Ocean Science

Global surface-ocean p^{CO_2} and sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme

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 Research Park, 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, Norway
[8]NOAA Pacific Marine Environmental Laboratory, Seattle, USA

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

Abstract. A temporally and spatially resolved estimate of the global surface-ocean CO_2 partial pressure field and the sea–air CO_2 flux is presented, obtained by fitting a simple data-driven diagnostic model of ocean mixed-layer biogeochemistry to surface-ocean CO_2 partial pressure data from the SOCAT v1.5 database. Results include seasonal, interannual, and short-term (daily) variations. In most regions, estimated seasonality is well constrained from the data, and compares well to the widely used monthly climatology by Takahashi et al. (2009). Comparison to independent data tentatively supports the slightly higher seasonal variations in our estimates in some areas. We also fitted the diagnostic model to atmospheric CO_2 data. The results of this are less robust, but in those areas where atmospheric signals are not strongly influenced by land flux variability, their seasonality is nevertheless consistent with the results based on surface-ocean data. From a comparison with an independent seasonal climatology of surface-ocean nutrient concentration, the diagnostic model is shown to capture relevant surface-ocean biogeochemical processes reasonably well. Estimated interannual variations will be presented and discussed in a companion paper.

1 Introduction

The oceans are considered the dominant player in the global carbon cycle on long timescales, e.g. in the glacial–interglacial cycles (e.g. Sigman and Boyle, 2000). On a multi-millennial timescale, the oceans will be the sink for 80–95 % of the anthropogenic CO_2 emissions, and 70–80 % on a timescale of several hundred years (Archer et al., 1997). Currently, the oceans take up about 25 % of the emissions (Sarmiento et al., 2010). Concerns exist, however, that the sink efficiency may decrease in the coming decades as a consequence of anthropogenic climate change, 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, two data streams are 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 parameterization. As p^{CO_2} measurements only exist at discrete locations and ship 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) multi-linear 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.

– The CO_2 exchange between the Earth surface and the atmosphere can be quantified based on atmospheric CO_2 mixing ratio measurements (e.g. Conway et al., 1994) 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) (see Sect. 2.1). This estimation does not involve any parameterization 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 ocean-internal processes (biology, transport) cause mixed-layer carbon sources and sinks of comparable variability as the land biosphere, the resulting sea–air CO_2 exchange is much less variable and smoothed out in time because the carbonate chemistry of seawater slows the equilibration rate of dissolved CO_2 with the atmosphere. In addition, the responses to warming/cooling partially counteract the effects of ocean-internal sources/sinks. 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.

A further data-based method to estimate sea–air CO_2 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 parameterizations 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) as well as from ^{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 a 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)). Out of this context, the data-based sea–air CO_2 fluxes presented here have been developed aiming to (1) provide not only a monthly seasonal cycle but variability also on short-term (daily) and interannual timescales, (2) use an assimilation scheme that can easily and self-consistently compare or even combine the observational information of atmospheric and surface-ocean data, and (3) use a framework that can be extended to also incorporate constraints from other tracers, such as oxygen.

The presented results are based on the newly available Surface Ocean CO_2 Atlas (SOCAT) database (version 1.5) of CO_2 fugacity measurements (Pfeil et al., 2012, http://www.socat.info/). We describe the estimation method and test its performance, also using independent data. As a first step of analysis, this study mainly focuses on the mean seasonal cycle, because on this timescale (1) we can compare to the widely used monthly p^{CO_2} climatology by Takahashi et al. (2009), (2) we can investigate mutual consistency between the surface-ocean p^{CO_2} constraint and that by atmospheric CO_2 data as the signal/noise ratio is large, and (3) we can test the plausibility of our process representations invoking a seasonal surface-ocean phosphate (PO_4) climatology. Estimated interannual variations of the sea–air CO_2 flux are presented and discussed in a companion paper (Rödenbeck et al., 2013).

2 Method

2.1 Concept – overview

In a classical atmospheric inversion, a spatio-temporal field of surface-to-atmosphere carbon 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: The atmospheric transport model is supplemented by a chain of parameterizations of gas exchange, carbonate chemistry, and a carbon budget equation, which determine the sea-to-air CO_2 exchange ($f_{ma}^{CO_2}$) as a function of ocean-internal carbon sources and sinks (Fig. 1). Then the inversion is not directly adjusting the sea-to-air fluxes any more, but only indirectly through adjusting the ocean-internal fluxes instead (Fig. 2, from left to middle). This offers two advantages: (1) the equations describing
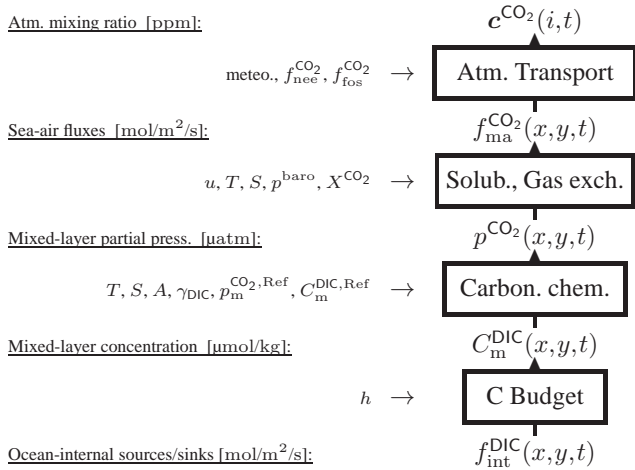
Fig. 1. Summary of the main quantities and process parameterizations involved in the diagnostic ocean mixed-layer model. Thick boxes denote the process parameterizations given in Sect. 2.2, each expressing the quantity right above as a function of the quantity right below. Quantities at the arrows on the left represent driver data entering the parameterizations. See Table 2 for mathematical symbols, and Appendix A1 for the equations and further explanation.

the individual processes impose spatial and temporal structure to the sea–air fluxes, based on the spatio-temporal information from the driver data (sea-surface temperature (SST), wind speed, etc., Fig. 1). (2) The explicit representation of oceanic processes involves further quantities, notably sea-surface 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 the atmospheric data (Fig. 2, right). In this mode of operation, atmospheric transport models and atmospheric data are actually no longer used in the inverse calculation, but the Bayesian framework, including a-priori spatial and temporal correlations, is still applied in the same way as in the atmospheric mode.

The calculation is global over the time period 1985–2011, with a temporal resolution of 1 day and a horizontal resolution of approximately 4° latitude × 5° longitude (TM3 model grid).

2.2 Process parameterizations

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 CO_2 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 ≈ 4° lat. × 5° long. × 19 vertical levels. The model is driven by 6-h interannual meteorological fields derived from the National Centers for Environmental Prediction (NCEP) reanal-
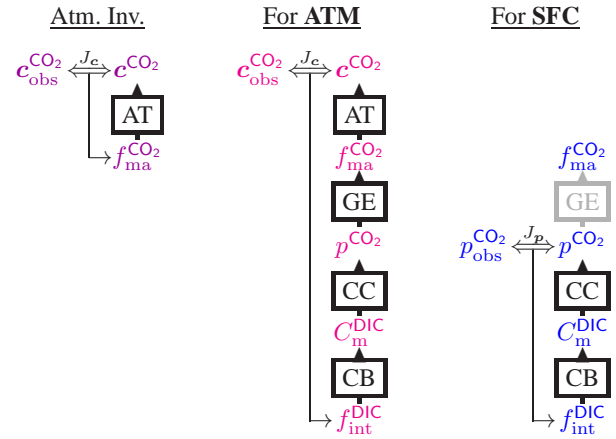

Fig. 2. Illustration of the inverse procedure (three modi). Boxes denote the process parameterizations from Fig. 1, causally linking quantities from bottom to top. Double arrows symbolize the matching between observed and modelled quantities, as gauged by a cost function J (Appendix A2). Thin arrows indicate the adjustments of the respective unknowns done to minimize the model–data mismatch. Left: pure atmospheric transport inversion not used here but given for reference: the sea–air fluxes $f_{ma}^{CO_2}$ are adjusted to match atmospheric mixing ratios (terrestrial net ecosystem exchange is adjusted as well but omitted from this schematic for clarity). Middle: in the case of **ATM**, ocean-internal sources/sinks f_{int}^{DIC} are adjusted instead of sea–air fluxes to match atmospheric CO_2 mixing ratios (again net ecosystem exchange is further adjusted). Right: in the case of **SFC**, ocean-internal sources/sinks are adjusted to match the surface-ocean CO_2 partial pressure observations. Sea–air fluxes are then calculated from the estimated partial pressure field.

ysis (Kalnay et al., 1996). The model fields are sampled at the location and time of the individual mixing ratio measurements used. TM3 has performed well in intercomparisons of state-of-the-art global tracer transport models (e.g. Stephens et al., 2007; Law et al., 2008).

