the tropical Pacific are well anti-correlated to ENSO.
- In the eastern part (right panel), sea–air CO<sub>2</sub> flux anomalies lag ENSO more than in the western part (left).
- The OIS contribution (thin light-blue) accounts for most of the sea–air CO<sub>2</sub> flux anomalies, but does not show much West-to-East temporal phase shift.
- Ocean–internal DIC sources and sinks ( $f_{\text{int}}^{\text{DIC}}$ , lower panels) have a larger amplitude than the corresponding OIS contribution to sea–air CO<sub>2</sub> flux anomalies, and lead them by several months, due to the memory effect of the mixed-layer DIC budget (the time lag has been checked to be broadly consistent with the time constant of the budget equation of mixed-layer DIC concentration, using realistic values for piston velocity, Revelle factor, and mixed-layer depth, and assuming periodic forcing on a 4-year time scale).



**Fig. S5.** Size of the residuals between the estimated  $p^{\text{CO}_2}$  field and the observations, measured by the Scatter diagnostic of Eq. (S3.6) for the individual regions and yearly intervals.

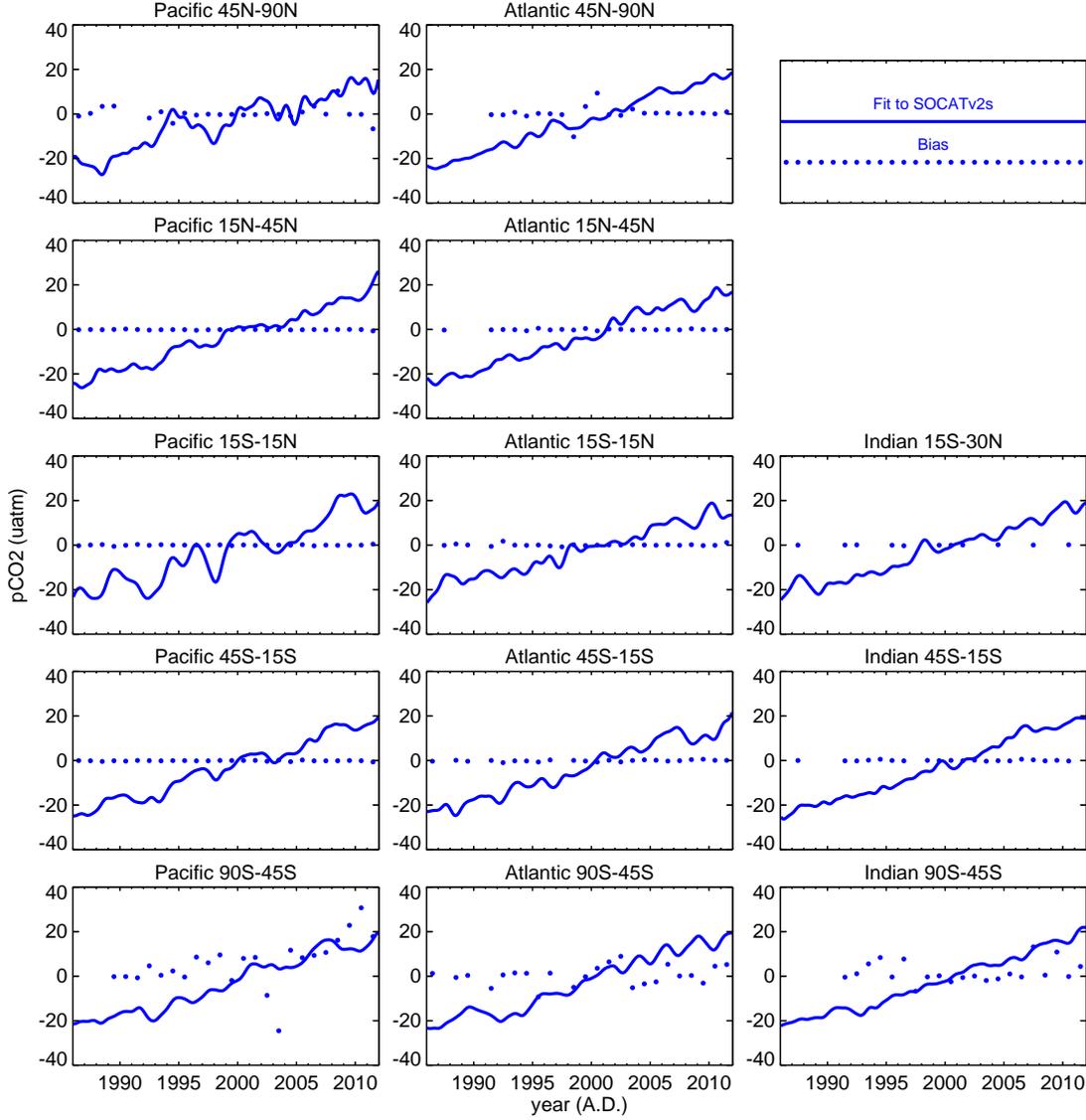
### S3 Fit to data

#### S3.1 Fit to the SOCAT data

The ability of the scheme to fit the SOCAT data has been measured by two diagnostics, calculated individually for regions and yearly intervals: The size of the remaining deviations between the estimated  $p^{\text{CO}_2}$  field (sampled at the locations and times of the measurement) and the observed values is quantified by

$$\text{Scatter} = \sqrt{\frac{1}{n} \sum_i \left( p_{\text{est},i}^{\text{CO}_2} - p_{\text{obs},i}^{\text{CO}_2} \right)^2} \quad (\text{S3.6})$$

where the sum is taken over all the  $n$  data points within the considered region and yearly interval. Due to its relation to the reduced  $\chi^2$  of Bayesian estimation theory (Tarantola, 1987), its global value should be close to the assumed model-data mismatch uncertainty (10 µatm). According to Fig. S5, this is, on average, even the case for the regions and years individually, in particular for the tropical Pacific. The larger residuals in the subpolar regions are exclusively due to areas North or South of 60°.

