Sea-air CO2 flux estimated from SOCAT surface-ocean CO2 partial pressure data and atmospheric CO2 mixing ratio data

Christian Rödenbeck, Ralph F. Keeling, Dorothee C. E. Bakker, Nicolas Metzl, Are Olsen, Christopher L. Sabine, M. Heimann

Ocean Sci. Discuss., 9, 2273–2326, 2012
www.ocean-sci-discuss.net/9/2273/2012/
doi:10.5194/osd-9-2273-2012

Ocean Science Discussions

Sea-air CO_2 flux estimated from SOCAT surface-ocean CO_2 partial pressure data and atmospheric CO_2 mixing ratio data

C. Rödenbeck[1], R. F. Keeling[2], D. C. E. Bakker[3], N. Metzl[4], A. Olsen[5,6,7], C. Sabine[8], and M. Heimann[1]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany
[2]Scripps Institution of Oceanography, University of California, San Diego, USA
[3]School of Environmental Sciences, University of East Anglia, Norwich, UK
[4]LOCEAN-IPSL, CNRS, Paris, France
[5]Institute of Marine Research, Bergen, Norway
[6]Uni Bjerknes Centre, Bergen, Norway
[7]Bjerknes Centre for Climate Research, Bergen, Noway
[8]NOAA Pacific Marine Environmental Laboratory, Seattle, USA

Received: 27 April 2012 – Accepted: 15 May 2012 – Published: 27 June 2012

Correspondence to: C. Rödenbeck (christian.roedenbeck@bgc-jena.mpg.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Surface-ocean CO_2 partial pressure data have been assimilated into a simple diagnostic model of surface-ocean biogeochemistry to estimate the spatio-temporal CO_2 partial pressure field and ultimately the sea-air CO_2 fluxes. Results compare well with the widely used monthly climatology by Takahashi et al. (2009) but also contain some short-term and interannual variations. Fitting the same model to atmospheric CO_2 data yields less robust but consistent estimates, confirming that using the partial pressure based estimates as ocean prior in atmospheric CO_2 inversions may improve land CO_2 flux estimates. Estimated seasonality of ocean-internal carbon sources and sinks is discussed in the light of observed nutrient variations.

1 Introduction

The oceans are considered the dominant player in the global carbon cycle on long time scales, e.g., in the glacial-interglacial cycles (e.g., Sigman and Boyle, 2000). On a multi-millenial time scale, the oceans will be the sink for 80–95 % of the anthropogenic CO_2 emissions, and 70–80 % on a time scale of several hundred years (Archer et al., 1997). Currently, the oceans take up about 25 % of the emissions. Concerns exist, however, that the sink efficiency may decrease in the coming decades as a consequence of global warming, as suggested by model projections (e.g., Sarmiento and Le Quéré, 1996; Matear and Hirst, 1999; Joos et al., 1999) and tentatively confirmed by data analysis (e.g. Le Quéré et al., 2007, 2010). As a prerequisite to understanding the involved processes, one needs to quantify sea-air CO_2 fluxes, their variability, and their response to forcing.

Currently, there are two data streams used to estimate the variability of sea-air CO_2 fluxes:

– Based on measurements of the CO_2 partial pressure (p^{CO_2}) of surface water, the sea-air CO_2 flux is calculated through a gas exchange parametrization. As p^{CO_2}

OSD
9, 2273–2326, 2012

measurements only exist along ship cruise tracks, spatial and temporal interpolation is needed. The following interpolation methods have been proposed: (1) statistical interpolation combined with an advection-diffusion equation (Takahashi et al., 2009); (2) purely statistical interpolation with error quantification (Jones et al., 2012b); (3) Multilinear Regressions between p^{CO_2} and ocean state variables (e.g., Park et al., 2010; Chen et al., 2011); (4) neural networks learning the relationship between p^{CO_2} and oceanic state variables available from remote sensing platforms and ocean reanalysis projects (e.g., Lefèvre et al., 2005; Telszewski et al., 2009); (5) assimilation of the p^{CO_2} data into a process model of ocean biogeochemistry (Valsala and Maksyutov, 2010; While et al., 2012). These methods are complementary in certain aspects: the more statistical methods are dominated by the data themselves but are strongly affected by spatial and temporal gaps, while the more complex methods more easily spread the information but are dependent on driver data sets or even on model formulations.

– Surface-atmosphere CO_2 fluxes can be quantified based on atmospheric CO_2 mixing ratio measurements (e.g., Conway et al., 1994, and many others) by the atmospheric transport inversion (e.g., Newsam and Enting, 1988; Rayner et al., 1999; Bousquet et al., 2000; Rödenbeck et al., 2003; Baker et al., 2006). This estimation does not involve any parametrization of gas exchange. However, ocean fluxes are difficult to detect by this method in most parts of the globe because their imprint on the atmospheric mixing ratio records is small compared to that of the much more variable land fluxes: even if inorganic carbon sources and sinks in the ocean interior are comparable in strength to those of the land biosphere, they only lead to dampened and delayed sea-air CO_2 fluxes due to the buffer effect of marine carbonate chemistry. A further problem of atmospheric inversions based on data from a discrete set of measurement sites is that information is only provided on scales comparable to or larger than the distance between the sites or the sampling frequency, while variability exists also on smaller scales.

OSD
9, 2273–2326, 2012

A further data-based method to estimate sea-air fluxes uses ocean-interior carbon data in inverse calculations of oceanic transport (Gloor et al., 2003; Mikaloff Fletcher et al., 2006). This method is independent of parametrizations of gas-exchange as well. However, it only yields long-term sea-air CO_2 fluxes over large spatial regions, not its temporal or high-resolution spatial variability. The long-term global sea-air CO_2 flux has also been estimated from observed trends in atmospheric oxygen (Keeling and Shertz, 1992) or ^{13}C isotopic ratios in atmospheric CO_2 (Ciais et al., 1995).

In view of the above-mentioned problem of the atmospheric CO_2 inversions, most of these studies use the sea-air flux climatology by Takahashi et al. (2009) (or earlier versions) as Bayesian prior. (In some cases, the long-term fluxes from ocean-interior inversions are used in the prior as well, e.g., Rödenbeck et al. (2003), or formalized as joint inversion by Jacobson et al. (2007).) To combine the two considered data streams directly, this study proposes an extension of the atmospheric inversion method by a diagnostic data-driven model of mixed-layer biogeochemistry that can alternatively or simultaneously be constrained by atmospheric CO_2 data or surface-ocean p^{CO_2} data. Among the range of methods for interpolating p^{CO_2} as listed above, the proposed diagnostic scheme is intermediate in that it spreads the data information in a simple statistical way but also incorporates small-scale patterns from wind and SST dependencies. It offers an easy and self-consistent way to compare or to combine atmospheric and oceanic observational information, because both are used in a common framework. While we focus here exclusively on using carbon data, the diagnostic scheme has been developed as part of a system that can also use constraints from other tracers, such as oxygen.

The presented results are based on the newly available Surface Ocean CO_2 Atlas (SOCAT) data base (v1.5) of p^{CO_2} measurements (Pfeil et al., 2012). As a first step, this study will mainly be focused on the mean seasonal cycle of mixed-layer biogeochemistry. We compare the results based on SOCAT p^{CO_2} data to those based on atmospheric CO_2 data, and discuss the mutual consistency. The origin of ocean-internal carbon sources and sinks is discussed in the light of observed nutrient variations.

OSD
9, 2273–2326, 2012

2 Method

2.1 Concept – overview

In a classical atmospheric inversion, a spatio-temporal field of surface-to-atmosphere tracer fluxes is estimated such that its corresponding mixing ratio field – as simulated by an atmospheric tracer transport model – matches as closely as possible a set of mixing ratio observations. The match is gauged by a quadratic cost function to be minimized (Appendix A2).

Here we extend the inversion framework by not only considering the process of atmospheric transport but also processes in the oceanic mixed layer determining sea-to-air CO_2 exchanges in dependence of ocean-internal carbon sources and sinks (Fig. 1). The atmospheric transport model is extended into a chain of parametrizations of gas exchange, carbonate chemistry, and a carbon budget equation. Then the inversion is not adjusting the sea-to-air fluxes any more, but the ocean-internal fluxes instead. This has two purposes: (1) the representation of oceanic processes involves further quantities, notably CO_2 partial pressure. Through cost function contributions gauging the match of modelled and measured p^{CO_2}, these data can be used as an observational constraint replacing (or in addition to) the atmospheric data. (2) The equations modelling the individual processes impose spatial and temporal structure to the sea-air fluxes, using information from the driver data (SST, wind speed, etc.).

2.2 Process parametrizations

All processes considered explicitly are given in Fig. 1 and summarized in the following. Details, including the equations used, are found in Appendix A.

Atmospheric transport. Atmospheric tracer mixing ratio fields in response to surface-to- atmosphere fluxes are simulated by the global off-line atmospheric transport model TM3 (Heimann and Körner, 2003) with a spatial resolution of $\approx 4^\circ$ lat. $\times 5^\circ$ long. $\times$ 19 vertical levels. The model is driven by 6-hourly interannual meteorological fields derived from the NCEP reanalysis

OSD
9, 2273–2326, 2012

(Kalnay et al., 1996). The model fields are sampled at the location and time of the individual mixing ratio measurements used.

Solubility and gas exchange. Diffusive sea-to-air gas exchange is proportional to the over/under-saturation and to the piston velocity with a quadratic wind speed dependence (Appendix A1.1).

Carbonate chemistry. The carbon species relevant for gas exchange (CO_2) only accounts for a small part of the carbon relevant in the ocean-internal budget (dissolved inorganic carbon, DIC). The link between CO_2 abundance (expressed in terms of partial pressure $p_m^{CO_2}$) and DIC abundance (in terms of its concentration C_m^{DIC}) is determined by chemical equilibria, which we assume to be attained instantaneously. The dependence $p_m^{CO_2} = p_m^{CO_2}(C_m^{DIC})$ is actually non-linear, but has been linearized around the measured long-term values. It also contains a temperature-dependent factor, and contributions from seasonal variations in alkalinity and salinity (Appendix A1.2).

Mixed-layer tracer budget. Changes in the spatio-temporal field of dissolved tracer in the mixed layer need to be balanced by the sum of fluxes (Appendix A1.3). As the sea-air flux itself depends on tracer concentration, this leads to a differential equation the solution of which gives tracer concentration as a function of ocean-internal sources and sinks f_{int}.
An additional important item in the budget is the re-entrainment during mixed-layer deepening of tracer left behind previously during mixed-layer shoaling, which we represent as a "history flux" f_{hist}. We further consider the influence of freshwater fluxes.

2.3 Data constraints – base runs

We will present results of two main cases, that use the same diagnostic scheme but different main data sets as constraint:

SFC. Fit to surface p^{CO_2} data points from the SOCAT v1.5 data base (Pfeil et al., 2012). Data pre-treatment and further details are given in Table 1.

ATM. Fit to atmospheric CO_2 mixing ratios measured approximately weekly or hourly at a set of sites by various institutions. Details in Table 1. (This case is similar to a classical atmospheric CO_2 inversion, except that sea-air fluxes are not estimated directly as adjustable degrees of freedom, but indirectly through the ocean process parametrizations by adjusting f_{int}^{DIC}

OSD
9, 2273–2326, 2012

(Appendix A2.2); the parametrizations thus also co-determine the a-priori value and uncertainty structure of the sea-air fluxes.)

2.4 Adjustable degrees of freedoms

The basic unknown to be adjusted to match the surface-ocean or atmospheric data is the spatially and temporally varying field of ocean-internal carbon sources and sinks ($f_{\mathrm{int}}^{\mathrm{DIC}}$, Sect. A1.3). Similar to the unknown sea-air flux in the pure atmospheric transport inversion, Bayesian a-priori spatial and temporal correlations have been implemented to enforce the flux field to be smooth on scales smaller than around 1910 km (longitude), 960 km (latitude), and about 2 weeks (time). These correlations represent the main mechanism by which the information from the discrete atmospheric or oceanic data points is interpolated into space and time. The chosen lengths are identical to the ocean flux correlations in the standard Jena Inversion (update of Rödenbeck, 2005), and similar to those of $p^{\mathrm{CO_2}}$ found from correlation analysis along cruise tracks by Jones et al. (2012a). In the time domain, we allow in the present study only seasonal degrees of freedom in $f_{\mathrm{int}}^{\mathrm{DIC}}$, such that $f_{\mathrm{int}}^{\mathrm{DIC}}$ is forced to be a smooth seasonal climatology. This suppresses any spurious interannual variations that may otherwise arise from the very different coverage of the SOCAT data in the individual years. Though estimated $p^{\mathrm{CO_2}}$ and sea-air fluxes will still involve interannual variations from the driver quantities (sea surface temperature T, wind speed u), we will correspondingly mainly consider the seasonality only. (On prospects to also obtain interannual variations see Sect. 4.3 below.)

In addition to the unknown ocean-internal carbon fluxes, the inversion also adjusts land-atmosphere $\mathrm{CO_2}$ fluxes (in case **ATM** only) as well as technical degrees of freedom related to initial conditions and mean state (Sect. A2.2).

OSD
9, 2273–2326, 2012

2.5 Testing the method

2.5.1 Synthetic data test

As a prerequisite to interpreting the results of fitting the scheme to surface-ocean p^{CO_2} data (run **SFC**), Appendix B establishes by "synthetic data testing" that (1) the scheme contains sufficient degrees of freedom to reproduce the p^{CO_2} seasonality in the different parts of the ocean, and (2) the information available in the SOCAT data set is sufficient to constrain it. Analogous tests for run **ATM** are given in the Supplement, Sect. S4.

2.5.2 Sensitivity test

To what extent do the results depend on choices taken in the algorithm, and potential errors in the driver data sets used? Various detailed influences of individual model elements were tested in sensitivity runs (Supplement, Sect. S2). In particular, the tightness of the Bayesian priors (weighting of prior constraint vs. data constraint) was varied, to test how robustly the results depend on the data. Further sensitivity cases are related to correlation length, driver data, and others.