Solubility and gas exchange. Diffusive sea-to-air gas exchange is proportional to the over-/undersaturation of CO_2 in the surface ocean and to the piston velocity with the quadratic wind speed dependence by Wanninkhof (1992), scaled to match the global average piston velocity of Naegler (2009) (Appendix A1.1).

Carbonate chemistry. The carbon species relevant for gas exchange (CO_2) only account 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 non-linear dependence of $p_m^{CO_2}$ on C_m^{DIC} has been linearized in standard ways (Sarmiento and Gruber, 2006). The chemistry parameterization also contains a temperature-dependent factor, and contributions from seasonal variations in alkalinity and salinity (Appendix A1.2).

Mixed-layer carbon budget. Changes in the spatio-temporal field of dissolved inorganic carbon in the ocean mixed layer need to be balanced by the sum of fluxes (Appendix A1.3). As the sea–air flux itself depends on carbon concentration, this balance can be expressed as a linear first-order differential equation. Its solution gives DIC concentration C_m^{DIC} as a function of the ocean-internal sinks f_{int}^{DIC} (representing the total effect of biological conversion and vertical/horizontal advection/diffusion on mixed-layer carbon concentration). An additional important item in the budget is the re-entrainment during mixed-layer deepening of carbon 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 Main adjustable degrees of freedoms

The ocean-internal carbon sources and sinks (f_{int}^{DIC}) at the end of the described chain of process parameterizations is the basic unknown to be adjusted to match the surface-ocean p^{CO_2} or atmospheric CO_2 data (next section). It is treated similarly to the unknown sea–air flux in the pure atmospheric transport inversion (Fig. 2), including Bayesian a-priori spatial and temporal correlations. The detailed specification and further explanations are found in Appendix A2.2.

2.4 Data constraints – base runs

We will present results of two cases that differ in the data set used as main constraint (Fig. 2):

SFC. In the main case used to create the primary product of this study, the ocean part of the diagnostic scheme is fit to surface p^{CO_2} data points from the SOCAT v1.5 database (Pfeil et al., 2012, http://www.socat.info/). Data pre-treatment and further details are given in Table 1. SOCAT data cover about 6 million pixels/time steps globally within the calculation period (until 2007, see Supplement Figs. S7.3 and S7.4 for data distribution).

ATM. As a comparison case used to assess the consistency between surface-ocean and atmospheric data, the diagnostic scheme is fit to atmospheric CO_2 mixing ratios measured approximately weekly or hourly at a set of sites by various institutions. Details are given in Table 1.

3 Results and discussion

The main product of this study is a data-based estimate of the global spatio-temporal CO_2 partial pressure field and ultimately the sea–air CO_2 fluxes. Here we characterize and evaluate its robustness, errors, and information content. In addition, we compare these results with independent data and previous estimates. The consideration of interannual variations is done in the companion paper (Rödenbeck et al., 2013).

3.1 Overview

To illustrate the characteristics of the quantities and the temporal scales involved in the scheme, Fig. 3 shows time series of key quantities (as of run **SFC**, blue) at the example pixel around Station M in the Norwegian Sea (66° N, 2° E). The ocean-internal sources/sinks (bottom panel) are relatively smooth, as the a-priori temporal correlations (Appendix A2.2) suppress fast variations. 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. In addition, the DIC concentration is rising in response to the atmospheric CO_2 increase. The surface-ocean CO_2 partial pressure (3rd panel from bottom) shares the rise and the variations with the DIC concentration, but the seasonality is again slightly shifted because of the temperature and alkalinity effects. Finally, the sea–air CO_2 flux (top panel) is dominated by the p^{CO_2} variability; in addition, it shows high-frequency (daily) variability due to wind speed and solubility changes. Figure 3 also illustrates the dampening effect of the carbonate chemistry on sea–air exchange, as the seasonal amplitude of the sea–air flux (top) is much lower than that of the ocean-interior flux (bottom).

The spatial resolution and domain of the calculation is illustrated by Fig. 4, showing the amplitude of the mean seasonal cycle of p^{CO_2} for each pixel.

3.2 Data and model constraints

The estimated p^{CO_2} and sea–air CO_2 flux fields combine information from the partial pressure data and from the process parameterizations. To illustrate this, Fig. 3 also shows the a-priori state (thin dashed grey) defined by $f_{int,pri}^{DIC} = 0$ (Appendix A2.2, see bottom panel). The choice $f_{int,pri}^{DIC} = 0$ means that, without data knowledge, it is equally likely to have an internal source or an internal sink at any given location and time, and thus corresponds to a state of no information on f_{int}^{DIC}. The variations in the prior values of C_m^{DIC}, p^{CO_2}, and the sea–air fluxes $f_{ma}^{CO_2}$, which then follow from the process parameterizations, only comprise responses to variations in sea surface temperature and the other driving variables (including rising atmospheric CO_2), but miss variations in the ocean-internal sources/sinks. As this a-priori p^{CO_2} field contradicts the data (black dots), the estimation procedure now adjusts the internal flux f_{int}^{DIC} in such a way that the data are matched as closely as possible (blue). Most prominently, this reduces or – as at the example pixel of Fig. 3 – even reverses the seasonal variations of p^{CO_2} (reflecting that thermal, biological, and physical effects partially oppose each other). The fit at further locations is given below (Sect. 3.4).