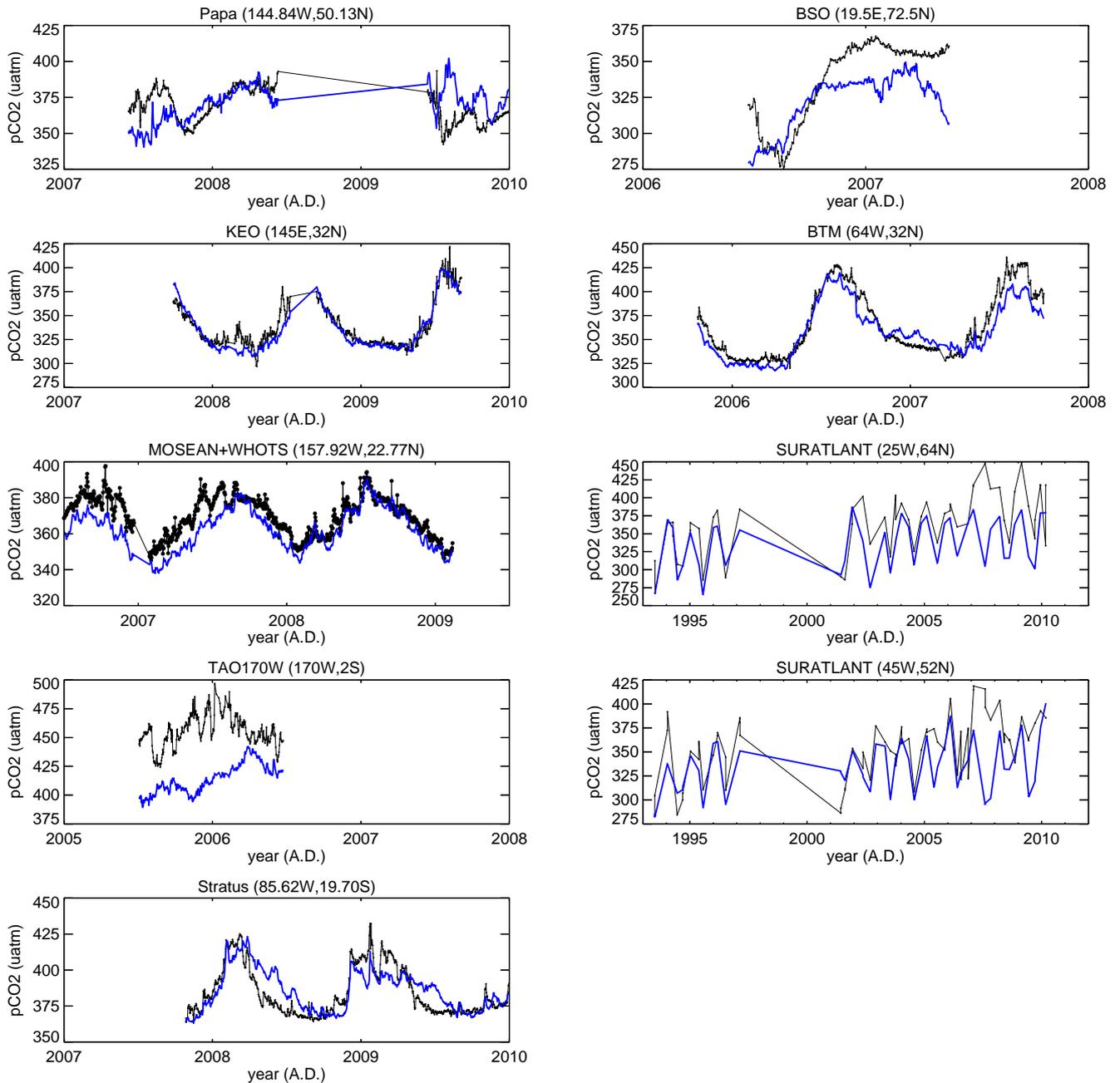


**Fig. S6.** Systematic biases (dots) of the estimated  $p^{\text{CO}_2}$  field with respect to the observations, measured by the diagnostic of Eq. (S3.7) for the individual regions and yearly intervals. For comparison, the interannual anomalies of the  $p^{\text{CO}_2}$  regional averages are overplotted (lines).

In order to reveal any remaining systematic structure in the residuals on interannual time scales, we also consider a mean bias,

$$\text{Bias} = \frac{1}{n} \sum_i \left( p_{\text{est},i}^{\text{CO}_2} - p_{\text{obs},i}^{\text{CO}_2} \right) \quad (\text{S3.7})$$