3 Results

3.1 Overview

To illustrate the characteristics of the quantities and the temporal scales involved in the scheme, Fig. 2 shows time series of key quantities, averaged/integrated over the ocean South of 45° S as an arbitrary example region. The ocean-internal sources/sinks (bottom panel) are a smooth seasonal climatology by construction (Sect. 2.4). The mixed-layer DIC concentration (panel above) responds to these ocean-internal fluxes and the emerging sea-air exchange. Being their temporal integral, the DIC concentration is slightly shifted in phase with respect to the ocean-internal fluxes. Besides its

OSD
9, 2273–2326, 2012

seasonality, the DIC concentration is rising in response to the atmospheric CO_2 increase. The surface-ocean equivalent CO_2 partial pressure (3rd panel from bottom) shares the rise and the seasonality with the DIC concentration, but the seasonality is again slightly shifted because of the temperature and alkalinity dependent relation. The sea-air CO_2 flux (top panel) is dominated by the p^{CO_2} variability. In addition, it shows high-frequency variability due to wind speed and solubility changes.

The remainder of the paper focuses on the mean seasonal cycle of p^{CO_2} and sea-air CO_2 flux. Figure 3 shows the amplitude of the p^{CO_2} seasonality for each pixel, thereby illustrating the spatial resolution and domain of the calculation.

3.2 Estimates based on SOCAT p^{CO_2} data (run SFC)

Figure 4 shows the mean seasonal cycle of p^{CO_2} estimated from SOCAT data. The p^{CO_2} field has been integrated over a set of regions splitting the ocean into basins and latitude bands. The narrow grey range of sensitivity results (Sect. 2.5.2) indicates that the results are robust against several potential error influences.

In most extra-tropical regions, the seasonal cycle estimated from SOCAT is similar in phase to the p^{CO_2} climatology by Takahashi et al. (2009). However, the amplitude is larger in many regions, in particular in the high-latitude North Pacific and the temperate North Atlantic. Both these regions contain spots of particular large seasonal amplitudes in our results (Fig. 3). In the *North Pacific*, direct comparison to the SOCAT data at the individual pixels tends to confirm the larger amplitudes (see Fig. 5 (top) for a typical example). Nevertheless, the regional average is considerably influenced by the Northernmost pixels where the data do not offer sufficient seasonal coverage for such a comparison; it therefore remains open whether the direct spatial extrapolation of p^{CO_2} as in Takahashi et al. (2009) or the extrapolation of f_{int}^{DIC} as in the diagnostic model is more realistic (note that the fit to synthetic data (Appendix B) confirms that the diagnostic model would be flexible enough to also follow the Takahashi et al. (2009) pattern if data would be there to constrain it that way; not also, however, that our fit is less robust in the North Pacific than elsewhere as seen from the slightly wider grey band

OSD
9, 2273–2326, 2012

in Fig. 4). In the *temperate North Atlantic*, the discrepancy between the two estimates is mainly related to the Mediterranean (compare Supplement, Fig. S7.1). Importantly, Takahashi et al. (2009)'s climatology does not actually cover this region, why we used p^{CO_2} extrapolated from its Atlantic open-ocean values. However, direct comparison to SOCAT data (Fig. 5, bottom) does again support the unusually high seasonal amplitude, though seasonal data coverage at this pixel is far from complete (it is still the only pixel in the Mediterranean covering more than a short individual period). Also in other regions, coastal areas and marginal seas account for part of the differences in Fig. 4. See Supplement (Sect. S5) for further analysis of residuals (differences between the fitted p^{CO_2} field and the original data points).

Partly, the differences in Fig. 4 can be traced to the fact that both estimates are based on different p^{CO_2} data sets (Supplement, Sect. S6), or to interannual variations. Differences in the Tropics may be related to the exclusion of El Niño data in Takahashi et al. (2009). Moreover, especially in the Indian, the synthetic data test (Appendix B) reveals spuriously high seasonality in our Tropical results, maybe due to a too variable and too strong prior.

3.3 Comparison of p^{CO_2} based and atmospheric results (run ATM)

Figure 6 gives the result of fitting the diagnostic scheme to atmospheric CO_2 data (run **ATM**). The result of the fit to SOCAT from Fig. 4 is given again for comparison. In the Southern Hemisphere, there is broad agreement of the atmospheric-based and SOCAT-based results in the phasing of the seasonal cycle. Less agreement is found in the Tropics and the Northern Hemisphere, in particular in the temperate latitudes.

The grey band around the atmospheric results in Fig. 6 indicates its sensitivity to the tightness of the Bayesian a-priori constraint. Though this is only a subset of sensitivity runs considered for run **SFC** in Fig. 4, the range is much wider, revealing that the fit to the atmospheric CO_2 data is much less robust than the fit to the SOCAT data (run **SFC**). As investigated by synthetic data testing (Supplement, Sect. S4), this is largely related to the additional degrees of freedom in adjusting land-to-atmosphere

OSD
9, 2273–2326, 2012

fluxes: signals in the atmospheric data are partially wrongly attributed to land or ocean. Consistently, best agreement between **ATM** and **SFC** occurs in the Southern regions away from land influence, while the large disagreement in the North temperate Pacific is due to the short but strong sink period during summer, typical for terrestrial boreal ecosystems.

3.4 Comparison to DIC measurements

Figure 7 compares the estimates based on SOCAT with independent DIC measurements from the BATS station (Bates, 2011). There is relatively good agreement in phase and amplitude of the seasonal cycle. The a-priori estimates (Sect. A2.2) do not yet show such an agreement, confirming that the agreement is related to the information from the SOCAT data. The comparison also indicates that the approximations in our carbonate chemistry parametrization (Sect. A1.2) are acceptable.

4 Discussion

4.1 Combining oceanic and atmospheric data

The results establish that oceanic and atmospheric data are consistent in regions away from land influences, but that estimates based on p^{CO_2} data provide much more reliable constraints on internal ocean processes, in particular in ocean areas closer to land. Combining the two data streams is therefore both possible and beneficial.

Such a combination can be implemented by adding together the cost function contributions of p^{CO_2} data (from **SFC**) and of atmospheric CO_2 (from **ATM**). The estimated fields in this combined case (not shown) are almost identical to the **SFC** run, as expected from the weakness of the atmospheric constraint on oceanic fluxes. However, as the atmospheric data constrain the sum of ocean and land fluxes, improvements of the estimated land fluxes can be expected. Of course, the combined run is essentially equivalent to using the sea-air fluxes of the **SFC** run as priors in a subsequent

OSD
9, 2273–2326, 2012

"classical" atmospheric inversion (similar as in the joint inversion by Jacobson et al., 2007).

4.2 Short-term variations

In addition to the seasonal variations considered so far, surface-ocean biogeochemical quantities also vary on faster time scales (synoptic, day-to-day) in response to variations in wind speed (e.g., accelerated depletion of an over-saturated mixed layer in a high wind speed event, Bates et al. (1997)) or temperature (changes in solubility and chemical equilibrium) as well as due to short-term variations in biology and upwelling. The results of the diagnostic scheme also include such short-term variations, as illustrated in Fig. 8 (top, blue line). For the location and time period chosen in this figure, a dense time series of observations exists for comparison (orange). Despite the vertical offset between observed and fitted time series, they agree reasonably well not only in terms of seasonal cycle but also in many of the short-term patterns (especially in winter and spring, less so in summer).

Importantly, even though the observed time series in Fig. 8 is part of SOCAT and has thus been used in the fit, the diagnostic scheme cannot respond to its short-term variations because we are only adjusting smooth seasonal degrees of freedom (Sect. 2.4). Rather, the short-term variations in the results are only brought in by the process parametrizations (Appendix A1), both to the a-priori fields and the adjustment during the fit, as illustrated by the a-priori time series in Fig. 8 (bottom). Thus, the level of agreement of short-term variations between the results and the observations (Fig. 8, top) can be regarded as an independent validation of the short-term behaviour of our parametrizations (note that misfits may partly also result from scale mismatch between the model pixel ($\approx 4^\circ$ lat. $\times 5^\circ$ long.) and the point measurement)[1].

[1] In addition to the short-term variations not resolved by our degrees of freedom, misfits are also seen in the resolved features (seasonal cycle and long-term average), which may represent missing interannual variations, or result from the rigidity due to the a-priori spatial

OSD
9, 2273–2326, 2012

Though the presented validation of short-term variations is far from comprehensive, it indicates that our results reproduce at least part of the real short-term variations. This applies both to the **SFC** and **ATM** runs, as they use the same parametrizations. Of course, short-term variations related to ocean-internal processes, such as upwelling events or algal blooms, are neither parametrized nor constrained and thus at best partly contained in the resolved slowlier time scales (though suitable parametrizations could potentially be added, see Sect. 4.5 below).

4.3 Prospects: interannual variability

In the present implementation of the diagnostic scheme, ocean-internal sources/sinks (f_{int}) are purely seasonal (Sect. 2.4). Nevertheless, analogously to the short-term variations, the results also contain interannual variations as simulated by the parametrizations, mainly from temperature-induced variations in carbonate chemistry (factor β). In some areas, this already explains much of the real variability, e.g., in the oligotrophic gyres like at the BATS station, as seen from the match of interannual variations in Fig. 7.

In other regions (e.g., the El Niño driven variations in the Tropical Pacific), however, the temperature-related variability of the prior is actually opposite to the true interannual variability, which is dominated by variability in marine biology and transport that overcompensate the temperature-related variability. This full variability can be retrieved from the SOCAT data base by adding non-seasonal degrees of freedom to the unknown f_{int} (note that most of the driver data sets – wind speed, SST, etc. – comprise all modes of variability). This will be presented in a subsequent step (Rödenbeck et al., 2012).

correlations (Sect. 2.4) which forces the fit to compromise between potentially contradicting constraints at neighbouring locations (see the Supplement, Sect. S5, for a more detailed consideration of misfits). (Errors in the long-term average of $p_{\mathrm{m}}^{\mathrm{CO_2}}$ may also create errors in the direction of the sea-air flux [and thus unrealistic rise or fall of mixed-layer carbon], if it changes the sign of $p_{\mathrm{m}}^{\mathrm{CO_2}} - p_{\mathrm{a}}^{\mathrm{CO_2}}$; this is however not the case at the example location of Fig. 8.)

OSD
9, 2273–2326, 2012

Potential alternative ways to retrieve interannual information include: (1) the diagnostic scheme may be extended to make use of further data streams. An envisaged example is atmospheric oxygen measurements (starting in 1991), already used in Rödenbeck et al. (2008) to estimate interannual variations in oceanic biogeochemical behaviour, or surface-ocean oxygen data by the ARGO project (starting in 2000, http://www.argo.ucsd.edu). (2) The diagnostic scheme may be extended by a relation between $f_{\mathrm{int}}^{\mathrm{DIC}}$ and suitable driving variables, as done in regression methods (see Sect. 4.5 below).

4.4 The ocean-internal carbon sources and sinks

Considering ocean-internal processes, mixed-layer carbon addition/removal occurs both through biological respiration/photosynthesis in the mixed layer and through mixing-in of water masses with higher/lower carbon concentration. In the second case, in turn, there is a contribution from the pre-formed carbon concentration originating from the last contact of the arriving water parcel with the atmosphere, and a contribution from carbon added to the water parcel along its way through the ocean interior by remineralization.

Both mixed-layer and ocean-interior biological sources and sinks of carbon are linked to sources and sinks of nutrients. If we assume Redfield proportions between these biological carbon and nutrient changes, and further assume the hypothetical case that seasonal variations in the differences between the pre-formed carbon and nutrient concentrations of arriving water parcels and those of the destination mixed-layer are in Redfield ratios as well[2], then seasonal ocean-internal nutrient sources and sinks ($f_{\mathrm{int}}^{\mathrm{PO_4}}$ and $f_{\mathrm{int}}^{\mathrm{NO_3}}$) are proportional to seasonal $f_{\mathrm{int}}^{\mathrm{DIC}}$. We calculated the potential seasonal

[2] This is equivalent to assuming that the internal sources/sinks of the tracer $C_{\mathrm{m}}^{\mathrm{DIC}^*} = C_{\mathrm{m}}^{\mathrm{DIC}} - r_{\mathrm{C:P}} C_{\mathrm{m}}^{\mathrm{PO_4}}$ (Gloor et al., 2003) have no seasonality.

OSD
9, 2273–2326, 2012

changes in the mixed-layer PO_4 concentration using a PO_4 budget equation[3]. The resulting seasonal changes in the mixed-layer phosphate concentration were compared to an observation-based monthly climatology (World Ocean Atlas, Garcia et al. (2006)).

Despite the assumption of purely Redfieldian seasonality of sources/sinks, the inferred PO_4 concentration seasonality is broadly similar in phase and amplitude to the observations (Fig. 9). Overestimation is seen especially in the temperate latitudes, indicating that the issing non-Redfieldian sources/sinks cannot be neglected here. Differences between predicted and observed $C_m^{PO_4}$ seasonality also arise from (1) errors in the diagnostic model and its driver data sets (note that, unlike the p^{CO_2} field, $C_m^{PO_4}$ is sensitive to further model details, see Supplement, Sect. S2), and (2) uncertainties in the $C_m^{PO_4}$ climatology[4].

The PO_4 comparison indicates that the scheme's unknowns (f_{int}^{DIC}) are not only a mathematical device in the fit but also meaningful in terms of ocean-internal processes. However, individual processes cannot be distinguished based on carbon data alone; this rather requires the simultaneous use of further data like nutrients, and parametrizations relating them to f_{int}^{DIC}.