For most pixels, the SOCAT data set only sporadically contains data points (in particular, periods of dense
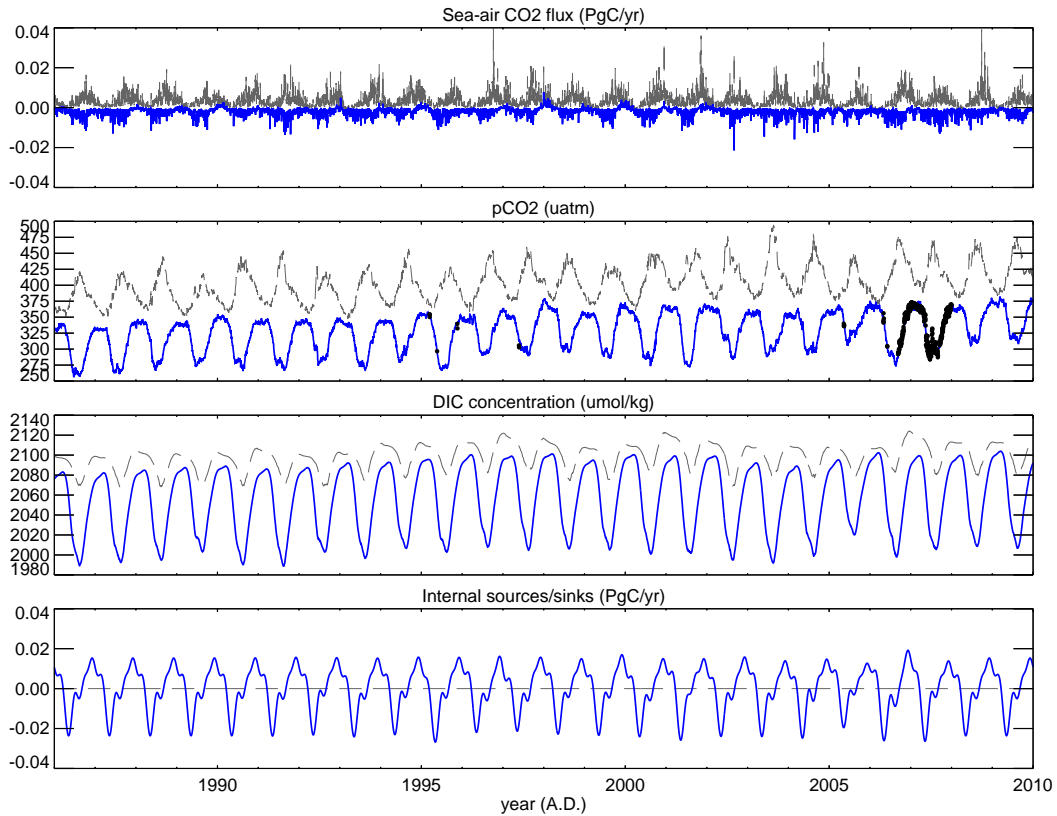
Fig. 3. Example time series (full daily resolution) at the model pixel around Station M (66° N, 2° E) of the main quantities of the diagnostic model (Fig. 1): sea–air CO_2 fluxes ($f_{ma}^{CO_2}$), surface-ocean CO_2 partial pressure ($p_m^{CO_2}$), mixed-layer DIC concentration (C_m^{DIC}), as well 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). The SOCAT data points that happen to fall in the example pixel are shown as black dots in the $p_m^{CO_2}$ panel. The Bayesian prior ($f_{int}^{DIC} = 0$ and corresponding C_m^{DIC}, $p_m^{CO_2}$, and $f_{ma}^{CO_2}$, Appendix A2.2, Sect. 3.2) is also given (thin dashed grey lines).
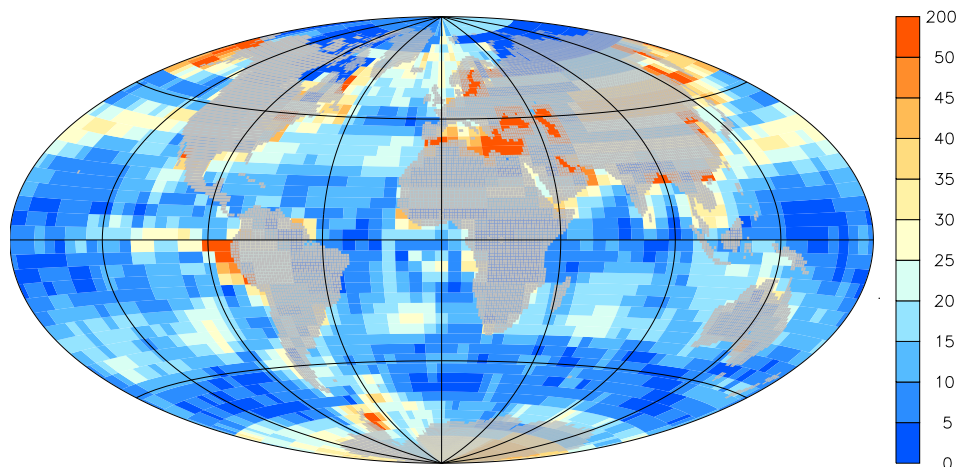

Fig. 4. Amplitude of the seasonal cycle of surface-ocean CO_2 partial pressure (µatm) estimated by fitting the diagnostic scheme to the SOCAT data (run **SFC**). The amplitude is given as temporal standard deviation of the 1997–2009 monthly mean $p_m^{CO_2}$ at each pixel.
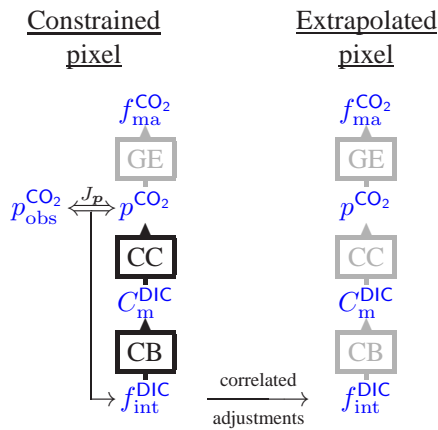
Fig. 5. Illustration of the information flow in the spatio-temporal extrapolation through a-priori correlations. Where data values exist, pixels are directly constrained (left column, which is identical to Fig. 2). Due to the a-priori spatial correlations (Appendix A2.2), adjustments to f_{int}^{DIC} at the constrained pixel also change f_{int}^{DIC} at neighbouring unconstrained pixels. p^{CO_2} is then calculated passively from the extrapolated f_{int}^{DIC} by the process parameterizations, using the local values of the driving fields (right column).

In data-dense areas, local adjustments to f_{int}^{DIC} (both for constrained and for extrapolated pixels) will need to compromise between several data constraints within the correlation radius (de-weighted with distance due to the decay of the correlation, Fig. A1). As with any smoothing, this may be beneficial in suppressing outliers, but also adverse in partly suppressing small-scale signals (the structures from the driving data of the parameterizations are never smoothed however, as the smoothing only acts on f_{int}^{DIC}). In data-poor areas, there may exist pixels that do not have any constrained pixel within the correlation radius; then f_{int}^{DIC} will stay close to the prior.

In time, extrapolation happens not only due to the a-priori correlations, but additionally due to the memory effect resulting from the relaxation time of the budget equation Eq. (A21).

regular sampling, as at Station M since 2006, are a rare exception). Thus, only a small fraction of pixels/time steps is constrained directly in the described way. However, most parts of the space–time domain are still constrained indirectly through extrapolation via the spatial and temporal a-priori correlations defined in Appendix A2.2. The mechanism is illustrated in Fig. 5. In particular, as seasonal variations in f_{int}^{DIC} are allowed to be adjusted much more freely than non-seasonal variations (equivalent to a-priori correlations between adjustments to f_{int}^{DIC} at any given day of year in each year, Appendix A2.2), the mean seasonal cycle is extrapolated throughout the time period of the calculation. This also means that data points in any year can contribute to constrain this mean seasonality. Thus, though there are regions or periods too far away from data points for extrapolation, larger-scale seasonal variations are constrained almost everywhere in the ocean. This has been confirmed by assessing how well a given p^{CO_2} field can be retrieved by the scheme

from pseudo data sampled from this field at the times and locations of the SOCAT data (Appendix B). Constraining inter-annual variations requires more even data distribution in time than constraining seasonal variability; this density of data is only available is some areas of the ocean (Rödenbeck et al., 2013).

The spatio-temporal a-priori correlations also act to smooth f_{int}^{DIC} on small scales (Fig. 3, bottom), as they damp small-scale and short-term variations. This has been done because the available data do not have the spatial or temporal density required to constrain these fine-scale features adequately. Nevertheless, despite the smooth f_{int}^{DIC} the corresponding p^{CO_2} and sea–air CO_2 flux fields (Fig. 3, top) do comprise fast variations as represented in the process parameterizations, including responses to variations in temperature (changes in solubility and chemical equilibrium) or in wind speed (e.g. accelerated depletion of an oversaturated mixed layer in a high wind speed event, Bates et al., 1997). These effects already account for a considerable part of short-term variations (see Sect. 3.4 below). Only variations related to fast ocean-internal processes, such as the sub-weekly variations of upwelling events or algal blooms, will be missing. In any case, the pseudo-data tests (Appendix B) confirm that the degrees of freedom in the diagnostic scheme provide sufficient flexibility to follow variability as contained in the p^{CO_2} fields from a biogeochemical process model simulation (as the model has less day-to-day variability than the real p^{CO_2} field, however, this test may underestimate errors from aliasing such high-frequency variations into variations on the resolved timescales).

3.3 Robustness

Figures 6 and 7 show the seasonality of the sea–air CO_2 flux and the surface-ocean CO_2 partial pressure as estimated from SOCAT data by run **SFC**. The fields have been integrated/averaged over a set of regions splitting the ocean into basins and latitude bands.