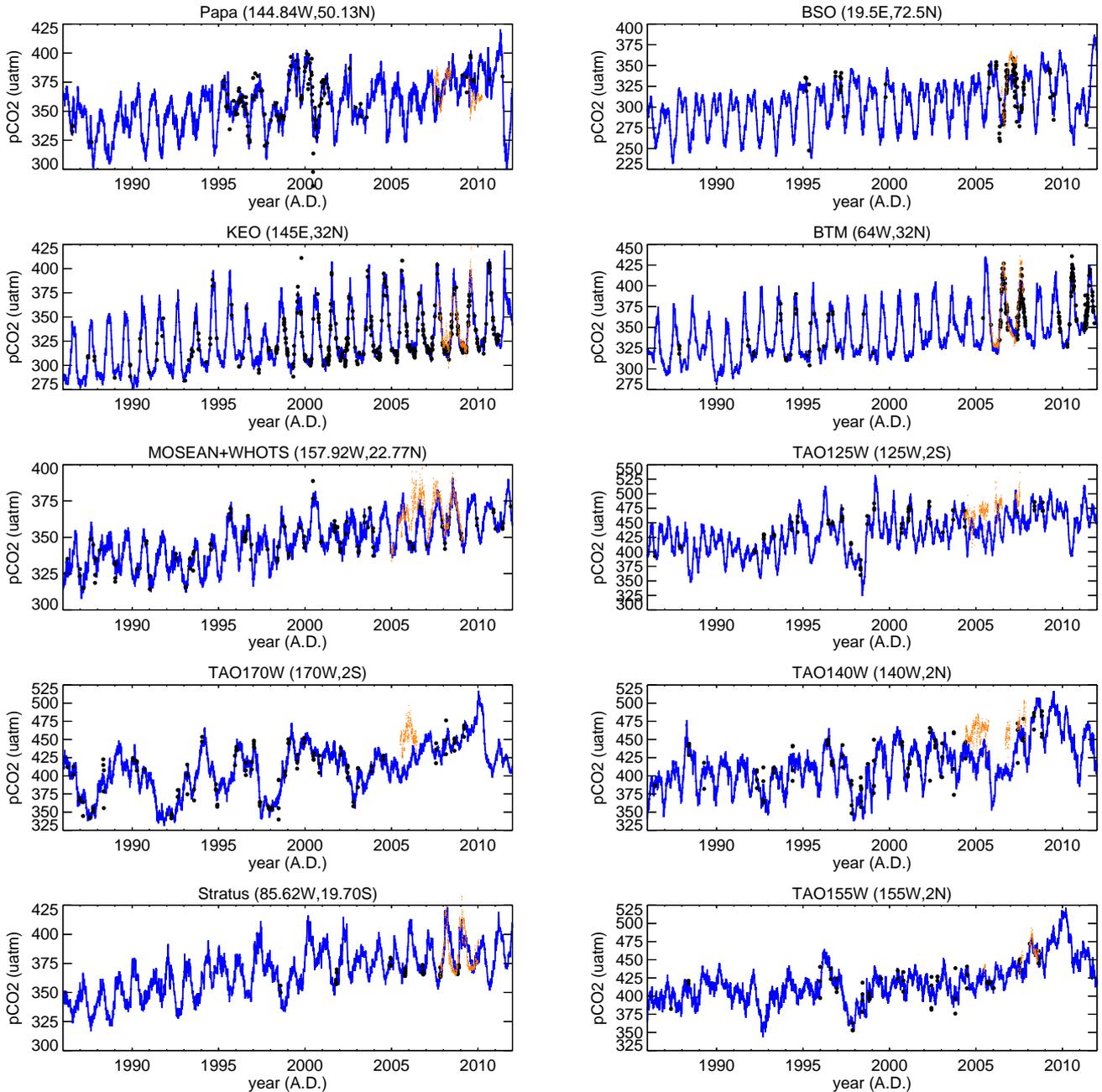
Yearly biases in tropical and subtropical regions (in particular the tropical Pacific) are below  $0.5 \mu\text{atm}$ , in most years even below  $0.1 \mu\text{atm}$ . According to Fig. S6, this is negligible compared with the year-to-year variations in  $p^{\text{CO}_2}$  (which are, e.g., on the order of  $20 \mu\text{atm}$  in the tropical Pacific). We conclude that the diagnostic scheme indeed uses the interannual information from the data completely (again, the larger biases in the subpolar regions are restricted to the high Northern and Southern latitudes). Note that possible sampling biases –for example aliasing of seasonal variability into year-to-year variability due to sampling the seasons differently in different years– cannot be revealed by the bias test, because the fit and the test would be affected in the same way.



**Fig. S7.** Comparison between the estimated  $p^{\text{CO}_2}$  field (run SFC, blue) and independent  $p^{\text{CO}_2}$  measurements not in SOCAT v2\*, at various locations in the Pacific (left) or Atlantic (right): BSO – Arthun et al. (2012), SURATLANT – Corbière et al. (2007); MetzI et al. (2010), others – Sabine et al. (2010). The estimated field has been picked at the times where measurements exist (connected by straight lines for clarity); the time axes have been limited to respective years with data.

### S3.2 Comparison to independent data – Seasonal and day-to-day variations

Updating Fig. S7.6 of Rödenbeck et al. (2013) based on SOCAT v1.5, a comparison between the presented  $p^{\text{CO}_2}$  field based on SOCAT v2 (shipboard data) and independent observations at various locations is shown in Fig. S7 (note that some of these comparison time series (moorings except for BSO) are actually part of SOCAT v2 but have been retained from the fit to be able to use them in this validation). The additional data after 2007 improve the correspondence at several locations (KEO, MOSEAN+WHOTS, Stratus). A slight deterioration is seen at BSO. Broad agreement is also found in many day-to-day