[3] The PO_4 budget equation is as Eq. (A16) for DIC (including a history flux), except (1) that the ocean-internal flux was reduced by the Redfield ratio $r_{C:P} = 106$ ($f_{int}^{PO_4} = f_{int}^{DIC}/r_{C:P}$), (2) that the long-term mean was subtracted from $f_{int}^{PO_4}$, and (3) that there was no sea-air flux of phosphate ($f_{ma}^{PO_4} = 0$).

[4] Gouretski (2000) reports systematic data uncertainties on the order of 0.04 µmol/kg, partly due to missing standards. Additional uncertainties in the climatology are expected due to data sparseness in many regions. Month-to-month scatter may partially reflect uncertainties.

The seasonality of nitrate as an alternative nutrient ($C_m^{NO_3}/r_{N:P}$ according to WOA climatology, thin line in Fig. 9) is similar but tends to be smaller than $C_m^{PO_4}$. NO_3 may be advantageous from an observational point of view as it are less prone to, e.g., storage issues than PO_4. However, the NO_3 concentration is subject to additional processes (nitrogen fixation and denitrification).

OSD
9, 2273–2326, 2012

4.5 The diagnostic mixed-layer scheme in the context of other $p_m^{CO_2}$ mapping methods

The suitability of individual $p_m^{CO_2}$ mapping methods (Sect. 1) depends on the intended application. The primary motivation of the diagnostic mixed-layer scheme proposed here was to be able to process different data streams (so far p^{CO_2} and atmospheric CO_2 mixing ratio) in a consistent way, and to combine them. The intention was to achieve this in a data-dominated way, with the least complex model possible.

Statistical methods like the reference $p_m^{CO_2}$ climatology by Takahashi et al. (2009) or the interpolation by Jones et al. (2012b) have the advantage to be data-dominated, quite independent of driver data sets and process parametrizations. Regression methods (multilinear regression, neural networks) use driver data sets to bring in information about modes of variability not resolved by the data, thus addressing the problem of data sparsity; still they leave flexibility in the way how $p_m^{CO_2}$ and these driving fields are related. Data assimilation into process models then also prescribe this relation analytically, which brings in process knowledge as a further constraint, with the risk of suppressing real variability not captured in the chosen parametrizations.

The diagnostic scheme is similar to the regression methods, thus reproducing some of the short-term (Fig. 8) or interannual (Fig. 7) variability. It is more rigid, like data assimilation, for some processes that may be considered relatively well known (sea-air gas exchange, carbonate chemistry, and mixed-layer mass conservation). It is more free, on the other hand, for the ocean-internal sources and sinks (independent of any assumptions except for smoothness as needed to regularize the mathematics), i.e., it is purely statistical here. The smoothness assumption represents the mechanism of spatio-temporal interpolation[5].

[5] This results in different notions of output resolution: formally, the resolution is daily and $\approx 4^\circ \times 5^\circ$ pixels, but part of this fine-scale variability is only coming from the driving data, while only the larger scales (mean seasonal cycle and > 1500 km areas) are actually informed by the SOCAT data.

OSD
9, 2273–2326, 2012

An advantage of explicit parametrizations, as in the proposed mixed-layer scheme, is that the unknowns (here: internal sources/sinks) are physical quantities which are open for interpretation (cmp. Sect. 4.4). This also allows to use measurements of further quantities of the scheme (such as mixed-layer DIC concentration, nutrients, etc.) as data constraints as well. These further data streams can either be used instead of the present ones ($p_{\mathrm{m}}^{\mathrm{CO_2}}$ or atmospheric data) to create alternative estimates for comparison, or in conjunction with the present ones to potentially distinguish more processes (e.g., carbonate chemistry).

All carbon sources/sinks in the ocean and on land are linked to oxygen sinks/sources, most of them in rather well-defined stoichiometric ratios. The diagnostic scheme can therefore be extended to also make use of atmospheric or oceanic measurements of oxygen (Rödenbeck et al., 2012).

Beyond the scheme as presented here, multilinear regression could be brought in by expressing the unknown ocean-internal fluxes as linear combination of (further) driving variables and then match the data by adjusting time-independent weights. As in the regression methods, this may be an alternative way to add short-term and interannual variability (Sects. 4.2–4.3), at the expense of being less data-driven. Instead of linear combinations, also parametrizations of the ocean-internal processes involving adjustable parameters could be used if available; this would also open up the way to use even further data streams, such as ocean color from satellite observation.

5 Summary and conclusions

We considered two independent data streams to constrain surface ocean biogeochemistry: (1) surface-ocean CO_2 partial pressure, and (2) atmospheric CO_2 mixing ratios. As a tool to relate these data streams to each other and to the sea-air CO_2 exchange, we developed a diagnostic model comprising simple representations of sea-air gas exchange, tracer solubility, carbonate chemistry, mixed-layer tracer budgets, and

OSD
9, 2273–2326, 2012

seasonal re-entrainment (history). Focusing on seasonal variations, the following conclusions were drawn:

- The diagnostic scheme can be robustly fit to SOCAT p^{CO_2} data, to estimate the spatio-temporal field of p^{CO_2} and sea-air CO_2 exchange.
- In terms of seasonality, the resulting p^{CO_2} field agrees well with the climatology by Takahashi et al. (2009), with most differences related to coastal areas and marginal seas, especially where data density is low. The results of our diagnostic scheme additionally contain short-term and – to some extent – interannual variations.
- The atmospheric CO_2 constraint on surface ocean biogeochemistry (run **ATM**) is not robust in general, but consistent with the oceanic constraint (run **SFC**) in areas away from land influences. This confirms that it is appropriate and advisable to use the oceanic constraint as oceanic prior in an atmospheric inversion, to improve the inferred land fluxes.
- The diagnostic scheme involves interpretable quantities. It can be extended by parametrizations of ocean-internal sources and sinks, that allow to use further data streams to constrain individual biogeochemical processes.
- Using synthetic data tests (as in Appendix B), the diagnostic mixed-layer model can also be applied to assess possible impacts of additional data to constrain ocean biogeochemistry, to help designing future observation strategies.

Appendix A

Model documentation

For reference, all mathematical symbols are listed in Table 2 and Table 3.

OSD
9, 2273–2326, 2012

A1 Ocean process parametrizations

The variables of the following parametrizations are spatio-temporal fields, with a temporal resolution of 1 day and a horizontal resolution of ≈4°×5°, i.e., the equations are applied to every pixel of the atmospheric transport model. Vertically, we consider an oceanic mixed layer assumed to be perfectly homogeneous in temperature and chemical composition, exchanging inorganic carbon with the overlying atmosphere, the underlying ocean interior, neighbouring pixels, and the organic carbon pool.

A1.1 Solubility and gas exchange

Sea-air flux is expressed in the usual way in terms of a partial pressure difference,

$$f_{\mathrm{ma}}^{\mathrm{CO_2}} = k^{\mathrm{CO_2}} \varrho L^{\mathrm{CO_2}} \left(p_{\mathrm{m}}^{\mathrm{CO_2}} - p_{\mathrm{a}}^{\mathrm{CO_2}} \right) \tag{A1}$$

(atmospheric sign convention: positive=source to atmosphere) where $p_{\mathrm{a}}^{\mathrm{CO_2}}$ is the CO_2 partial pressure in the atmosphere, and $p_{\mathrm{m}}^{\mathrm{CO_2}}$ the CO_2 partial pressure in the homogeneous mixed layer[6]. The temperature-dependent solubility $L^{\mathrm{CO_2}}$ relates the partial pressure difference to a dissolved CO_2 concentration difference. The seawater density ϱ is needed to relate volume-based and mass-based quantities. The diffusive resistance to sea-air gas exchange is described by the piston velocity $k^{\mathrm{CO_2}}$. We use the formulation according to Wanninkhof (1992)

$$k^{\mathrm{CO_2}} = \Gamma u^2 \left(Sc^{\mathrm{CO_2}} / Sc^{\mathrm{Ref}} \right)^{-0.5} \tag{A2}$$

calculated from 6-hourly NCEP wind speed u (Kalnay et al., 1996) and Schmidt number Sc given in Supplement Sect. S1.1. The global scaling factor Γ is chosen such that

[6] More precisely, it is the CO_2 partial pressure in air in equilibrium with sea water of the present chemical composition.

OSD
9, 2273–2326, 2012

the global mean CO_2 piston velocity $k_{Glob}^{CO_2}$ matches the value from Table 3 given by Naegler (2009) (average of values by Naegler and Levin (2006), Krakauer et al. (2006), Sweeney et al. (2007), and Müller et al. (2008) inferred from radiocarbon data).

The atmospheric partial pressure $p_a^{CO_2}$ is, via Eq. (S1.1), proportional to the atmospheric CO_2 molar mixing ratio X^{CO_2} just above the ocean surface, which has considerable increasing trend, seasonality, and –even over the ocean– synoptic variability. It is taken from the atmospheric CO_2 mixing ratio fields provided by the Jena inversion *s81_v3.4* (see http://www.bgc-jena.mpg.de/~christian.roedenbeck/download-CO2/). This X^{CO_2} field has been obtained from a forward run of the atmospheric transport model (TM3, resolution $\approx 4^\circ \times 5^\circ \times 19$ layers), with surface fluxes from a classical atmospheric transport inversion based on atmospheric CO_2 measurements. The X^{CO_2} field constructed that way is compatible with measured CO_2 mixing ratios at the sites used, and therefore realistic at least in the rising trend and large-scale seasonality; the sensitivity of the results presented here to the X^{CO_2} field is small anyway, see Supplement, Sect. S2. (Of course, the use of inversion-based X^{CO_2} already creates some dependence of our prior from atmospheric data. This formally violates assumptions of Bayesian inference, which however could only affect calculations of a-posteriori covariances. Even there, we would not expect problems due to the small sensitivity.)

A1.2 Carbonate chemistry

The CO_2 partial pressure in the mixed layer is a function of mixed-layer DIC concentration C_m^{DIC}, alkalinity A, temperature T, and salinity S,

$$p_m^{CO_2} = p_m^{CO_2}(C_m^{DIC}, A, T, S) \tag{A3}$$

At first, we single out the temperature dependence into a temperature-dependent factor and $p_m^{CO_2}$ at a reference temperature T^{LT},

$$p_m^{CO_2} = \beta(T, T^{LT}) \cdot p_m^{CO_2}(C_m^{DIC},\ A,\ T^{LT},\ S) \tag{A4}$$

OSD
9, 2273–2326, 2012

For iso-chemical sea water (i.e., constant C_m^{DIC}, A, and S), Takahashi et al. (2009) experimentally determined the ratio of $p_m^{CO_2}$ of samples at different temperatures to be

$$\beta(T_1,T_2) = \exp\left\{0.0433\,K^{-1}\,(T_1 - T_2)\right. \tag{A5}$$

$$\left. -4.35\times10^{-5}\,K^{-2}\left(T_1^2 - T_2^2\right)\right\}$$

Even though this relation has been fitted to data in the temperature range about 2 to 25°C, it is used here for all temperatures. The actual temperature T is taken from the daily OAFlux product (Table 1). As reference temperature, we use its long-term mean T^{LT} over the inversion period.

The remaining dependences of $p_m^{CO_2}$ are non-linear, but monotonic and can be linearized to good approximation (Sarmiento and Gruber, 2006). Linearization of the dependence on C_m^{DIC} is important here in order to be able to use the fast minimization algorithm of Rödenbeck (2005) (Sect. A2.1). We linearize around the long-term averages (superscript LT)

$$p_m^{CO_2,LT} = p_m^{CO_2}(C_m^{DIC,LT} + C_m^{DIC,\Delta LT}, A^{LT}, T^{LT}, S^{LT}) \tag{A6}$$

where the additional term $C_m^{DIC,\Delta LT}$ takes into account that the chosen reference value $p_m^{CO_2,LT}$ (mean of data-based climatology by Takahashi et al. (2009)) may not be consistent with the chosen $C_m^{DIC,LT}$ (see Table 1), A^{LT}, or T^{LT} (alternatively, mutually consistent reference values could be calculated from full carbonate-chemistry formulas, which would however require solving a 5th order system of equations, Zeebe and Wolf-Gladrow, 2001). This gives

OSD
9, 2273–2326, 2012

$$p_{\mathrm{m}}^{\mathrm{CO_2}} = \beta(T, T^{\mathrm{LT}}) \times \tag{A7}$$

$$\left[p_{\mathrm{m}}^{\mathrm{CO_2,LT}} + \frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}} \left(C_{\mathrm{m}}^{\mathrm{DIC}} - C_{\mathrm{m}}^{\mathrm{DIC,LT}} - C_{\mathrm{m}}^{\mathrm{DIC,\Delta LT}} \right) \right.$$

$$\left. + \frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial A} \left(A - A^{\mathrm{LT}} \right) + \frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial S} \left(S - S^{\mathrm{LT}} \right) \right]$$

The DIC sensitivity is related to the Revelle factor R by

$$\frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}} = \frac{p_{\mathrm{m}}^{\mathrm{CO_2,LT}}}{C_{\mathrm{m}}^{\mathrm{DIC,LT}}} R \tag{A8}$$

Revelle factors are based on the response factors γ_{DIC} from Egleston et al. (2010),

$$R = \frac{C_{\mathrm{m}}^{\mathrm{DIC,LT}}}{\gamma_{\mathrm{DIC}}} \tag{A9}$$

The alkalinity sensitivity has been calculated using the approximation (Eq. (8.3.17) of Sarmiento and Gruber, 2006) (note missing minus sign),

$$\frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial A} = \frac{p_{\mathrm{m}}^{\mathrm{CO_2,LT}}}{A^{\mathrm{LT}}} \frac{-(A^{\mathrm{LT}})^2}{(2C_{\mathrm{m}}^{\mathrm{DIC,LT}} - A^{\mathrm{LT}})(A^{\mathrm{LT}} - C_{\mathrm{m}}^{\mathrm{DIC,LT}})} \tag{A10}$$

The salinity sensitivity (direct effect on chemical equilibrium only) has been set to

$$\frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial S} = \frac{p_{\mathrm{m}}^{\mathrm{CO_2,LT}}}{S^{\mathrm{LT}}} \tag{A11}$$

OSD
9, 2273–2326, 2012

(Eq. (8.3.5) of Sarmiento and Gruber, 2006). Any seasonal and interannual variations in the proportionalities $\partial p_m^{CO_2}/\partial C_m^{DIC}$ etc. are neglected. The variations in alkalinity A and salinity S are fixed according to input data sets (Table 1). Their effects on $p_m^{CO_2}$ is illustrated in the Supplement, Fig. S7.2. Note that the effect of the alkalinity variations, which are strongly related to freshwater fluxes, is mainly offset by the freshwater effects on DIC considered below (Eq. A17).