To test how robustly the results are determined from the data and to assess the impact of errors in the parameterizations and their driving fields, individual important model elements have been changed in sensitivity runs: (i) increase and decrease of the a-priori uncertainty by a factor 2, leaving more/less freedom for inverse adjustments, (ii) decrease of the a-priori uncertainty of non-seasonal variations in f_{int}^{DIC} by a factor of 2, (iii) increase in the spatial correlation lengths by a factor of 3, (iv) increase and decrease of the global mean piston velocity by 3.2 cm (range given by Naegler, 2009), and (v) increase and decrease mixed-layer depth h by a factor of 2. The resulting range of results (grey bands in Figs. 6 and 7) is generally very narrow compared to the seasonal amplitude, in particular for p^{CO_2}. In most regions, the sensitivity of the sea–air flux is slightly larger than that of p^{CO_2}, almost entirely due to the uncertainty of gas exchange (variation in piston velocity). Larger spread in p^{CO_2} than in the sea–air flux
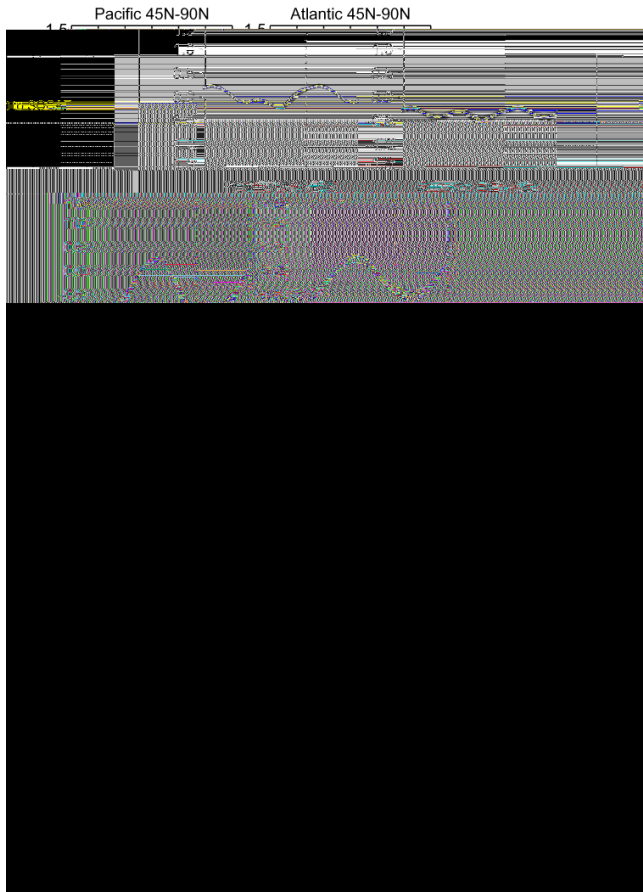
Fig. 6. Monthly mean seasonal cycle of sea–air CO_2 fluxes, estimated by fitting the diagnostic scheme to SOCAT data (run **SFC**, blue). The grey shading around the standard result comprises sensitivity cases listed in Sect. 3.3. Fluxes have been integrated over latitude bands split into ocean basins (including open ocean, coastal areas and marginal seas; panels in geographical arrangement). Seasonal cycles have been calculated over the period 1997–2009; the first half year is repeated for clarity.

Fig. 7. Analogously to Fig. 6, monthly mean seasonal cycle of surface-ocean carbon dioxide partial pressure $p_m^{CO_2}$, estimated by fitting the diagnostic scheme to SOCAT data (run **SFC**, blue). The grey shading comprises the same sensitivity cases. For comparison, the $p_m^{CO_2}$ climatology by Takahashi et al. (2009) is given (violet, with a small additive adjustment to transfer the nominal year 2000 average into a 1997–2009 average according to the rising atmospheric CO_2 partial pressure field $p_a^{CO_2}$, Eq. A3).

is found mostly in the North Pacific and the temperate North Atlantic, entirely due to a few marginal areas (including pixels in the far north of the Pacific, or the Black Sea and the Caspian Sea added to the Atlantic region, see Fig. 4) where the p^{CO_2} seasonality is exceptionally high due to high SST seasonality, but where gas exchange is low such that their contribution to regional sea–air flux is nevertheless small. If p^{CO_2} is averaged only over the open ocean, the p^{CO_2} spread becomes narrow in all regions (see Supplement Fig. S7.8).

3.4 Fit to the data

The match of the estimated p^{CO_2} field and measured data points somewhat depends on the considered location (Fig. 8). At the North Pacific example pixel (i.e. within the region of largest spread among the sensitivity cases in Fig. 7), the seasonality is underestimated by some µatm. This may result
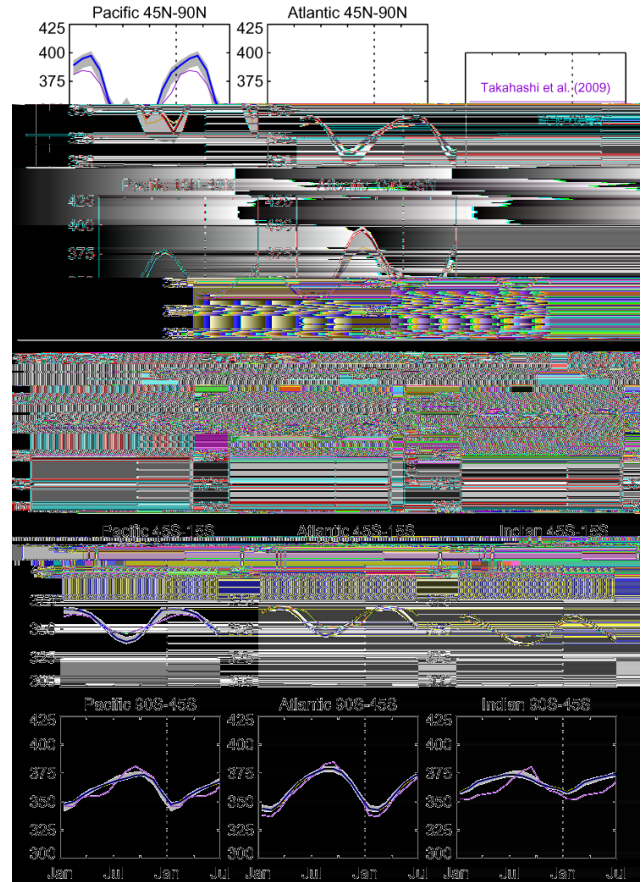
from scale mismatch between the model pixel ($\approx 4°$ lat. $\times 5°$ long.) and the point measurement, but likely also from the need to compromise between the signals of neighbouring locations in the fit: the statistics of the misfit in the surrounding area (Pacific $45°$–$90°$ N) reveal very small seasonal biases (well below 1 µatm, Fig. 9). The distribution of residuals deviates somewhat from the Gaussian assumed mathematically in the Bayesian inference by having wider tails, but is symmetric. Example locations STM and – to a lesser degree – BTM show close seasonal fit (Fig. 8), and agreement also in some high-frequency features (see below). A less successful fit is found at Stratus (South Pacific) in an area of low data density (note that the independent data at this location are mainly from after the end of the SOCAT v1.5 data period; thus the agreement relies on the extrapolating effect of the dominating seasonal degrees of freedom (Appendix A2.2)
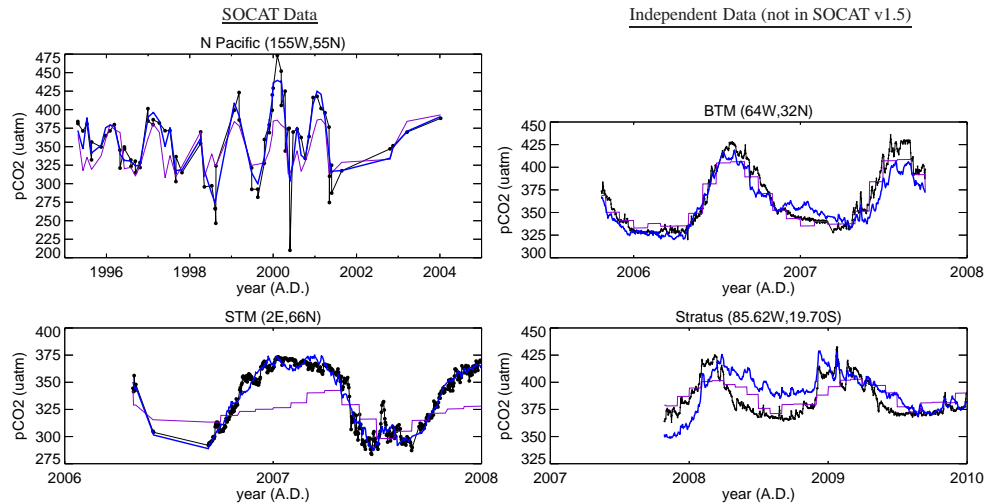
Fig. 8. Comparison between the estimated p^{CO_2} field (run **SFC**, blue) and measurements (black dots) at test locations. Left: data contained in SOCAT v1.5 (larger dots) at a pixel in the North Pacific, and the pixel around Station M in the North Atlantic. Right: two mooring sites in the North Atlantic (BTM, 32° N, 64° W) and South Pacific (Stratus, 19.70° S, 85.62° W) by Sabine et al. (2010) that are not part of SOCAT v1.5 (smaller dots). The p^{CO_2} field has been picked at the times where the measurements exist (connected by straight lines for clarity); the time axes have been limited to respective years with data. The monthly climatology by Takahashi et al. (2009) (regridded and with added trend parallel to the atmospheric CO_2 increase, violet) is also given. See Supplement for comparisons to SOCAT data (Fig. S7.5) and independent data (Fig. S7.6) at further locations.
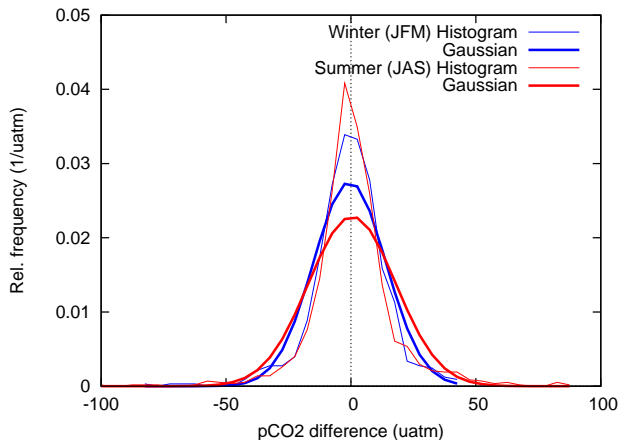