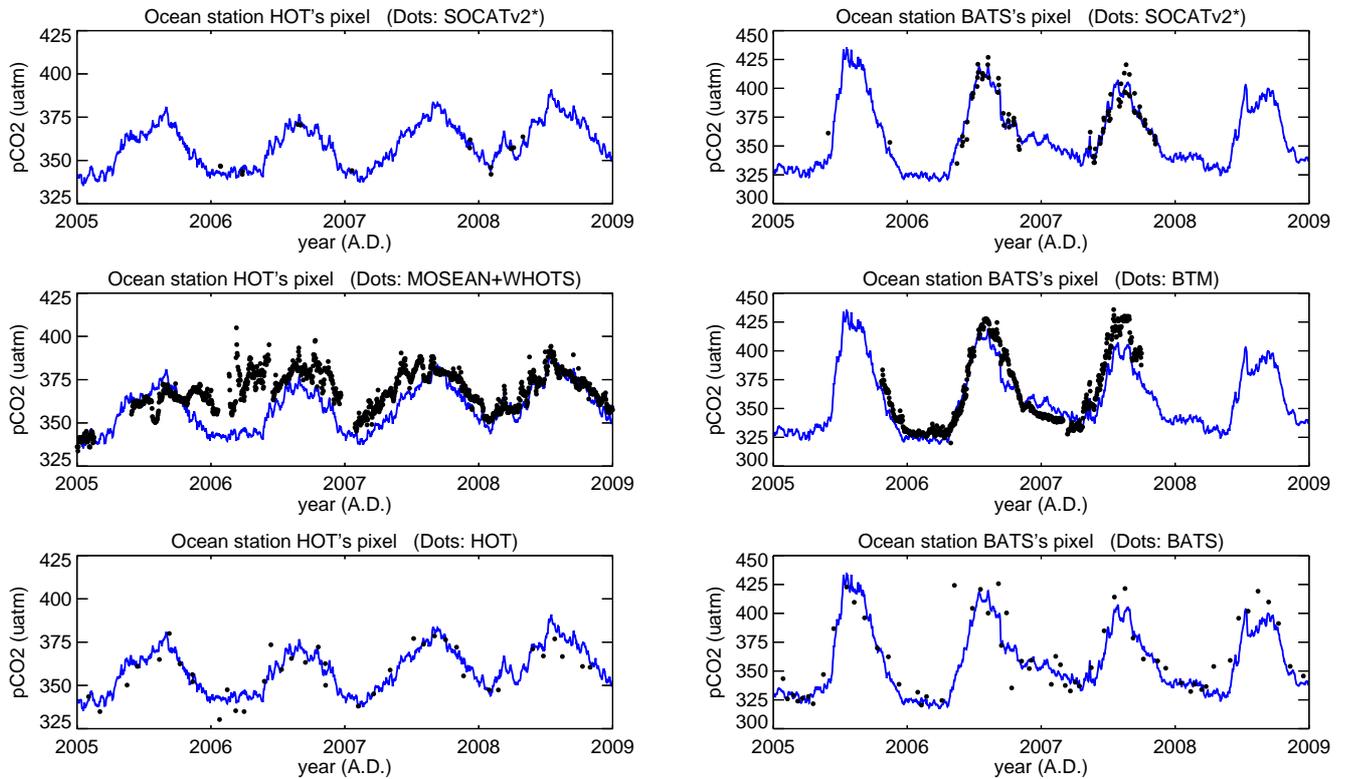


**Fig. S8.** Check of the fit to the data: Comparison between the estimated  $p^{\text{CO}_2}$  field (run **SFC**, blue) and all the data points from SOCAT v2\* (black) present at the respective pixels around several mooring sites. The orange points repeat the independent measurements from Fig. S7.

variations, even though these are generally not constrained by the  $p^{\text{CO}_2}$  data (which normally are much less frequent in time) but only parametrized (mainly following SST variations) (compare Fig. S7.7 of Rödenbeck et al., 2013).