In summary, Eq. (A7) provides a linear link between $p_m^{CO_2}$ and C_m^{DIC} according to carbonate chemistry. To simplify the formulas, we define an apparent DIC concentration in equilibrium with the atmosphere,

$$C_a^{DIC} = C_m^{DIC,LT} + \left(\frac{\partial p_m^{CO_2}}{\partial C_m^{DIC}}\right)^{-1}\left[\frac{p_a^{CO_2}}{\beta(T,T^{LT})} - p_m^{CO_2,LT} - \frac{\partial p_m^{CO_2}}{\partial A}\left(A - A^{LT}\right) - \frac{\partial p_m^{CO_2}}{\partial S}\left(S - S^{LT}\right)\right] \quad \text{(A12)}$$

Using this definition, Eq. (A7) becomes

$$p_m^{CO_2} = p_a^{CO_2} + \beta(T,T^{LT})\frac{\partial p_m^{CO_2}}{\partial C_m^{DIC}}\left(C_m^{DIC} - C_a^{DIC} - C_m^{DIC,\Delta LT}\right) \quad \text{(A13)}$$

Now also defining an "apparent piston velocity" of DIC as

$$k^{DIC} = k^{CO_2} \cdot L^{CO_2}\beta(T,T^{LT})\frac{\partial p_m^{CO_2}}{\partial C_m^{DIC}} \quad \text{(A14)}$$

OSD
9, 2273–2326, 2012

Eqs. (A1) and (A13) can be combined and simplified into

$$f_{\mathrm{ma}}^{\mathrm{CO_2}} = k^{\mathrm{DIC}} \varrho \left(C_{\mathrm{m}}^{\mathrm{DIC}} - C_{\mathrm{a}}^{\mathrm{DIC}} - C_{\mathrm{m}}^{\mathrm{DIC,\Delta LT}} \right) \tag{A15}$$

expressing the sea-air CO_2 flux as a linear function of the surface DIC concentration.

A1.3 Mixed-layer DIC budget

Changes in the spatio-temporal field of dissolved tracer in the mixed layer need to be balanced by the sum of fluxes,

$$\frac{\mathrm{d}}{\mathrm{d}t} C_{\mathrm{m}}^{\mathrm{DIC}} = \frac{1}{h\varrho} \left(-f_{\mathrm{ma}}^{\mathrm{CO_2}} + f_{\mathrm{hist}}^{\mathrm{DIC}} + f_{\mathrm{frw}}^{\mathrm{DIC}} + f_{\mathrm{int}}^{\mathrm{DIC}} \right) \tag{A16}$$

Fluxes (tracer exchange per unit horizontal area and unit time) comprise loss through sea-air exchange $f_{\mathrm{ma}}^{\mathrm{CO_2}}$, seasonal re-entrainment $f_{\mathrm{hist}}^{\mathrm{DIC}}$ and freshwater effects $f_{\mathrm{frw}}^{\mathrm{DIC}}$ explained below, and fluxes $f_{\mathrm{int}}^{\mathrm{DIC}}$ due to all other ocean-internal process (biological conversion, vertical advection/diffusion, or horizontal advection/diffusion). This formulation assumes that fluxes are immediately diluted vertically throughout the mixed layer. Mixed-layer depth h is prescribed as a climatology by de Boyer Montégut et al. (2004) (downloaded from "http://www.locean-ipsl.upmc.fr/~cdblod/mld.html" using the MLD climatology "`mldT02_sk.nc`" according to a temperature criterion; chosen because the MLD climatology based on a density criterion has incomplete coverage, and there is no data-based MLD product also representing interannual variations). The density ϱ of surface water (assumed constant, Table 3) is needed as $C_{\mathrm{m}}^{\mathrm{DIC}}$ is mass-based (given in µmol/kg) while fluxes are volume/area-based.

Two ocean-internal processes are parametrized explicitly in the budget Eq. (A16):

– The freshwater effect $f_{\mathrm{frw}}^{\mathrm{DIC}}$ describes the dilution or enhancement of the mixed-layer DIC concentration by precipitation, evaporation, river freshwater input, or

OSD
9, 2273–2326, 2012

sea-ice formation/melting. Assuming that freshwater does not contain either carbon or salt (and that all salinity variations are related to freshwater fluxes), the freshwater effect on DIC is calculated from salinity changes,

$$f_{\mathrm{frw}}^{\mathrm{DIC}} = h\varrho \frac{C_{\mathrm{m}}^{\mathrm{DIC,LT}}}{S^{\mathrm{LT}}} \frac{\mathrm{d}S}{\mathrm{d}t} \tag{A17}$$

– The "history" flux $f_{\mathrm{hist}}^{\mathrm{DIC}}$ singles out part of the tracer exchange between the mixed layer and the underlying ocean, namely the re-entrainment during mixed-layer deepening of concentrations left behind previously during mixed-layer shoaling, if nothing had changed below the mixed layer in the meanwhile. It is therefore proportional to the difference between the current concentration $C_{\mathrm{m}}^{\mathrm{DIC}}(t)$ and that at time t_{prev} when the mixed layer was as deep for the last time:

$$f_{\mathrm{hist}}^{\mathrm{DIC}}(t) = \varrho \cdot \left(C_{\mathrm{m}}^{\mathrm{DIC}}(t_{\mathrm{prev}}) - C_{\mathrm{m}}^{\mathrm{DIC}}(t) \right) \cdot \Theta\left(\frac{\mathrm{d}h}{\mathrm{d}t} \right) \tag{A18}$$

The Heaviside step function Θ ensures that f_{hist} only acts during mixed-layer deepening ($\mathrm{d}h/\mathrm{d}t > 0$). The history flux is non-zero in the long-term mean: the deeper waters that get disconnected from the shoaling mixed layer will preserve the high DIC concentrations of the winter season. A budget equation that does not include the re-entrainment of this amount of DIC during mixed-layer deepening would imply a systematic downward "leak" of DIC. The history flux is thus essential to balance the long-term mean internal sources/sinks $f_{\mathrm{int}}^{\mathrm{DIC}}$ with the mean sea-air exchange $f_{\mathrm{ma}}^{\mathrm{CO_2}}$ (plus the mean DIC accumulation in the mixed layer itself), as derived in Sect. S3.

The explicit treatment of these processes is not strictly needed if the only purpose of the scheme was mapping of $p^{\mathrm{CO_2}}$, because they would otherwise be absorbed into $f_{\mathrm{int}}^{\mathrm{DIC}}$ during the fit to the data. However, splitting them off allows easier interpretation of $f_{\mathrm{int}}^{\mathrm{DIC}}$ (Sect. 4.4), also preparing the envisaged link to oxygen.

OSD
9, 2273–2326, 2012

Substituting Eq. (A15) into Eq. (A16) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}C_{\mathrm{m}}^{\mathrm{DIC}} = -\frac{k^{\mathrm{DIC}}}{h}C_{\mathrm{m}}^{\mathrm{DIC}} + \frac{k^{\mathrm{DIC}}}{h}C_{\mathrm{a}}^{\mathrm{DIC}} + \frac{k^{\mathrm{DIC}}}{h}C_{\mathrm{m}}^{\mathrm{DIC,\Delta LT}} + \frac{1}{h\varrho}f_{\mathrm{hist}}^{\mathrm{DIC}}(C_{\mathrm{m}}^{\mathrm{DIC}}) + \frac{1}{h\varrho}f_{\mathrm{frw}}^{\mathrm{DIC}} + \frac{1}{h\varrho}f_{\mathrm{int}}^{\mathrm{DIC}} \quad \text{(A19)}$$

This represents a 1st order differential equation that can be solved (numerically) for given "forcing" $f_{\mathrm{int}}^{\mathrm{DIC}}$; it thus provides a linear dependence of $C_{\mathrm{m}}^{\mathrm{DIC}}$ on $f_{\mathrm{int}}^{\mathrm{DIC}}$ for every pixel. The history flux $f_{\mathrm{hist}}^{\mathrm{DIC}}$ represents a linear operator of $C_{\mathrm{m}}^{\mathrm{DIC}}$ according to Eq. (A18) and therefore does not destroy the linearity. Nevertheless, we simplify the numerical implementation by not calculating $f_{\mathrm{hist}}^{\mathrm{DIC}}$ for the actual concentration field $C_{\mathrm{m}}^{\mathrm{DIC}}$, but rather always from seasonal $C_{\mathrm{m}}^{\mathrm{DIC}}$ variations inferred from the $p^{\mathrm{CO_2}}$ climatology (Takahashi et al., 2009) via

$$C_{\mathrm{m}}^{\mathrm{DIC}} = C_{\mathrm{m}}^{\mathrm{DIC,LT}} + \left(\frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}}\right)^{-1}\left[\frac{p_{\mathrm{m}}^{\mathrm{CO_2}}}{\beta(T,T^{\mathrm{LT}})} - p_{\mathrm{m}}^{\mathrm{CO_2,LT}} - \frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial A}\left(A - A^{\mathrm{LT}}\right) - \frac{\partial p_{\mathrm{m}}^{\mathrm{CO_2}}}{\partial S}\left(S - S^{\mathrm{LT}}\right)\right] \quad \text{(A20)}$$

(cmp. Eq. (A7)). Proper treatment without this simplification would require to include the history flux into the numerical solver of the differential equation. The approximation introduces errors to the mass balance which however are small.

For the numerical solution of the budget equation and the implementation of fixed (a-priori) and adjustable contributions, we decompose the DIC concentration

$$C_{\mathrm{m}}^{\mathrm{DIC}} = C_{\mathrm{m}}^{\mathrm{DIC,ref}} + C_{\mathrm{m}}^{\mathrm{DIC,\Delta LT}} + C_{\mathrm{m}}^{\mathrm{DIC,\Delta ini}} + C_{\mathrm{m}}^{\mathrm{DIC,\Delta pri}} + C_{\mathrm{m}}^{\mathrm{DIC,\Delta adj}} \quad \text{(A21)}$$

OSD
9, 2273–2326, 2012

into a fixed temporally constant reference concentration $C_m^{DIC,ref}$, the fixed temporally constant correction $C_m^{DIC,\Delta LT}$ of the chemistry linearization points, an adjustable temporally constant deviation $C_m^{DIC,\Delta ini}$ from the reference determining the initial concentration at the start $t = t_i$ of the inversion period, a fixed deviation $C_m^{DIC,\Delta pri}$ in the absence of interior sources/sinks ($f_{int}^{DIC} = 0$) just responding to solubility changes, history and freshwater fluxes, and a deviation $C_m^{DIC,\Delta adj}$ responding to the adjustable ocean-internal fluxes f_{int}^{DIC} and to $C_m^{DIC,\Delta ini}$. Correspondingly, Eq. (A19) can be decomposed into

$$\frac{d}{dt}\left(C_m^{DIC,ref} + C_m^{DIC,\Delta LT} + C_m^{DIC,\Delta ini}\right) = 0 \tag{A22}$$

$$\begin{aligned}\frac{d}{dt}C_m^{DIC,\Delta pri} &= -\frac{k^{DIC}}{h}C_m^{DIC,\Delta pri} \\ &+\frac{k^{DIC}}{h}\left(C_a^{DIC} - C_m^{DIC,ref}\right) \\ &+\frac{1}{h\varrho}f_{hist}^{DIC} + \frac{1}{h\varrho}f_{frw}^{DIC}\end{aligned} \tag{A23}$$

$$\begin{aligned}\frac{d}{dt}C_m^{DIC,\Delta adj} &= -\frac{k^{DIC}}{h}C_m^{DIC,\Delta adj} \\ &-\frac{k^{DIC}}{h}C_m^{DIC,\Delta ini} + \frac{1}{h\varrho}f_{int}^{DIC}\end{aligned} \tag{A24}$$

We take $C_a^{DIC}(t = t_i)$ at the beginning of the time period as reference $C_m^{DIC,ref}$. Eqs. (A23) and (A24) are solved with zero initial condition. (Note that C_m^{DIC} will approach a well-defined solution even for any initial condition after some spin-up time because the coefficient $-k^{DIC}/h$ is negative. Nevertheless, the possibility to change the initial condition through $C_m^{DIC,\Delta ini}$ is needed to avoid distortion of the seasonal cycle by some initial transient.) Equation (A24) provides the sought-after link of C_m^{DIC} to the unknown

OSD
9, 2273–2326, 2012

internal fluxes $f_{\mathrm{int}}^{\mathrm{DIC}}$ (plus a minor contribution from the response to the unknown initial condition $C_{\mathrm{m}}^{\mathrm{DIC},\Delta\mathrm{ini}}$).

Note that the DIC concentration contains the unknown correction $C_{\mathrm{m}}^{\mathrm{DIC},\Delta\mathrm{LT}}$ for the chemistry linearization point which however cancels out both for $p_{\mathrm{m}}^{\mathrm{CO_2}}$ (Eq. A13) and the sea-air flux (Eq. A15). Therefore, $C_{\mathrm{m}}^{\mathrm{DIC},\Delta\mathrm{LT}}$ does not affect the calculation anywhere except for the absolute DIC levels.