Fig. 9. Histograms of the residuals of the fit to SOCAT data (i.e. p^{CO_2} estimated by run **SFC** minus data, µatm, horizontal axis). Thin lines show the relative abundance (within bins of 5 µatm) of residuals over all pixels from the North Pacific (north of 45° N) and all time steps with data in winter (January–February–March, blue) or in summer (July–August–September, red). Respective thick lines give a Gaussian of the same mean and standard deviation. See Supplement Fig. S5.2 for global maps of seasonal residuals.

but misses any interannual anomalies). See the Supplement (Sect. S5, Figs. S7.5 and S7.6) for more comparisons and statistics.

In order to validate the high-frequency variations, we compare them to mooring observations independent of the SO-CAT v1.5 data set (Fig. 10, upper panels). In the extra-tropics (example sites BSO or MOSEAN+WHOTS), many short-term features are reproduced in run **SFC**, though mostly with reduced amplitude. Discrepancies (e.g. beginning of 2007 at BSO) can largely be explained by the deviation of sea surface temperature in our driving field – regridded to the $\approx 4° \times 5°$ pixels around the stations – from the local temperature (lower panels). Note that this discrepancy does not preclude that the pixel average of p^{CO_2} is actually reproduced more accurately than the local value, to the extent that our SST field represents the pixel's average temperature; however, this cannot be validated or falsified based on the point data.

In contrast to the extra-tropics, the high-frequency variations in our estimate are unrelated or even opposite to those of the measurements in the tropics (example site TAO170W), even though the SST field performs similarly as at the extra-tropical sites. While in the extra-tropics warmer SST leads to higher $p_m^{CO_2}$ as reproduced by the solubility and chemistry parameterizations, tropical $p_m^{CO_2}$ increases during cold events, as in the upwelling of cold and carbon-rich deep water. Such events would have to be reproduced as high-frequency variations in the ocean-internal sources/sinks f_{int}^{DIC}, which however cannot be constrained from the SO-CAT data set (potentially except at very few locations in periods of high-frequency sampling).

As similar conditions are found at other extra-tropical and tropical sites (Supplement Figs. S7.6, S7.7), we conclude that our results reproduce at least part of the real short-term variations in the extra-tropics dominated by solubility and

Table 1. Data sets used as constraints or as driver data in the process parameterizations (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), http://www.socat.info/	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 (approximately weekly) or hourly averages, 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		Lee et al. (2006)	Gridded monthly climatology (interpolated)	All
γ^{DIC}		Egleston et al. (2010)	Mean spatial pattern	All
$C_m^{DIC, Ref}$	GLODAP	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:				
$p_m^{CO_2}$		Takahashi et al. (2009)	Gridded monthly climatology	Sect. 3.5
$C_m^{PO_4}$	WOA 2005	Garcia et al. (2006)	Gridded monthly climatology	Sect. 3.7

Glossary: GLODAP = Global Ocean Data Analysis Project; LOCEAN = Laboratoire d'océanographie et du climat: expérimentations et approches numériques; Meteo. = Meteorological variables as listed in Heimann and Körner (2003); NCEP = National Centers for Environmental Prediction; OAFlux = Objectively Analysed air–sea Fluxes; SOCAT = Surface Ocean CO_2 Atlas; WOCE = World Ocean Circulation Experiment; WOA = World Ocean Atlas; Math symbols see Table 2.

chemistry. In the tropics, more realistic short-term variations could potentially be obtained by parameterization of fast processes (upwelling events) in f_{int}^{DIC}.

3.5 Comparison to the Takahashi climatology

The monthly mean seasonal cycle calculated from our spatiotemporal p^{CO_2} field has been compared to the widely used observation-based climatology by Takahashi et al. (2009), which differs from our method in the p^{CO_2} data set used (Lamont-Doherty Earth Observatory (LDEO) database, Takahashi et al., 2010) and in the way of spatio-temporal extrapolation of the point data into pixels without data: while the diagnostic scheme extrapolates the internal flux field f_{int}^{DIC} and then infers p^{CO_2} from it via the process representations (Fig. 5), Takahashi et al. (2009) extrapolates the p^{CO_2} field itself, using a 2-dimensional advection–diffusion equa-

tion[1]. In data-dense areas, results should not depend much on the extrapolation method.

In most regions, the seasonal cycle estimated from SOCAT by the diagnostic scheme is similar in phase to the p^{CO_2} climatology by Takahashi et al. (2009) (Fig. 7), though the seasonal amplitude is somewhat larger in many regions. Even better agreement is found when only considering the open ocean: the open-ocean seasonalities in the temperate North Pacific and Atlantic almost fully coincide (Supplement Fig. S7.8). Open-ocean agreement is also confirmed at several of our test locations (Fig. 8).

Differences between our seasonalities and the Takahashi et al. (2009) climatology are somewhat bigger in areas of low data density (e.g. the Southern Ocean), as expected. Comparison to data at typical test locations tends to confirm the slightly larger seasonal amplitudes in our estimate (Fig. 8

---

[1] As a further methodological difference, the diagnostic scheme uses each data point at its original sampling time; thus no transfer to a nominal year as in Takahashi et al. (2009) is needed and interannual variations in atmospheric CO_2 above the ocean are fully taken into account.