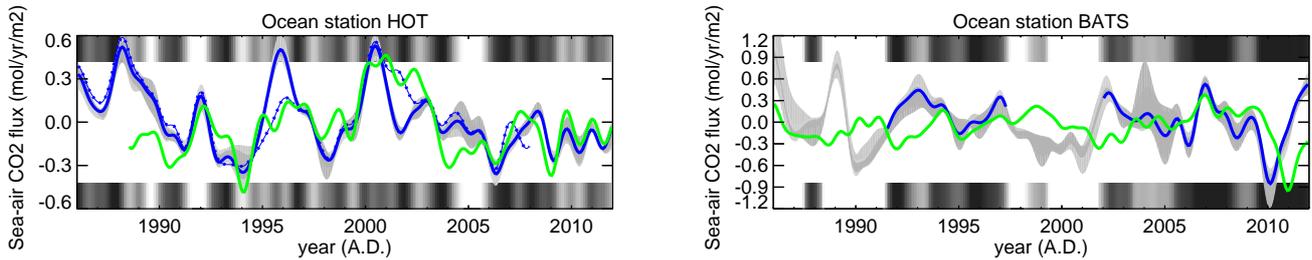
To set the agreement with the independent data into perspective, Fig. S8 compares the  $p^{\text{CO}_2}$  field at the same locations to the data at these pixels that have been used in the fit. For example, at TAO170W where discrepancies to the independent mooring data are particularly large, Fig. S8 reveals that the fit to the shipboard data is much more close, but that there are no shipboard data during or around the mooring period in 2005. Thus the estimated  $p^{\text{CO}_2}$  field is not constrained to follow the excursion towards larger values even though features of this amplitude are seen to happen in many other years.

An even more detailed look is possible at two particular locations where 3 independent  $p^{\text{CO}_2}$  observation records exist (Fig. S9): Besides the shipboard data used in the fit (top panels), there are independent moorings (middle panels) and  $p^{\text{CO}_2}$



**Fig. S9.** CO<sub>2</sub> partial pressure in the respective pixel around the ocean stations HOT (left) and BATS (right) for 4 selected years. Top panels show the data points in SOCATv2\* that happen to fall into these pixels, together with the results of run **SFC** fitting these data (blue line). Middle and bottom panels compare the same run **SFC** results (blue lines) to independent data not used in the fit: Moorings MOSAN and WHOTS (middle left), mooring BTM (middle right),  $p^{\text{CO}_2}$  calculated from HOT station's DIC and Alk by Dore et al. (2009) (bottom left), and  $p^{\text{CO}_2}$  calculated from BATS station's DIC and Alk by Bates et al. (2012) (bottom right).

values calculated from measured DIC (bottom panels). While the fit to the shipboard data points is quite good (as it should), it is obvious that the scheme cannot possibly fit the independent time series simultaneously. As the measurements are done at close-by but not exactly the same places (they fall into the same  $4^\circ \times 5^\circ$  pixel), spatial heterogeneity within the respective pixel may explain part of this discrepancy. Additional differences may arise from unavoidable technical and environmental factors affecting the oceanographic measurements. Note that the relatively high values in the MOSEAN+WHOTS moorings in the first half of 2006 (Feb-02 to June-08) have been identified as suspect (SOCAT release note 2013-07-15, <http://cdiac.ornl.gov/ftp/oceans/SOCATv2/>).



**Fig. S10.** Interannual anomalies of the sea–air CO<sub>2</sub> fluxes around the ocean stations HOT (left) and BATS (right). SOCAT-based results (run **SFC**, blue) are compared to estimates based on independent regular data time series (Dore et al. (2009) and Bates et al. (2012), respectively, green). Periods with low “Reduction of uncertainty” in run **SFC** (shading at top and bottom), or not covered by the independent time series, respectively, have been omitted from the plots. At station HOT (left) results based on SOCATv1.5 have been added as dotted line.

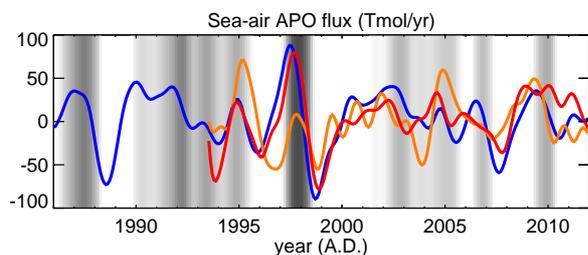
### S3.3 Comparison to independent time series carbon data – interannual variations