A2 Inversion

The basic method of this paper is a linear Bayesian inversion (Newsam and Enting, 1988). The implementation is based specifically on the atmospheric transport inversion ("Jena CO_2 inversion") described in Rödenbeck (2005) and reviewed in Sect. A2.1. This atmospheric transport inversion is extended by representing sea-air fluxes by the ocean parametrizations (Sect. A2.2) and by using the $p^{\mathrm{CO_2}}$ constraint instead of or in addition to the atmospheric constraint (Sect. A2.3).

A2.1 The classical atmospheric transport inversion

The atmospheric inversion is based on a set of observed mixing ratios $\boldsymbol{c}_{\mathrm{obs}}$ (time series at observation sites, see Table 1). The inversion calculation seeks fluxes $\boldsymbol{f}$ that lead to the best match between $\boldsymbol{c}_{\mathrm{obs}}$ and corresponding modelled mixing ratios $\boldsymbol{c}_{\mathrm{mod}}(\boldsymbol{f})$, in the sense that the value of the cost function

$$J_{\mathrm{c}} = \frac{1}{2}(\boldsymbol{c}_{\mathrm{obs}} - \boldsymbol{c}_{\mathrm{mod}})^{\mathrm{T}}\mathbf{Q}_{\mathrm{c}}^{-1}(\boldsymbol{c}_{\mathrm{obs}} - \boldsymbol{c}_{\mathrm{mod}}) \qquad \text{(A25)}$$

is minimal, where the (diagonal) matrix $\mathbf{Q}_{\mathrm{c}}$ introduces a weighting among the concentration values, involving assumed measurement uncertainty, location-dependent model uncertainty, and a data density weighting, as described in Rödenbeck (2005). The modelled mixing ratios, taken at the same time and location as the observations, are

OSD
9, 2273–2326, 2012

simulated by a numerical transport model (here, the TM3 model, Heimann and Körner, 2003), which can formally be written as

$$\boldsymbol{c}_{\mathrm{mod}} = \mathbf{A}\boldsymbol{f} + \mathbf{c}_0 \tag{A26}$$

with the transport matrix $\mathbf{A}$ and the initial concentration $\mathbf{c}_0$.

Since the atmospheric data at the discrete set of sites cannot fully constrain the flux field at all pixels and time steps, Bayesian a-priori constraints are introduced to regularize the estimation. They are implemented here by writing the fluxes as a function of a set of parameters $\boldsymbol{x}$,

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}_{\mathrm{fix}} + \mathbf{F}\boldsymbol{x} \tag{A27}$$

In this linear flux model (Rödenbeck, 2005), each parameter in $\boldsymbol{x}$ acts as a multiplier for one of the columns of the matrix $\mathbf{F}$ to be explained below. A-priori, the parameters $\boldsymbol{x}$ are assumed to have zero mean and unit variance and to be uncorrelated, expressed by adding the Bayesian cost function contribution

$$J = J_{\mathrm{c}} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \tag{A28}$$

This is equivalent to assuming fluxes with a-priori mean $\boldsymbol{f}_{\mathrm{fix}}$ and a-priori covariance matrix $\mathbf{Q}_{\boldsymbol{f},\mathrm{pri}} = \mathbf{F}\mathbf{F}^{\mathrm{T}}$.

The specification of the flux model elements $\boldsymbol{f}_{\mathrm{fix}}$ and $\mathbf{F}$ is detailed in Rödenbeck (2005). The flux is first written as a sum of contributions from several components,

$$\boldsymbol{f} = \boldsymbol{f}_{\mathrm{ma}} + \boldsymbol{f}_{\mathrm{nee}} + \boldsymbol{f}_{\mathrm{fos}} + \boldsymbol{f}_{\mathrm{ini}} \tag{A29}$$

($\boldsymbol{f}_{\mathrm{ma}}$ = sea-air flux, $\boldsymbol{f}_{\mathrm{nee}}$ = terrestrial Net Ecosystem Exchange (NEE), $\boldsymbol{f}_{\mathrm{fos}}$ = fossil fuel burning emissions, $\boldsymbol{f}_{\mathrm{ini}}$ = flux pulse at beginning of inversion period to adjust initial atmospheric mixing ratio field). Each component i is represented by a structure as Eq. (A27), i.e.,

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_i \left(\boldsymbol{f}_{\mathrm{fix},i} + \mathbf{F}_i\boldsymbol{x}_i\right) \tag{A30}$$

OSD
9, 2273–2326, 2012

where the vectors $\boldsymbol{x}_i$ denote subsets of parameters in $\boldsymbol{x}$, and the matrices $\mathbf{F}_i$ denote the corresponding groups of columns of $\mathbf{F}$. The columns of $\mathbf{F}_i$ represent spatio-temporal base functions out of which the spatially and temporally varying adjustments to the flux field $\boldsymbol{f}_i$ are composed. In space, the base functions represent flux elements localized at each of the grid cells of the transport model. These elements overlap each other by exponential tails, acting to smooth the flux field. In time, the base functions represent Fourier modes. Smoothing is achieved by down-weighting the higher-frequency modes. In addition to the a-priori correlations implemented that way, all columns of $\mathbf{F}_i$ are proportional to a spatio-temporal weighting allowing flux adjustments in areas of activity of the component (e.g., over the ocean for $\boldsymbol{f}_{\mathrm{ma}}$) while suppressing flux adjustments elsewhere.

The detailed flux model settings for the components $\boldsymbol{f}_i$ of the surface-to-atmosphere CO_2 flux in the "Jena inversion" are:

- $\boldsymbol{f}_{\mathrm{ma}}$: sea-air flux, being of central interest here. In the pure transport inversion as in (Rödenbeck, 2005), this component is estimated directly by decomposition into flux elements as just described.

- $\boldsymbol{f}_{\mathrm{nee,lt}} + \boldsymbol{f}_{\mathrm{nee,seas}} + \boldsymbol{f}_{\mathrm{nee,var}}$: terrestrial Net Ecosystem Exchange (NEE), further split into long-term flux (superscript "lt"), mean seasonality ("seas"), and non-seasonal variations ("var", interannual and high-frequency variations). The specification of the priors, a-priori sigmas, and correlation structure is taken from the Jena inversion v3.2 (as documented in Rödenbeck (2005) except for a-priori uncertainties tightened by the factor $\sqrt{8}$, and length scales of a-priori spatial correlations increased by a factor 3 in longitude direction and by 1.5 in latitude direction.)

- $\boldsymbol{f}_{\mathrm{fos}}$: fossil fuel burning. Only a fixed (prior) term, as in the Jena inversion v3.2 (yearly emissions from the EDGAR v4.0 data base, update of Olivier and Berdowski, 2001)

- $\boldsymbol{f}_{\mathrm{ini}}$: purely technical component: flux pulse at beginning of inversion period to adjust initial atmospheric mixing ratio field (see Rödenbeck, 2005).

OSD
9, 2273–2326, 2012

The numerical minimization of the cost function is done by Conjugate Gradient descent with re-orthonomalization (Rödenbeck, 2005).

A2.2 Extension: using ocean parametrization

The inversion set-up used in this study is identical to the Jena inversion system as of Sect. A2.1 in most respects, except that the sea-air flux $\boldsymbol{f}_{\mathrm{ma}}$ is not estimated directly but implemented as a function of the ocean-internal flux $\boldsymbol{f}_{\mathrm{int}}$ using the parametrizations of Sect. A1. Then, instead of $\boldsymbol{f}_{\mathrm{ma}}$, the ocean-internal flux $\boldsymbol{f}_{\mathrm{int}}$ is adjusted to match the data constraint. The detailed flux-model specifications for $\boldsymbol{f}_{\mathrm{int}}$ are similar to those for $\boldsymbol{f}_{\mathrm{ma}}$, using the same spatial a-priori correlations (correlation length scale about 1910 km in longitude direction and about 960 km in latitude direction, as the ocean flux in the Jena inversion v3.2, Rödenbeck, 2005).

Temporal correlations are twice as long as in the Jena inversion v3.2, involving frequencies up to about biweekly. Only seasonal degrees of freedom are included, suppressing spurious non-seasonal variations from unequal temporal distribution of SOCAT data in many parts of the globe (see discussion in Sect. 4.3 and Fig. S7.3); this restriction will be relaxed in a subsequent step (Rödenbeck et al., 2012).

As a-priori state we assume to have no information on the internal fluxes, i.e., $f^{\mathrm{DIC}}_{\mathrm{int,fix}} = 0$. Consequently, the prior values of $C^{\mathrm{DIC}}_{\mathrm{m}}$, $p^{\mathrm{CO_2}}$, and the sea-air fluxes $f^{\mathrm{CO_2}}_{\mathrm{ma}}$ arise from solution equilibrium (plus freshwater and history fluxes) according to Eq. (A23). They only depend on the input data sets of T, h, u, $X^{\mathrm{CO_2}}$, p_{baro}, S, A as well as the parameters of Table 3. The "error" of this prior model $f^{\mathrm{DIC}}_{\mathrm{int,fix}} = 0$ is obviously identical with the process fluxes themselves. Therefore, the a-priori uncertainty around this prior should reflect the expected amplitude of variations. In the standard set-up, we make no assumptions about spatial structure of the amplitude, and take a constant per-ocean-area uncertainty over all the ocean. No adjustments are allowed in ice-covered regions (Supplement, Sect. S1.2). The a-priori uncertainties are scaled such that the implied

OSD
9, 2273–2326, 2012

a-priori uncertainty of the globally integrated sea-air flux is 1200 Tmol (with respect to 3-monthly anomalies).

In addition to the main degree of freedom $f_{\mathrm{int}}^{\mathrm{DIC}}$, there are adjustable degrees of freedom in the land fluxes as specified in Sect. A2.1 (run **ATM** only). Further, some degrees of freedom are needed for technical reasons. The initial conditions of the carbon budget equation needs to become consistent with the data, achieved by adjusting $C_{\mathrm{m}}^{\mathrm{DIC},\Delta\mathrm{ini}}$ (spatially resolved, but constant in time, zero a-priori). Additionally, as mentioned in Sect. A2.1, we need degrees of freedom adjusting the atmospheric initial condition for CO_2 (run **ATM** only).

Table 4 summarizes the variables that are adjusted in the cost function minimization.

A2.3 Extension: the p^{CO_2} constraint

Analogously to the observed atmospheric mixing ratios $\boldsymbol{c}_{\mathrm{obs}}$, all selected partial pressure values from the SOCAT data base (Table 1) are collected into a vector $\boldsymbol{p}_{\mathrm{obs}}$. A corresponding vector $\boldsymbol{c}_{\mathrm{mod}}$ of modelled partial pressures is formed by sampling the gridded $p_{\mathrm{m}}^{\mathrm{CO_2}}$ field of the model at the locations and times of the measurements (i.e., the modelled value is that found in the grid cell/time step which encloses the location/time of the corresponding observation). If several SOCAT data points fall into the same grid cell and time step, they are averaged together and only form a single element in $\boldsymbol{p}_{\mathrm{obs}}$ and $\boldsymbol{c}_{\mathrm{mod}}$.

The vector $\boldsymbol{c}_{\mathrm{mod}}$ is a function of the adjustable variable $\boldsymbol{f}_{\mathrm{int}}$, as is $\boldsymbol{c}_{\mathrm{mod}}$ (except that $\boldsymbol{c}_{\mathrm{mod}}$ is also a function of further adjustable variables, $\boldsymbol{f}_{\mathrm{nee}}$ and $\boldsymbol{f}_{\mathrm{ini}}$, which are irrelevant now). We can define an analogous cost function contribution

$$J_{\mathrm{p}} = \frac{1}{2}(\boldsymbol{p}_{\mathrm{obs}} - \boldsymbol{c}_{\mathrm{mod}})^{\mathrm{T}}\mathbf{Q}_{\mathrm{p}}^{-1}(\boldsymbol{p}_{\mathrm{obs}} - \boldsymbol{c}_{\mathrm{mod}}) \tag{A31}$$

The covariance matrix $\mathbf{Q}_{\mathrm{p}}$ is chosen diagonal as well; the uncertainty for every individual pixel with two or more SOCAT data points is set to 10 µatm (for pixels with a single data point to $\sqrt{2} \times 10$ µatm). The fit to the SOCAT data is then done in the same

OSD
9, 2273–2326, 2012

way as to the atmospheric data, just with J_c replaced by J_p. (As some degrees of freedom – land fluxes and atmospheric initial condition – do not depend on $p_m^{CO_2}$, they are unconstrained in this inversion and remain at their a-priori values.)

For the case combining atmospheric and oceanic constraints (Sect. 4.1) both cost function contributions are added ($J = J_c + J_p + \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}$).

Appendix B

Test of retrieval capability

As a prerequisite to interpreting the results of fitting the scheme to surface-ocean p^{CO_2} data, we need to test whether (1) the scheme contains sufficient degrees of freedom to reproduce the spatial and temporal variability of the p^{CO_2} field, and (2) the information available in the SOCAT data set is sufficient to constrain it. This test is done by a synthetic run: (1) we create a spatio-temporal p^{CO_2} field to be used as synthetic "truth" in the following way: the gridded monthly climatology by Takahashi et al. (2009) is mapped onto the grid of our scheme, and interpolated in time (using a spectral filter with a triangular pulse-response of one month half-width) to get smooth temporal variations on our daily time-steps. This seasonal cycle is repeated in each year of the inversion period. Moreover, a trend following the atmospheric CO_2 rise (based on the smoothed interannual variations of $p_a^{CO_2}$) is added. (2) This synthetic p^{CO_2} field is sampled at the locations and times of the SOCAT measurements, in the same way as the modelled p^{CO_2} field (Sect. A2.3). This gives a set of pseudo data representing the same amount of information (in terms of data density available to the scheme) as the actual data. (Note that SOCAT also contains along-track variability on scales much shorter than the size of our grid cells, which the scheme cannot make use of in any case.) (3) We then fit the diagnostic scheme to the pseudo data, and compare the result with the synthetic truth as the known correct answer.