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
$\mathbf{A}$	$\mathrm{ppm\,(\mu mol\,m^{-2}\,s^{-1})^{-1}}$	Atmospheric transport matrix, Eq. (A25)
$C_{\mathrm{m}}^{\mathrm{DIC}}$	$\mathrm{\mu mol\,kg^{-1}}$	Mixed-layer DIC concentration
$C_{\mathrm{a}}^{\mathrm{DIC}}$	$\mathrm{\mu mol\,kg^{-1}}$	DIC concentration in equilibrium with atmosphere
$\boldsymbol{c}_{\mathrm{obs}}$	ppm	Set of observed atmospheric mixing ratios
$\boldsymbol{c}_{\mathrm{mod}}$	ppm	Modelled mixing ratios, Eq. (A25)
$\boldsymbol{c}_0$	ppm	Initial mixing ratio
f_{ma}	$\mathrm{\mu mol\,m^{-2}\,s^{-1}}$	Sea-to-air tracer flux, Eq. (A17)
f_{hist}	$\mathrm{\mu mol\,m^{-2}\,s^{-1}}$	History (re-entrainment) flux, Eq. (A20)
f_{int}	$\mathrm{\mu mol\,m^{-2}\,s^{-1}}$	Ocean-internal tracer sources/sinks, Eq. (A18)
$\mathbf{F}$	$\mathrm{\mu mol\,m^{-2}\,s^{-1}}$	Flux model matrix, Eq. (A26)
$\boldsymbol{f}$	$\mathrm{\mu mol\,m^{-2}\,s^{-1}}$	Vector of flux discretized at pixels/time steps
h	m	Mixed layer depth
J	1	Cost function, Eq. (A27)
$J_{\boldsymbol{c}}$	1	Cost function, atmospheric data part, Eq. (A24)
$J_{\boldsymbol{p}}$	1	Cost function, surface-ocean data part, Eq. (A30)
k	$\mathrm{m\,s^{-1}}$	Piston velocity, Eq. (A2)
L	$\mathrm{mol\,kg^{-1}\,atm^{-1}}$	Solubility (including fugacity correction), Eqs. (A1) and (S1.1)
p_{m}	µatm	Partial pressure in mixed layer, Sect. A1
p_{a}	µatm	Atmospheric partial pressure, Eq. (A3)
p^{baro}	atm	Barometric pressure at ocean surface
p^{dry}	atm	Dry-air pressure at ocean surface
$p^{\mathrm{H_2O}}$	µatm	Saturation water vapor pressure above ocean, Eq. (S1.3)
S	‰	Mixed-layer salinity[b]
Sc	1	Schmidt number, Appendix S1.1
T	°C	Mixed-layer temperature (sea surface temperature, SST)
t	s	Time coordinate
$t_{\mathrm{i}}, t_{\mathrm{e}}$	s	Start time, end time of inversion period
u	$\mathrm{m\,s^{-1}}$	Wind speed at 10 m above surface
$\boldsymbol{x}$	1	Vector of adjustable parameters, Eq. (A26)
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. (A5) and (A6)
Γ	1	Scaling of piston velocities, Eq. (A2)
γ^{DIC}	$\mathrm{\mu mol\,kg^{-1}}$	Buffer factor (response factor), Eq. (A11)
ε	1	Ice-free fraction of ocean surface (0 = ice-covered, 1 = ice-free)

[a] ppm = $\mathrm{\mu mol\,mol^{-1}}$; [b] for the present purposes, salinity in ‰ or on the Practical Salinity Scale are considered interchangeable.

and Supplement Figs. S7.5 and S7.6). In particular, in the North Pacific where differences between the two methods are biggest, larger seasonalities are confirmed both at the individual pixels (see Fig. 8 for a typical example) and in the larger seasonal bias of about 8.5 µatm in the differences between Takahashi et al. (2009) and SOCAT data points in that area (however, the largest contribution to the difference in the $p^{\mathrm{CO_2}}$ average of Fig. 7 comes from the northernmost pixels where data are too sparse for a conclusive validation). Data at Station M (Fig. 8) also confirm larger seasonality in the North Atlantic.

Differences between the two estimates are also bigger towards the coast. Note that part of these coastal differences come from the fact that, in order to cover the entire ocean surface, we extrapolated the Takahashi et al. (2009) clima-

tology using its open-ocean values. For example, this contributes to the differences in the temperate North Atlantic due to the Mediterranean (see Supplement, Fig. S7.5).

Partly, the differences in Fig. 7 can be traced to the fact that the two estimates are based on different $p^{\mathrm{CO_2}}$ data sets, though there is considerable overlap between the SOCAT and LDEO databases (this influence has explicitly been tested in the Supplement Sect. S6). Sensitivity to data selection partially arises from interannual variations; in particular, differences in the tropics may be related to the exclusion of El Niño data in Takahashi et al. (2009). Moreover, the synthetic data test (Appendix B) reveals slightly too high seasonality in our results, though most of this mismatch is confined to certain marginal areas (that do not contribute much to the sea–air flux due to low gas exchange).
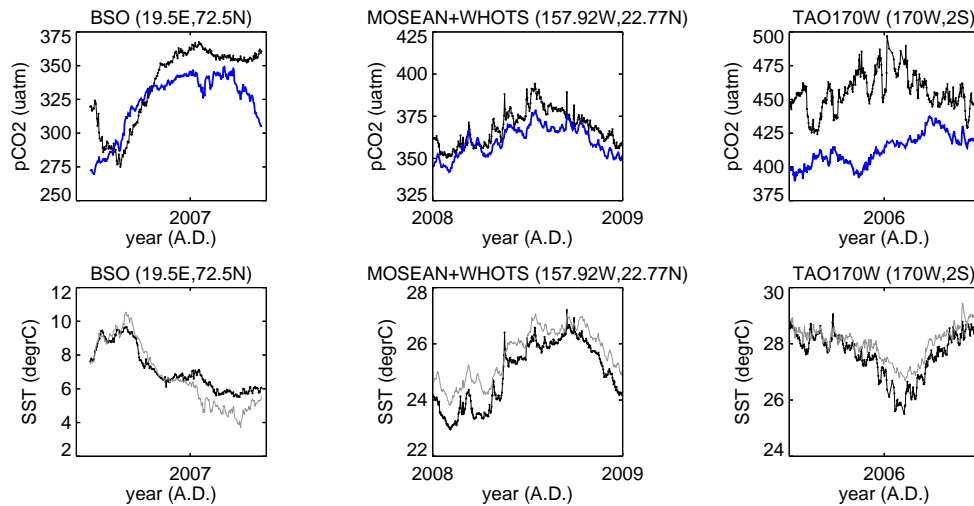
Fig. 10. Top: time series of surface-ocean p^{CO_2} (µatm) at independent mooring sites (not part of SOCATv1.5): Barents Sea Opening (72.5° N, 19.5° E, Arthun et al., 2012), combined MOSEAN and WHOTS moorings (22.77° N, 157.92° W, Sabine et al., 2010), and from the TAO array (2° S, 170° W, Sabine et al., 2010). Comparison between the measurements (black dots) and the p^{CO_2} field estimated by fitting to the SOCAT data (blue). The estimated field has been picked at the times where measurements exist; one year has been selected for each site. Bottom: sea surface temperature measured simultaneously with p^{CO_2} (black dots) compared to our SST driving field at this pixel (light grey).

Beyond seasonality, agreement is also found in the implied long-term average global sea–air flux. The 1990–2009 average from run **SFC** is $-1.2\,\mathrm{PgC\,yr^{-1}}$ (range between $-0.95\,\mathrm{PgC\,yr^{-1}}$ and $-1.45\,\mathrm{PgC\,yr^{-1}}$ due to sensitivity cases with lower/higher gas exchange rate, Sect. 3.3). Within uncertainties, this agrees well with the value of $(-1.36\pm 0.6)\,\mathrm{PgC\,yr^{-1}}$ recomputed from Takahashi et al. (2009) by Wanninkhof et al. (2012, Table 2, "Net flux" and "Coastal area") using CCMP winds (Atlas et al., 2011).

In summary, we take the similarity between the seasonalities from the two methods based on similar data as an indication that the estimated p^{CO_2} field in most regions is informed by the data and not primarily method-dependent.