Only very few regular multi-year time series of carbon observations independent from the SOCAT data base exist to validate the interannual variations in our estimates. One of these is the independent data by Dore et al. (2009) measured at the Hawaii Ocean Time-series station (HOT, 22° 45'N, 158° 00'W) in the North Pacific. We use  $p^{\text{CO}_2}$  values calculated by Dore et al. (2009) from the co-located measurements of DIC, Alk, temperature, and salinity. The seasonal amplitude in our  $p^{\text{CO}_2}$  estimates at the pixel of the HOT station is confirmed by the data (Fig. S9). In order to be able to compare interannual variations, we interpolated the HOT  $p^{\text{CO}_2}$  data set by fitting the diagnostic scheme to it<sup>10</sup>, such that we can apply the same interannual filter both to the SOCAT-based results and to the HOT validation data (Fig. S10, left). The interannual variations obtained from the HOT  $p^{\text{CO}_2}$  values agree relatively well to the SOCAT-based results both in the decadal and many year-to-year features, even though SOCAT does not contain regular data points within the pixel enclosing HOT. The agreement was even closer for SOCAT v1.5 than it is for SOCAT v2. To confirm that this agreement is not due to the common variability in the driving fields, we did test runs using purely seasonal driving fields (see Sect. 2.3.2): This hardly changes the signals of Fig. S10 (changes much smaller than differences between HOT-based and SOCAT-based interannual variations, not shown), such that the signals seen must come from the SOCAT or HOT data, respectively.

The agreement between the HOT-based and SOCAT-based results suggests that the interannual variations around HOT are spatially coherent enough such that local variations can also be picked up from nearby SOCAT data, spread laterally by the spatial correlations in our scheme (see Fig. 5 of the companion paper Rödenbeck et al., 2013, for illustration).

Less successful comparison is found for station BATS (Bates et al., 2012) (Fig. S10, right). This is probably due to the combined effect of less continuous data coverage (Fig. 4) and smaller (and maybe more heterogeneous) interannual variations in the Northern Atlantic.

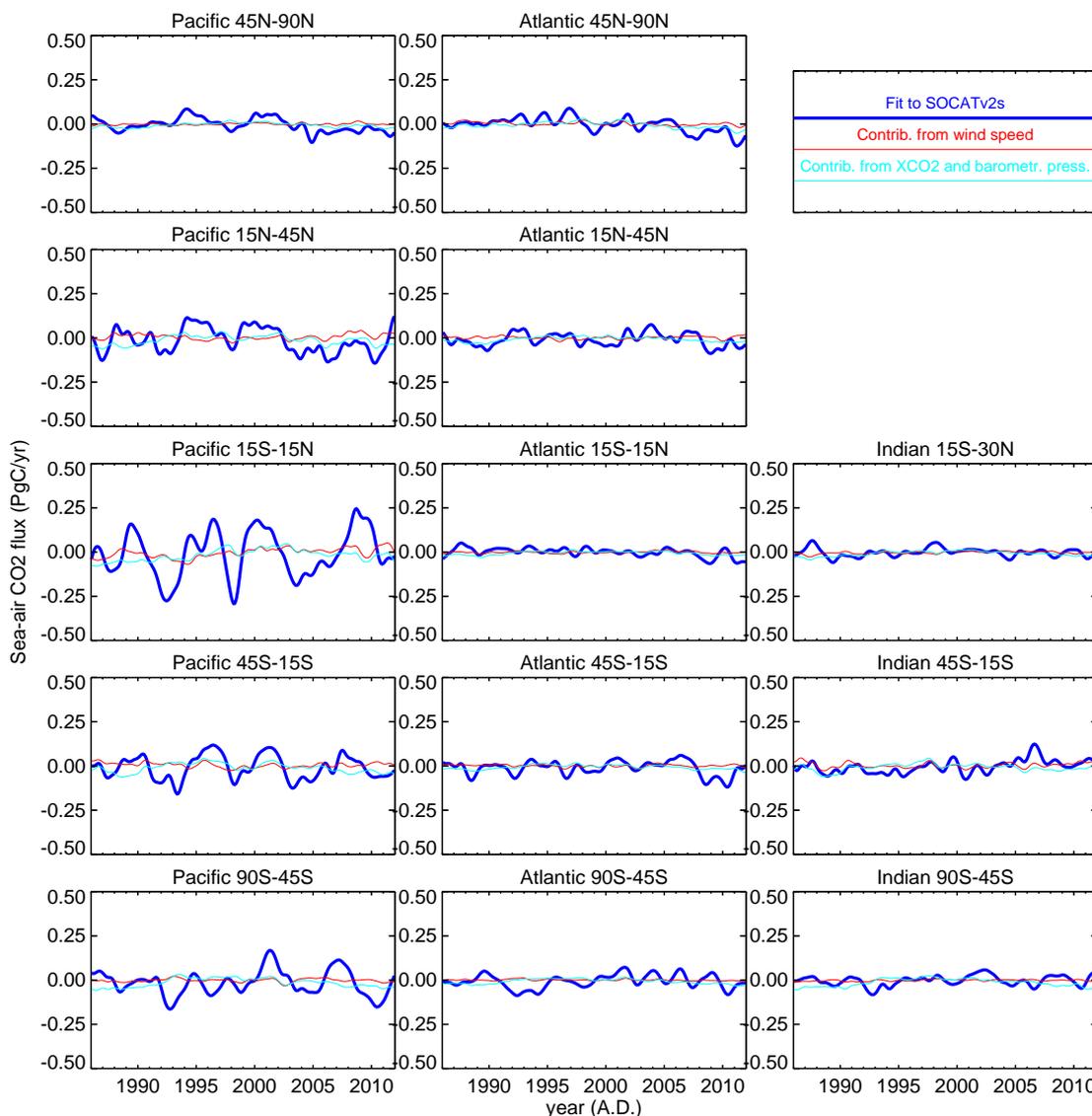
<sup>10</sup> We performed a fit of the diagnostic scheme just as in run **SFC**, except that the SOCAT data points were replaced by HOT data (of course, this only constrains the pixel considered here, i.e., the scheme is effectively operated locally without space dependence). This creates a time series that matches the HOT data points and is compatible with the mixed-layer budget, chemistry, and gas exchange in between.