OSD
9, 2273–2326, 2012

In terms of mean seasonality considered here, the retrieval (blue) fits the "known truth" (violet) closely (Fig. 10). Most region-to-region differences are reproduced. The only exception is the North Pacific and the Tropics, in particular the tropical Indian.

To be able to judge the match, we set it into perspective with two further cases representing situations with less or with more available information, respectively:

- The black line shows the Bayesian prior, which does not yet use any SOCAT or atmospheric data (Sect. 2.4). The p^{CO_2} seasonality thus just represents a-priori knowledge on the response to temperature-induced changes in the chemistry (and to lesser extend in solubility, Sect. A2.2). It is opposite in phase compared to the synthetic truth in high latitudes. The pseudo data are able to correct this large difference almost completely.

- We also did the retrieval based on maximum information about the p^{CO_2} field available, by constraining the scheme at every grid point and time step, not only where actual SOCAT data exist. The result (green) fits the known truth almost perfectly, showing that there are sufficient degrees of freedom in the scheme to reproduce the mean seasonality. However, as the performance of the SOCAT-sampled pseudo data is only slightly worse, we conclude that the SOCAT data density is essentially sufficient to constrain the seasonality of surface-ocean biogeochemistry on the scale of the considered regions (note that strong modes of variability existing in reality but not contained in our smooth synthetic "truth" potentially deteriorate the performance of our fit to real data).

The test confirms that the differences between our results and Takahashi et al. (2009) in the Northernmost part of the North Pacific are due to the lack of data in conjunction with spatial variability: in Takahashi et al. (2009), seasonal amplitudes are large in a region at around 50–60° N and strongly drop going Northward (see Supplement, Fig. S7.1). Though the diagnostic model would be able follow the spatial pattern in the climatology if data would say so (green line), in the absence of data it extrapolates the field northward keeping a high seasonal cycle of the pixels South of this area.

OSD
9, 2273–2326, 2012

Seasonal data coverage is also limiting in the the tropical Indian as the other problematic region (Supplement, Fig. S7.4), though it remains open why the the fit is not close even with full information (green).

Supplementary material related to this article is available online at:
http://www.ocean-sci-discuss.net/9/2273/2012/osd-9-2273-2012-supplement.pdf.

Acknowledgements. We would like to thank all contributors to the SOCAT data base, as well as all providers of atmospheric CO_2 data. We gratefully acknowledge helpful discussions with Nicolas Cassar, Andrew Dickson, Niki Gruber, Roberta Hamme, Steve Jones, Bob Key, Armin Köhl, Wolfgang Koeve, Iris Kriest, Corinne Le Quéré, Miguel Mahecha, Bill Munger, Andreas Oschlies, Taro Takahashi, Yasunori Tohjima, and Lisan Yu.

The service charges for this open access publication
have been covered by the Max Planck Society.

References

Archer, D., Kheshgi, H., and Maier-Reimer, E.: Multiple timescales for neutralization of fossil fuel CO_2, Geophys. Res. Lett., 24, 405–408, 1997. 2274

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cy., 20, GB1002, doi:10.1029/2004GB002439, 2006. 2275

Bates, N. R.: Multi-decadal uptake of carbon dioxide into subtropical mode water of the North Atlantic Ocean, Biogeosciences Discuss., 8, 12451–12476, doi:10.5194/bgd-8-12451-2011, 2011. 2283, 2313, 2323

Bates, N. R., Knap, A. H., and Michaels, A. F.: Contribution of hurricanes to local and global estimates of air-sea exchange of CO_2, Nature, 395, 58–61, 1997. 2284

OSD
9, 2273–2326, 2012

Bousquet, P., Peylin, P., Ciais, P., Le Quéré, C., Friedlingstein, P., and Tans, P. P.: Regional changes in carbon dioxide fluxes of land and oceans since 1980, Science, 290, 1342–1346, 2000. 2275

Chen, L., Xu, S., Gao, Z., Chen, H., Zhang, Y., Zhan, J., and Li, W.: Estimation of monthly air-sea CO_2 flux in the southern Atlantic and Indian Ocean using in-situ and remotely sensed data, Remote Sens. Environ., 115, 1935–1941, 2011. 2275

Ciais, P., Tans, P. P., Trolier, M., White, J. W. C., and Francey, R. J.: A Large Northern Hemisphere Terrestrial CO_2 Sink Indicated by the $^{13}C/^{12}C$ Ratio of Atmospheric CO_2, Science, 269, 1098–1102, 1995. 2276

Conkright, M., Locarnini, R. A., Garcia, H., O'Brien, T., Boyer, T., Stephens, C., and Antonov, J.: World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures, CD-ROM Documentation, National Oceanographic Data Center, Silver Spring, MD, 17 pp., 2002. 2313

Conway, T., Tans, P., Waterman, L., Thoning, K., Kitzis, D., Masarie, K., and Zhang, N.: Evidence for interannual variability of the carbon cycle from the national oceanic and atmospheric administration climate monitoring and diagnostics laboratory global air sampling network, J. Geophys. Res., 99, 22831–22855, 1994. 2275, 2313

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology, J. Geophys. Res., 109, C12003, doi:10.1029/2004JC002378, 2004. 2296, 2313

Egleston, E. S., Sabine, C. L., and Morel, F. M. M.: Revelle revisited: Buffer factors that quantify the response of ocean chemistry to changes in DIC and alkalinity, Global Biogeochem. Cy., 24, GB1002, doi:10.1029/2008GB003407, 2010. 2294, 2313

Garcia, H. E., Locarnini, R. A., Boyer, T. P., and Antonov, J. I.: World Ocean Atlas 2005, Volume 4: Nutrients (phosphate, nitrate, silicate)., in: NOAA Atlas NESDIS 64, edited by Levitus, S., US Government Printing Office, Washington, DC, 2006. 2287, 2313, 2325

Gloor, M., Gruber, N., Sarmiento, J., Sabine, C. L., Feely, R. A., and Rödenbeck, C.: A first estimate of present and preindustrial air-sea CO_2 flux patterns based on ocean interior carbon measurements and models, Geophys. Res. Lett., 30, 1010, doi:10.1029/2002GL015594, 2003. 2276, 2286

Gouretski, V.: Systematic errors as the cause for an apparent deep water property variability: global analysis of the WOCE and historical hydrographic data, Prog. Oceanogr., 48, 337–402, http://linkinghub.elsevier.com/retrieve/pii/S0079661100000495, 2000. 2287

OSD
9, 2273–2326, 2012

Heimann, M. and Körner, S.: The global atmospheric tracer model TM3, Tech. Rep. 5, Max Planck Institute for Biogeochemistry, Jena, Germany, 2003. 2277, 2301

Jacobson, A. R., Mikaloff Fletcher, S. E., Gruber, N., Sarmiento, J. L., and Gloor, M.: A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes, Global Biogeochem. Cy., 21, GB1019, doi:10.1029/2005GB002556, 2007. 2276, 2284

Jähne, B., Heinz, G., and Dietrich, W.: Measurement of the diffusion coefficients of sparingly soluble gases in water with a modified Barrer method, J. Geophys. Res., 92, 10767–10776, 1987.

Jones, S. D., Le Quéré, C., and Rödenbeck, C.: Autocorrelation characteristics of surface ocean pCO_2 and air-sea CO_2 fluxes, Global Biogeochem. Cy., accepted, 2012a. 2279

Jones, S. D., Le Quéré, C., Rödenbeck, C., and Manning, A.: A Global Interpolation of Surface Ocean pCO_2, J. Ad. Model. Earth Syst., in review, 2012b. 2275, 2288

Joos, F., Plattner, G.-K., Stocker, T. F., Marchal, O., and Schmittner, A.: Global warming and marine carbon cycle feedbacks an future atmospheric CO_2, Science, 284, 464–467, 1999. 2274

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, 1996. 2278, 2291, 2313

Keeling, R. F. and Shertz, S. R.: Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle, Nature, 358, 723–727, 1992. 2276

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H.: A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), Global Biogeochem. Cy., 18, GB4031, doi:10.1029/2004GB002247, 2004. 2313

Krakauer, N. Y., Randerson, J. T., Primeau, F. W., Gruber, N., and Menemenlis, D.: Carbon isotope evidence for the latitudinal distribution and wind speed dependence of the air-sea gas transfer velocity, Tellus, 58B, 390–417, 2006. 2292

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., and Heimann, M.: Satu-

OSD
9, 2273–2326, 2012

ration of the Southern Ocean CO_2 sink due to recent climate change, Science, 316, 1735–1738, 2007. 2274

Le Quéré, C., Takahashi, T., Buitenhuis, E. T., Rödenbeck, C., and Sutherland, S. C.: Impact of climate change and variability on the global oceanic sink of CO_2, Global Biogeochem. Cycles, 24, GB4007, doi:10.1029/2009GB003599, 2010. 2274

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., Park, G.-H., Wanninkhof, R., Feely, R. A., and Key, R. M.: Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, Geophys. Res. Lett., 33, L19605, doi:10.1029/2006GL027207, 2006. 2313

Lefèvre, N., Watson, A. J., and Watson, A. R.: A comparison of multiple regression and neural network techniques for mapping in situ pCO_2 data, Tellus, 57B, 375–384, 2005. 2275

Matear, R. J. and Hirst, A. C.: Climate change feedback on the future oceanic CO_2 uptake, Tellus B, 51, 722–733, 1999. 2274

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean, Global Biogeochem. Cy., 20, GB2002, doi:10.1029/2005GB002530, 2006. 2276

Müller, S. A., Joos, F., Plattner, G.-K., Edwards, N. R., and Stocker, T. F.: Modelled natural and excess radiocarbon – sensitivities to the gas exchange formulation and ocean transport strength, Global Biogeoch. Cycl., 22, GB3011, doi:10.1029/2007GB003065, 2008. 2292

Naegler, T.: Reconciliation of excess ^{14}C-constrained global CO_2 piston velocity estimates, Tellus, 61B, 372, doi:10.1111/j.1600-0889.2008.00408.x, 2009. 2292, 2315

Naegler, T. and Levin, I.: Closing the global radiocarbon budget 1945–2005, J. Geophys. Res., 111, D12311, doi:10.1029/2005JD006758, 2006. 2292

Newsam, G. N. and Enting, I. G.: Inverse problems in atmospheric constituent studies: I. Determination of surface sources under a diffusive transport approximation, Inverse Problems, 4, 1037–1054, 1988. 2275, 2300

Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in: The Climate System, edited by: Berdowski, J., Guicherit, R., and Heij, B. J., A. A. Balkema Publishers/Swets & Zeitlinger Publishers, Lisse, The Netherlands, ISBN 90 5809 255 0, 33–78, 2001. 2302

OSD
9, 2273–2326, 2012

Park, G.-H., Wanninkhof, R., Doney, S. C., Takahashi, T., Lee, K., Feely, R. A., Sabine, C. L., Triñanes, J., and Lima, I. D.: Variability of global net sea-air CO_2 fluxes over the last three decades using empirical relationships, Tellus B, 62, 352–368, 2010. 2275

Pfeil, B., Olsen, A., Bakker, D., et al.: A uniform, quality controlled, Surface Ocean CO_2 Atlas (SOCAT), Earth System Science Data, in preparation, 2012. 2276, 2278, 2313

Rayner, P., Enting, I., Francey, R., and Langenfelds, R.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}CO_2$ and O_2/N_2 observations, Tellus, 51B, 213–232, 1999. 2275

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Tech. Rep. 6, Max Planck Institute for Biogeochemistry, Jena, Germany, 2005. 2279, 2293, 2300, 2301, 2302, 2303

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, doi:10.5194/acp-3-1919-2003, 2003. 2275, 2276

Rödenbeck, C., Le Quéré, C., Heimann, M., and Keeling, R.: Interannual variability in oceanic biogeochemical processes inferred by inversion of atmospheric O_2/N_2 and CO_2 data, Tellus, 60B, 685, 2008. 2286

Sarmiento, J. L. and Gruber, N.: Ocean Biogeochemical Dynamics, Princeton Univ. Press, 2006. 2293, 2294, 2295

Sarmiento, J. L. and Le Quéré, C.: Oceanic carbon dioxide uptake in a model of century-scale global warming, Science, 274, 1346–1350, 1996. 2274

Sigman, D. M. and Boyle, E. A.: Glacial/interglacial variations in atmospheric carbon dioxide, Nature, 407, 859–869, 2000. 2274

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., and co authors: Constraining global air-sea gas exchange for CO_2 with recent Bomb ^{14}C measurements, Global Biogeochem. Cy., 21, GB2015, doi:10.1029/2006GB002784, 2007. 2292

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Kortzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tillbrook, B., Johannessen, T.and Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R., and de Baar, H. J. W.: Climatological mean and decadal change in surface ocean pCO_2 and net sea-air CO_2 flux over the global oceans, Deep-Sea Res. II, 56, 554–577, 2009. 2274, 2275, 2276, 2281, 2282, 2288, 2290, 2293, 2298, 2305, 2306, 2320, 2321, 2326

OSD
9, 2273–2326, 2012

Takahashi, T., Sutherland, S. C., and Kozyr, A.: Global Ocean Surface Water Partial Pressure of CO_2 Database: Measurements Performed during 1968–2010 (Version 2010), Tech. Rep. NDP-088, ORNL/CDIAC-152, 2010.

Telszewski, M., Chazottes, A., Schuster, U., Watson, A. J., Moulin, C., Bakker, D. C. E., González-Dávila, M., Johannessen, T., Körtzinger, A., Lüger, H., Olsen, A., Omar, A., Padin, X. A., Ríos, A. F., Steinhoff, T., Santana-Casiano, M., Wallace, D. W. R., and Wanninkhof, R.: Estimating the monthly pCO_2 distribution in the North Atlantic using a self-organizing neural network, Biogeosciences, 6, 1405–1421, doi:10.5194/bg-6-1405-2009, 2009. 2275

Valsala, K. V. and Maksyutov, S.: Simulation and assimilation of global ocean pCO_2 and air-sea CO_2 fluxes using ship observations of surface ocean pCO_2 in a simplified Biogeochemical offline model, Tellus, 62B, 821–840, 2010. 2275

Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, J. Geophys. Res.-Oceans, 97, 7373–7382, 1992. 2291

Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas, Marine Chemistry, 2, 203–205, 1974.