3.6 Comparison of p^{CO_2}-based and atmospheric results

Figure 11 compares the SOCAT-based fit from Fig. 6 (run **SFC**) with the fit of the diagnostic scheme to atmospheric CO_2 data (run **ATM**). In the Southern Hemisphere, there is broad agreement in the phasing and amplitude of the seasonal cycle. In particular, the difference between the two runs is much smaller than their differences from the (common) prior, i.e. than the adjustments due to the respective data constraints. Less agreement is found in the tropics and the Northern Hemisphere, in particular in the temperate latitudes.

The ocean fluxes estimated from the atmospheric data are much less robust (more sensitive to changes in uncertain model details, not shown) than those estimated from the p^{CO_2} data. As investigated by synthetic data testing (Supple-

ment Sect. S4), this is largely related to the additional degrees of freedom in adjusting land-to-atmosphere 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 temperate North Pacific is due to the short but strong sink period during summer, more typical for terrestrial boreal ecosystems.

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 stronger constraints on internal ocean processes, in particular in ocean areas closer to land. Combining the oceanic and atmospheric constraints is therefore both possible and beneficial (compare Jacobson et al., 2007). It 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 "classical" atmospheric inversion (similar as in the joint inversion by Jacobson et al., 2007). Details will be discussed further in the context of interannual variations in the companion paper (Rödenbeck et al., 2013).
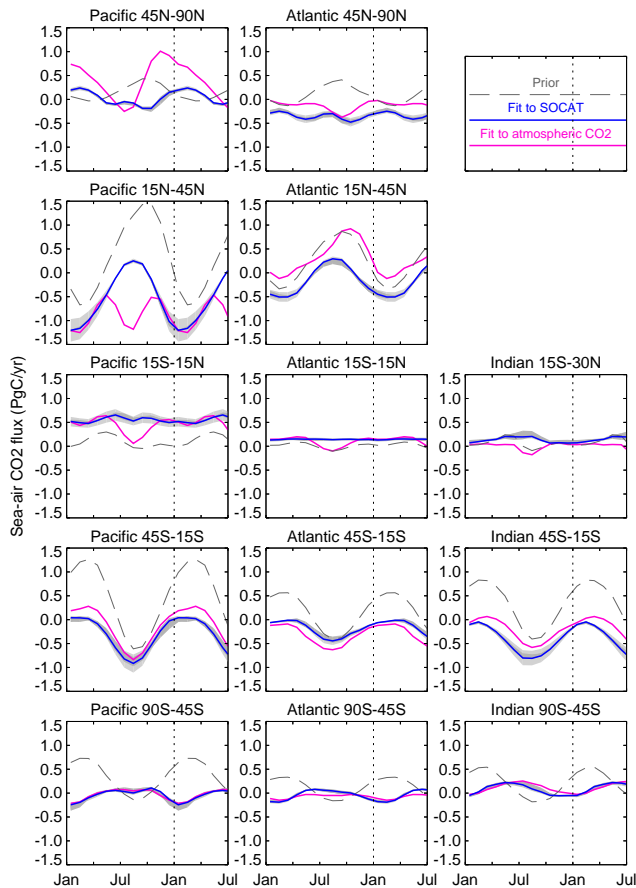
Fig. 11. Monthly mean seasonal cycle of sea–air CO_2 fluxes as Fig. 6. Estimates by fitting the diagnostic scheme to SOCAT data (run **SFC**, blue) are compared to the fit of the scheme to atmospheric CO_2 mixing ratio data (run **ATM**, magenta), as well as to the prior (no data constraint, thin dashed grey)
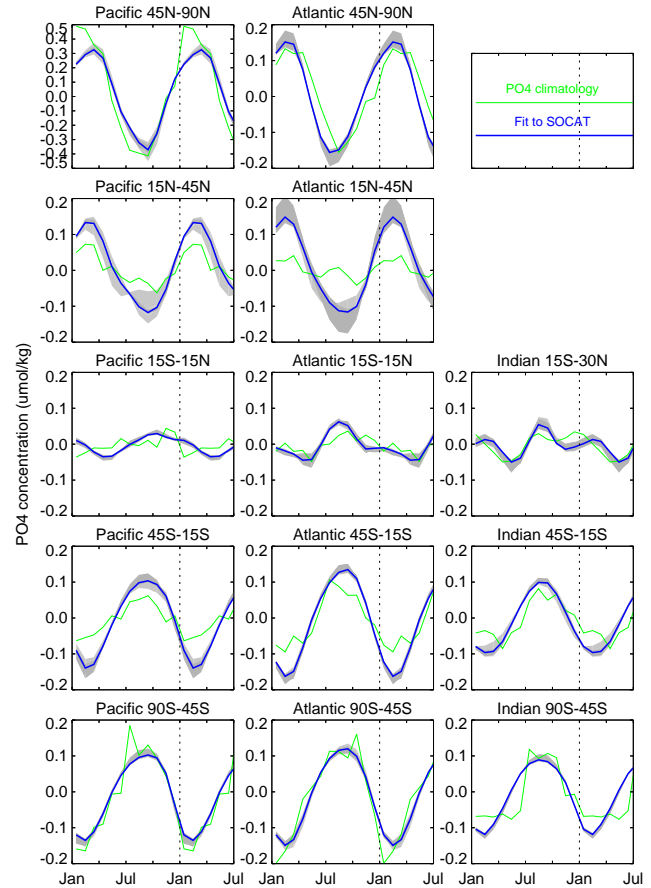

Fig. 12. Mixed-layer phosphate concentration (monthly climatology, mean subtracted). Seasonality inferred from the SOCAT fit (blue) assumes all ocean-internal carbon sources and sinks are related to phosphate sources and sinks in Redfield proportions. For comparison, the PO_4 climatology from the World Ocean Atlas (WOA, Garcia et al., 2006) is given (green). The grey uncertainty band around the SOCAT-based estimates again gives the range of sensitivity cases as in Fig. 7; note however that calculated PO_4 is also sensitive to several further model elements not included in this range.

3.7 Plausibility of the ocean-internal carbon sources and sinks

As the diagnostic scheme (in run **SFC**) is constructed to match the CO_2 partial pressure data, any errors in the parameterizations linking p^{CO_2} and f_{int}^{DIC} (namely carbonate chemistry (including temperature or alkalinity dependence), or mixed-layer carbon budget (including freshwater and history fluxes), see Fig. 1) will be compensated by spurious additional adjustments to the ocean-internal carbon sources and sinks. On the one hand, this means that the estimated p^{CO_2} (and sea–air flux) fields will be only little affected by imperfections in these parameterizations. Even at pixels not directly constrained by a p^{CO_2} data value but only indirectly via extrapolation of the f_{int}^{DIC} field through its a-priori spatio-temporal correlations (Fig. 5), any such bias will largely cancel out (to the extent that it is the same at the directly constrained and the extrapolated pixels). This leads to the robustness of the p^{CO_2} estimate shown in Sect. 3.3. Therefore, as long as we are only interested in the p^{CO_2} (and sea–air flux) fields, f_{int}^{DIC} could be just considered a mathematical device in the mapping of p^{CO_2}.

On the other hand, if the ocean-internal fluxes are of interest themselves (or in the envisaged extensions of the scheme using further data streams like oxygen coupled to carbon via f_{int}^{DIC}), correctness of the parameterizations is needed. On the seasonal timescale considered here, the plausibility of the estimated f_{int}^{DIC} field can be tested in the light of mixed-layer phosphate concentrations. As explained in Appendix C, a modelled mixed-layer phosphate field can approximately be calculated from our f_{int}^{DIC} estimate, and can be compared to an observation-based monthly climatology (World Ocean Atlas, Garcia et al., 2006). The inferred PO_4 concentration

seasonality is broadly similar in phase and amplitude to the observations (Fig. 12), including characteristic region-to-region differences, though overestimation is seen especially in the temperate latitudes. Given that part of the mismatch arises because the modelled PO_4 concentration neglects non-Redfieldian sources/sinks (as justified in Appendix C) and because of errors in the observed PO_4 climatology which are likely significant (and amplified by the Redfield ratio), we take this agreement as broad confirmation of the employed parameterizations.