**Fig. S11.** Interannual anomalies of the APO exchange inferred from SOCAT (blue) and atmospheric oxygen data (orange) as in Fig. 7 (top), with the result of an APO inversion of synthetic data added (red).

#### S4 Strength of the atmospheric oxygen constraint

A potential cause of mismatch between the APO flux estimates based on SOCAT or on atmospheric oxygen data (Fig. 7) is missing information about spatial details of the APO flux field in the atmospheric data, as the atmospheric APO inversion is based on 5 observation sites only. To quantify that effect, we performed a “synthetic” APO inversion where (1) the APO flux field inferred from SOCAT was transported in the atmospheric tracer transport model used in the inversion, (2) the resulting atmospheric APO field was sampled at the locations/times of the APO observations, (3) an APO inversion was performed using these pseudo-data (all other settings identical to the APO inversion with real data). The result is added in Fig. S11. The interannual anomalies retrieved from the pseudo-data generally reproduce the underlying anomalies on the SOCAT-based field. Largest deviations occur around 2006, where also the real APO inversion is missing the ENSO response. In other periods, deviations are too small to explain the discrepancies between the two data-based approaches. Nevertheless, a tendency of delayed peaks and troughs is evident in some ENSO events (1997/98, 2007), tentatively supporting the earlier timing.



**Fig. S12.** The interannual anomalies of the sea–air CO<sub>2</sub> exchange estimated from SOCAT data as in Fig. S1 (blue) and the contributions of variability in wind speed (red) and in atmospheric CO<sub>2</sub> mixing ratio and barometric pressure (cyan).

### S5 Contribution of gas exchange variability

Part of the interannual variability in the sea–air CO<sub>2</sub> flux is due to variations in gas exchange. We assessed the contributions of several driver variables to IAV, by calculating the differences between the full sea–air CO<sub>2</sub> flux and that obtained when the respective driver variable only varies seasonally without IAV. The two largest contributions come from IAV in wind speed and in atmospheric CO<sub>2</sub> partial pressure (product of atmospheric CO<sub>2</sub> mixing ratio and dry-air pressure). However, both of these contributions are small compared to the total variability (Fig. S12).

Through Schmidt number and solubility, gas exchange also depends on SST. However, as these two paths of influence largely cancel each other (Takahashi et al., 2009), the SST contribution is small and has been omitted from Fig. S12.