Weiss, R. F.: The solubility of nitrogen, oxygen and argon in water and seawater, Deep-Sea Res., 17, 721–735, 1970.

Weiss, R. F. and Price, B. A.: Nitrous oxide solubility in water and seawater, Marine Chemistry, 8, 347–359, 1980.

While, J., Totterdell, I., and Martin, M.: Assimilation of pCO_2 Data Into a Global Coupled Physical-Biogeochemical Ocean Model, J. Geophys. Res., 117, C03037, doi:10.1029/2010JC006815, 2012. 2275

Yu, L. and Weller, R. A.: Objectively Analyzed air-sea heat Fluxes (OAFlux) for the global oceans, B. Am. Meteorol. Soc., 88, 527–539, 2007. 2313

Zeebe and Wolf-Gladrow: CO_2 in Seawater: Equilibrium, Kinetics, Isotopes, Elsevier, 2001. 2293

OSD
9, 2273–2326, 2012

Table 1. Data sets used as constraints or as driver data in the process parametrizations (Sect. A1).

Quantity	Data set	Reference	Pre-treatment / resolution / remarks	Used for
Surface-ocean data:				
$p_m^{CO_2}$	SOCATv1.5	Pfeil et al. (2012)	Data are used having WOCE-flag=2 and valid fields for fugacity, temperature, and salinity. Values below 200 µatm or above 600 µatm have been excluded as being local compared to the grid cells. Values have been transferred from fugacity to partial pressure by dividing by 0.996.	**SFC**
Atmospheric data:				
c^{CO_2}	Various measurement programs	Conway et al. (1994) and many others	CO_2 mixing ratios c^{CO_2}. Data points are individual flask pair averages (approx. weekly) or hourly averages, respectively, taken at 57 sites globally	**ATM**
Driver data:				
T	OAFlux	Yu and Weller (2007)	Gridded, daily	all
ε	OAFlux	Yu and Weller (2007)	Gridded, daily	all
S	WOA 2001	Conkright et al. (2002)	Gridded monthly climatology (interpolated, values taken from A data set)	all
A	CDIAC	Lee et al. (2006)	Gridded monthly climatology (interpolated)	all
γ_{DIC}		Egleston et al. (2010)	mean spatial pattern	all
$C_m^{DIC,LT}$	CDIAC	Key et al. (2004)	mean spatial pattern	all
h	LOCEAN	de Boyer Montégut et al. (2004)	Temperature criterion; Monthly climatology (interpolated)	all
u	NCEP reanalysis	Kalnay et al. (1996)		all
Meteo.	NCEP reanalysis	Kalnay et al. (1996)	Driver for TM3 atmospheric transport model	all
Comparison data:				
C_m^{DIC}	BATS	Bates (2011)	Time series at (32.6° N, 65° W)	Fig. 7
$C_m^{PO_4}$	WOA 2005	Garcia et al. (2006)	Gridded monthly climatology (interpolated)	Fig. 9

Glossary:
BATS: Bermuda Atlantic Time-series Study
CDIAC: Carbon Dioxide Information Analysis Center
LOCEAN: Laboratoire d'océanographie et du climat: expérimentations et approches numériques
NCEP: National Centers for Environmental Prediction
OAFlux: Objectively Analysed air-sea Fluxes
SOCAT: Surface Ocean CO_2 Atlas
WOA: World Ocean Atlas
Math symbols: see Table 2

OSD
9, 2273–2326, 2012

Table 2. Mathematical symbols (also see Table 3). Some quantities are indexed by species or other distinctions explained in the text.

Quantity	Unit[a]	Meaning
$\mathbf{A}$	ppm/(μmol/m^2/s)	Atmospheric transport matrix, Eq. (A26)
C_m^{DIC}	μmol/kg	Mixed-layer DIC concentration
C_a^{DIC}	μmol/kg	DIC concentration in equilibrium with atmosphere
$\boldsymbol{c}_{obs}$	ppm	Set of observed atmospheric mixing ratios
$\boldsymbol{c}_{mod}$	ppm	Modelled mixing ratios, Eq. (A26)
$\mathbf{c}_0$	ppm	Initial mixing ratio
f_{ma}	μmol/m^2/s	Sea-to-air tracer flux, Eq. (A15)
f_{hist}	μmol/m^2/s	History (re-entrainment) flux, Eq. (A18)
f_{int}	μmol/m^2/s	Ocean-internal tracer sources/sinks, Eq. (A16)
$\mathbf{F}$	μmol/m^2/s	Flux model matrix, Eq. (A27)
$\boldsymbol{f}$	μmol/m^2/s	Vector of flux discretized at pixels / time steps
h	m	Mixed layer depth
J	1	Cost function, Eq. (A28)
J_c	1	Cost function, atmospheric data part, Eq. (A25)
J_p	1	Cost function, surface-ocean data part, Eq. (A31)
k	m/s	Piston velocity, Eq. (A2)
L	mol/kg/atm	Solubility (incl. fugacity corr.), Eqs. (A1) and (S1.1)
p_m	μatm	Partial pressure in mixed layer, Sect. A1
p_a	μatm	Atmospheric partial pressure, Eq. (S1.2)
p_{baro}	atm	Barometric pressure at ocean surface
p^{H_2O}	atm	Saturation water vapor pressure above ocean, Eq. (S1.3)
R	1	Revelle factor, Eq. (A8)
S	‰	Mixed-layer salinity[b]
Sc	1	Schmidt number, Sect. S1.1
T	°C	Mixed-layer temperature (Sea Surface Temperature, SST)
t	s	Time coordinate
t_i, t_e	s	Start time, end time of inversion period
u	m/s	Wind speed at 10 m above surface
$\boldsymbol{x}$	1	Vector of adjustable parameters, Eq. (A27)
X	ppm	Dry-air molar mixing ratio in the atmosphere above the ocean
z	m	Vertical coordinate
β	1	Temperature factor of CO_2 partial pressure, Eqs. (A4) and (A5)
Γ	1	Scaling of piston velocities, Eq. (A2)
γ_{DIC}	μmol/kg	Buffer factor (response factor), Eq. (A9)
ε	1	Ice-free fraction of ocean surface (0 = ice-covered, 1 = ice-free)

[a] ppm=μmol/mol

[b] For the present purposes, salinity in ‰ or on the Practical Salinity Scale are considered interchangeable.

OSD
9, 2273–2326, 2012

Table 3. Physical constants, and geophysical quantities assumed constant.

Quantity	Value		Meaning
c_p	3993	J/kg/K	Specific heat capacity of sea water (value at 20˚C)
$k_{Glob}^{CO_2}$	16.5	cm/hr	(= 4.58×10^{-5} m/s) Global mean piston velocity of CO_2 (Naegler, 2009)
R_{gas}	8.3144	J/mol/K	Gas constant
T_0	273.15	K	Absolute temperature at 0˚C
V_{ideal}	0.0224136	m^3/mol	Molar volume of an ideal gas
ϱ	1025	kg/m^3	Density of sea water

OSD
9, 2273–2326, 2012

Table 4. Adjustable degrees of freedom.

Quantity		A-priori value	Spatial resolution	Temporal resolution
Carbon cycle components:				
$f_{\mathrm{int}}^{\mathrm{DIC}}$	ocean-internal DIC sources/sinks	zero	corr. pix.	lt+seas
$f_{\mathrm{nee}}^{\mathrm{CO_2}}$	land NEE (run **ATM** only)		corr. pix.	lt, seas, var
Technical degrees of freedom:				
$C_{\mathrm{m}}^{\mathrm{DIC},\Delta\mathrm{ini}}$	initial condition of budget equation	zero	corr. pix.	lt
$f_{\mathrm{ini}}^{\mathrm{CO_2}}$	flux pulse for atmospheric initial condition (run **ATM** only)	zero	3 lat. bands	lt

Glossary:
lt: constant (long-term)
seas: seasonal
var: interannual and short-term variability
corr. pix.: correlated pixels
Math symbols: see Table 2.

OSD
9, 2273–2326, 2012

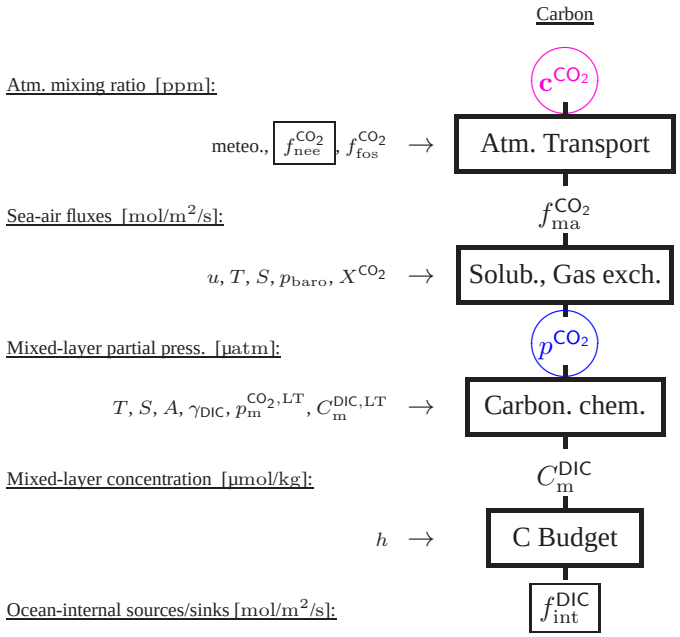

Fig. 1. Summary of the main quantities and process parametrizations involved in the diagnostic mixed-layer model. Thick boxes denote the process parametrizations given in Sect. 2.2, each expressing the quantity right above as a function of the quantity right below. Encircled quantities are those for which observations (Table 1) are used to constrain the scheme. The thin-framed quantities (ocean-internal carbon sources/sinks $f_{\mathrm{int}}^{\mathrm{DIC}}$ and – when fitting to atmospheric data – terrestrial Net Ecosystem Exchange $f_{\mathrm{nee}}^{\mathrm{CO_2}}$) are the main unknowns being adjusted as to satisfy the respective data constraints. Quantities at the arrows on the left represent driver data entering the parametrizations. See Table 2 for mathematical symbols, and Appendix A for further explanation.

OSD
9, 2273–2326, 2012

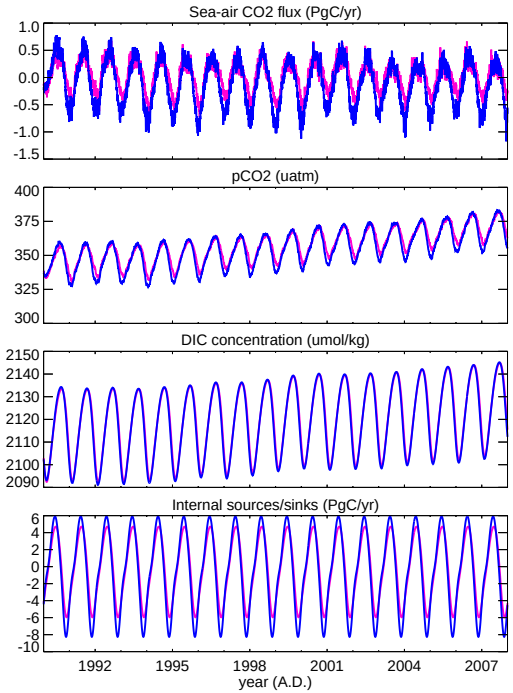

Fig. 2. Example time series of the main quantities of the diagnostic model (Fig. 1): sea-air CO_2 fluxes ($f_{ma}^{CO_2}$), surface-ocean CO_2 partial pressure ($p_m^{CO_2}$), mixed-layer DIC concentration (C_m^{DIC}), as well as ocean-internal carbon addition to (removal from) the mixed layer (f_{int}^{DIC}). Estimates have been obtained by fitting the diagnostic scheme to SOCAT surface-ocean CO_2 partial pressure data (**SFC**, blue) or atmospheric CO_2 mixing ratios (**ATM**, magenta, colors correspond to data source as in Fig. 1). Quantities are shown as integrals/averages over the example region (ocean South of 45° S).

OSD
9, 2273–2326, 2012

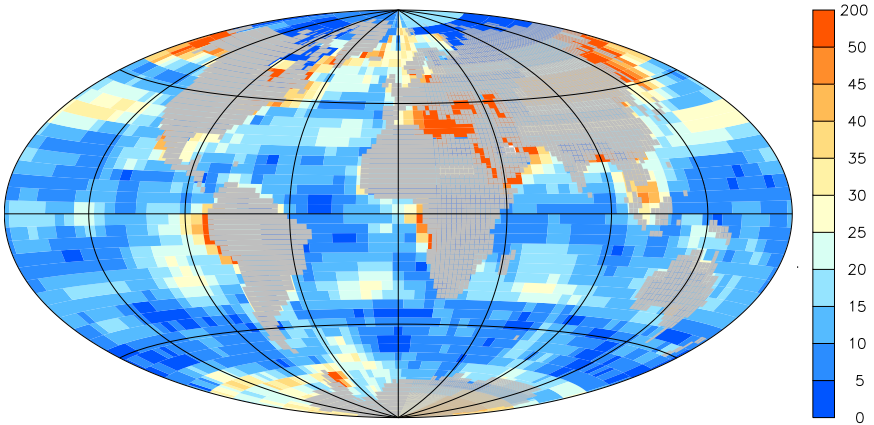

Fig. 3. Amplitude of the mean seasonal cycle of surface-ocean CO_2 partial pressure (µatm) estimated by fitting the diagnostic scheme to the SOCAT data (run **SFC**). The amplitude is given as temporal standard deviation of the monthly mean $p_m^{CO_2}$ at each pixel.

OSD
9, 2273–2326, 2012

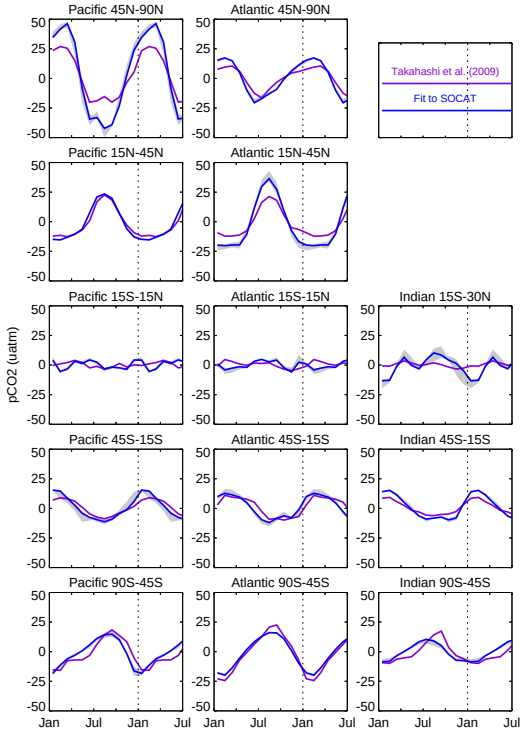

Fig. 4. Monthly mean seasonal cycle (calculated from the full inversion period after de-trending, first half year repeated in the figure, deviation from mean only) of surface-ocean carbon dioxide partial pressure $p_m^{CO_2}$, shown in a regional split of the entire ocean (averages per latitude band and per basin, panels in geographic arrangement). The fit of the diagnostic scheme to SOCAT data (blue) is compared to the $p_m^{CO_2}$ climatology by Takahashi et al. (2009) (violet). The grey shading around the SOCAT-based results comprises sensitivity cases listed in Sect. 2.5.2.

OSD
9, 2273–2326, 2012

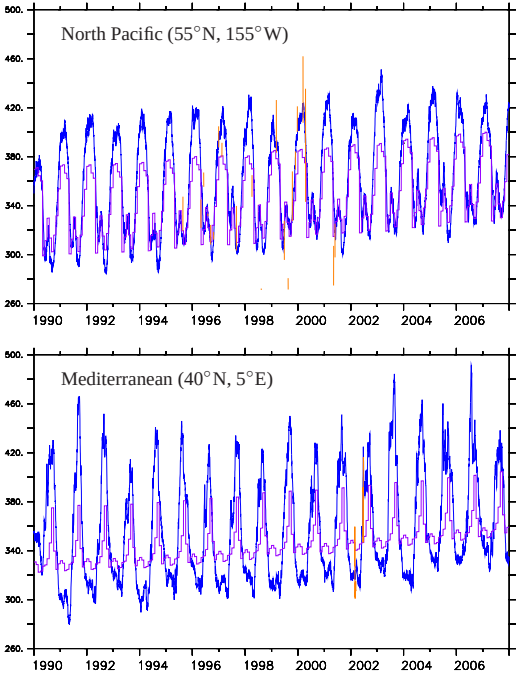

Fig. 5. Time series of mixed-layer p^{CO_2} (µatm) at two test locations. Comparison between the p^{CO_2} field estimated by fitting the SOCAT data (run **OCN**, blue), the monthly climatology by Takahashi et al. (2009) (regridded – in the Mediterranean extrapolated from Atlantic open-ocean values – and with added trend parallel to the atmospheric CO_2 increase, violet), and the measurements from the SOCAT data base contained in the respective pixel (orange).

OSD
9, 2273–2326, 2012

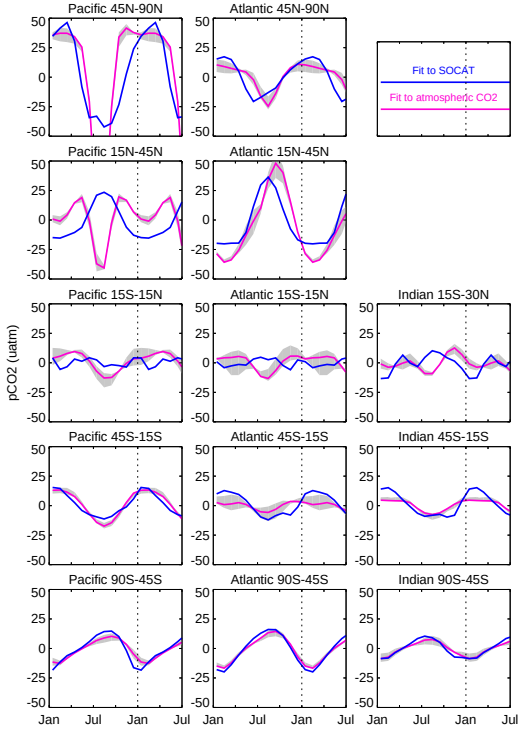

Fig. 6. Monthly mean seasonal cycle of surface-ocean carbon dioxide partial pressure $p_m^{CO_2}$ averaged over latitude bands and basin, as in Fig. 4. The diagnostic scheme has been fitted to surface-ocean p^{CO_2} data (blue, same as Fig. 4) or atmospheric CO_2 mixing ratios (magenta, results partially off-scale). The grey band indicates the sensitivity of the fit to atmospheric CO_2 data with respect to the strength of the Bayesian prior (a-priori uncertainty range decreased or increased by a factor of 2).

OSD
9, 2273–2326, 2012

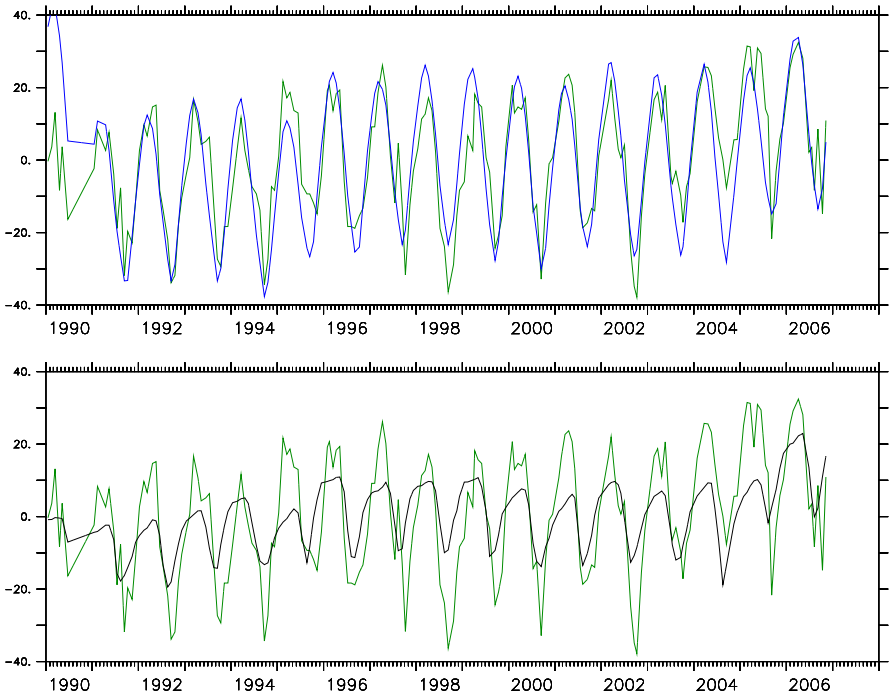

Fig. 7. Time series of mixed-layer DIC concentration C_m^{DIC} (µmol/kg, mean subtracted) at the ocean station BATS (Bermuda Atlantic Time-series Study, 32.6° N,65° W). Top: comparison between the measurements (Bates, 2011, dark green) and the DIC estimate based on SOCAT data (blue). The estimated field has been picked at times when DIC measurements exist; data points have been connected by straight lines for clarity. The first year is contaminated by a spin-up transient. Bottom: same measurements (dark green) compared to the a-priori prediction (black) of the diagnostic scheme before the SOCAT data constraint has been used (assuming no internal sources/sinks, but only response of mixed-layer DIC to changes in solubility and chemical equilibrium, mostly driven by temperature).

OSD
9, 2273–2326, 2012

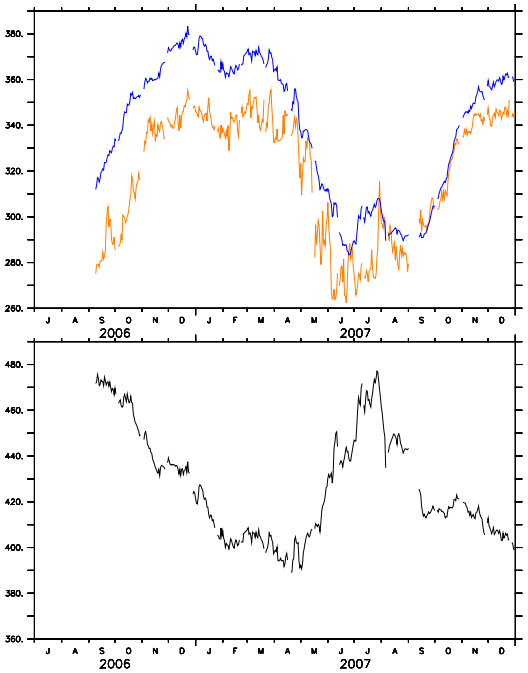

Fig. 8. Time series of mixed-layer p^{CO_2} (µatm) at the ocean station M (66.0° N, 2° E). Top: comparison between the measurements contained in the SOCAT data base (orange) and the p^{CO_2} field estimated by fitting to all SOCAT data (blue). The estimated field has been picked at the times where measurements exist; only 1.5 yr of intensive observation have been plotted. Bottom: a-priori p^{CO_2} field before the SOCAT data constraint has been used (assuming no internal sources/sinks, but only response of mixed-layer DIC to changes in solubility and chemistry induced by temperature and other driving variables). Note shifted vertical scale.

OSD
9, 2273–2326, 2012

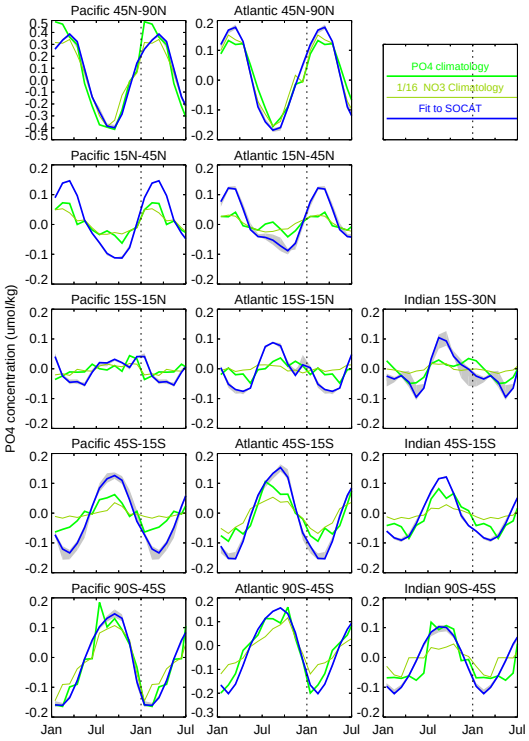

Fig. 9. Mixed-layer phosphate concentration (monthly climatology, mean subtracted). Seasonality inferred from the SOCAT fit assumes all ocean-internal carbon sources and sinks are related to phosphate sources and sinks in Redfield proportions. For comparison, the PO_4 climatology from the World Ocean Atlas (WOA, Garcia et al., 2006) is given, as well as the WOA NO_3 climatology scaled in Redfield proportions. The grey uncertainty band around the SOCAT-based estimates again gives the range of sensitivity cases as in Fig. 4; note however that calculated PO_4 is also sensitive to several further model elements not included in this range.

OSD
9, 2273–2326, 2012

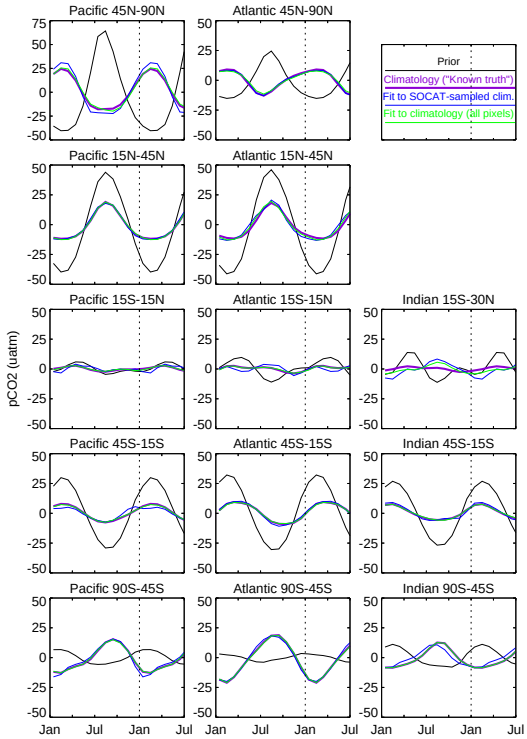

Fig. 10. Testing the capacity to retrieve the p^{CO_2} field from the SOCAT data set: fit of the diagnostic scheme to Takahashi et al. (2009)'s p^{CO_2} climatology (smoothed) subsampled at locations/times of the SOCAT measurements (blue), compared to the (smoothed) p^{CO_2} climatology itself ("known truth", violet). For further comparison, a fit of the scheme to the climatology at every grid cell and time step ("complete information", green) and the Bayesian prior ("no data information", black) is given. The p^{CO_2} field has been averaged over the regions as in Fig. 4.

OSD
9, 2273–2326, 2012