The carbonate chemistry parameterization (including its driving fields, in particular alkalinity) can specifically be checked using DIC data. This is done in the companion paper (Rödenbeck et al., 2013).

3.8 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 presented 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 of being data-dominated, quite independent of driver data sets and process parameterizations. Regression methods (multi-linear 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 of how $p_m^{CO_2}$ and these driving fields are related. Finally, data assimilation into process models then also prescribes this relation analytically, which brings in process knowledge as a further constraint, with the risk of suppressing real variability not captured in the chosen parameterizations.

The diagnostic scheme is similar to the regression methods, thereby also reproducing some of the short-term variability (Sect. 3.4). 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 (it is 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[2].

---

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

An advantage of explicit parameterizations, as in the proposed mixed-layer scheme, is that the unknowns (here: internal sources/sinks) are physical quantities which are potentially open for interpretation (compare Sect. 3.7). They also enable us 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 ($p_m^{CO_2}$ or atmospheric data) to create alternative estimates for comparison, or in conjunction with the present ones to potentially introduce further unknowns and so 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., 2013).

Beyond the scheme as presented here, multi-linear regression could be brought in by expressing the unknown ocean-internal fluxes as a linear combination of (further) driving variables and then match the data by adjusting time-independent weights. This would make the estimation more similar to data assimilation, with the benefits and risks mentioned above. Further, instead of linear combinations, parameterizations 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.

4 Summary and conclusions

We presented a temporally and spatially resolved estimate of the global surface-ocean CO_2 partial pressure field and the sea–air CO_2 flux, obtained on the basis of CO_2 partial pressure data from the SOCAT database. In order to interpolate the point data into a spatio-temporal field and to add short-term variations not constrained from the data, a simple model of surface-ocean biogeochemistry, including representations of sea–air gas exchange, solubility, carbonate chemistry, mixed-layer DIC budget, and seasonal re-entrainment ("history"), was fitted to the observations. The inverse procedure was similar to atmospheric inversion calculations, using spatial and temporal a-priori correlations to extrapolate the data information.

Focusing first 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 p^{CO_2} field and sea–air CO_2 flux (run **SFC**).

– In terms of seasonality, the resulting p^{CO_2} field agrees well with the climatology by Takahashi et al. (2009), confirming that the estimates are informed by the data.

Differences mainly occur in coastal areas and marginal seas, as well as in regions where data density is low.

– 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, as indicated by the comparison of predicted and observed seasonal phosphate variations. The scheme can be extended by parameterizations of ocean-internal sources and sinks, that allow us 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 B assess possible impacts of additional data on constraining ocean biogeochemistry, helping to design future observation strategies.

The results also include interannual variations, which will be considered in a companion paper (Rödenbeck et al., 2013).

The gridded p^{CO_2} and sea–air CO_2 flux estimates will be made available for download in digital form from the Jena inversion website, www.bgc-jena.mpg.de. Regular updates are planned.

Appendix A

Model documentation

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

A1 Ocean process parameterizations

This section details the equations used to express sea–air flux as a function of partial pressure (Sect. A1.1), partial pressure as a function of mixed-layer concentration (Sect. A1.2), and mixed-layer concentration as a function of ocean-interior source/sinks (Sect. A1.3), going down the chain of parameterizations in Fig. 1.

The variables of the following parameterizations are spatio-temporal fields, with a temporal resolution of 1 day and a horizontal resolution of $\approx 4° \times 5°$, 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,

$$f_{ma}^{CO_2} = k^{CO_2} \varrho L^{CO_2} \left(p_m^{CO_2} - p_a^{CO_2} \right) \tag{A1}$$

(atmospheric sign convention: positive = source to atmosphere), where $p_a^{CO_2}$ is the CO_2 partial pressure in the atmosphere, and $p_m^{CO_2}$ the CO_2 partial pressure in the homogeneous mixed layer[3]. The temperature-dependent solubility L^{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^{CO_2}. We use the formulation according to Wanninkhof (1992):

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

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

$$p_a^{CO_2} = X^{CO_2} p^{dry} \tag{A3}$$

inherits variability not only from atmospheric dry-air pressure p^{dry} (Supplement Sect. S1.1) but also from 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 $\approx 4° \times 5° \times 19$ 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 (see Supplement Sect. S2).[4]

---

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

[4] Of course, the use of inversion-based X^{CO_2} already creates some dependence of our prior from atmospheric data. In run **ATM**, 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 on X^{CO_2}.

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

Quantity	Value	Meaning
c_p	$3993\,\mathrm{J\,kg^{-1}\,K^{-1}}$	Specific heat capacity of sea water (value at $20\,^\circ\mathrm{C}$)
$k_{\mathrm{Glob}}^{CO_2}$	$16.5\,\mathrm{cm\,h^{-1}}$	$(= 4.58 \cdot 10^{-5}\,\mathrm{m\,s^{-1}})$ Global mean piston velocity of CO_2 (Naegler, 2009)
R_{gas}	$8.3144\,\mathrm{J\,mol^{-1}\,K^{-1}}$	Gas constant
T_0	$273.15\,\mathrm{K}$	Absolute temperature at $0\,^\circ\mathrm{C}$
V_{ideal}	$0.0224136\,\mathrm{m^3\,mol^{-1}}$	Molar volume of an ideal gas
ϱ	$1025\,\mathrm{kg\,m^{-3}}$	Density of sea water

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
$f_{\mathrm{nee}}^{CO_2}$	Land NEE (run **ATM** only)		corr. pix.	lt, seas, var
Technical degrees of freedom:				
$\Delta C_{\mathrm{m}}^{\mathrm{DIC,\,ini}}$	Initial condition of budget equation	Zero	corr. pix.	–
$f_{\mathrm{ini}}^{CO_2}$	Flux pulse for atmospheric initial condition (run **ATM** only)	Zero	3 lat. bands	–

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

A1.2 Carbonate chemistry

The CO_2 partial pressure in the mixed layer is a function of mixed-layer DIC concentration $C_{\mathrm{m}}^{\mathrm{DIC}}$, alkalinity A, temperature T, and salinity S:

$$p_{\mathrm{m}}^{CO_2} = p_{\mathrm{m}}^{CO_2}\left(C_{\mathrm{m}}^{\mathrm{DIC}},\, A,\, T,\, S\right). \tag{A4}$$

In the diagnostic scheme, the relevant dependence is that on $C_{\mathrm{m}}^{\mathrm{DIC}}$, while the influences of A, T, and S are calculated a-priori from driver data (Table 1). While the exact function is a 5th order polynomial (Zeebe and Wolf-Gladrow, 2001), we use existing approximations, as described in the following.

At first, we single out the temperature dependence into a temperature-dependent factor and $p_{\mathrm{m}}^{CO_2}$ at a reference temperature T^{Ref}:

$$p_{\mathrm{m}}^{CO_2} = \beta\left(T,\, T^{\mathrm{Ref}}\right) \cdot p_{\mathrm{m}}^{CO_2}(C_{\mathrm{m}}^{\mathrm{DIC}},\, A,\, T^{\mathrm{Ref}},\, S). \tag{A5}$$

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

$$\beta(T_1, T_2) = \exp\left\{0.0433\,\mathrm{K^{-1}}\,(T_1 - T_2)\right. \tag{A6}$$
$$\left. -4.35 \times 10^{-5}\,\mathrm{K^{-2}}\left(T_1^2 - T_2^2\right)\right\}.$$

Even though this relation has been fitted to data in the temperature range of about 2 to $25\,^\circ\mathrm{C}$, it is used here for all temperatures.

The remaining dependences of $p_{\mathrm{m}}^{CO_2}$ are non-linear, but monotonic and can be linearized to good approximation (Sarmiento and Gruber, 2006). Linearization of the dependence on $C_{\mathrm{m}}^{\mathrm{DIC}}$ is important here in order to be able to use the fast minimization algorithm of Rödenbeck (2005) (Sect. A2.1). We consider deviations (Δ) of the driver variables from temporally constant reference values (superscript "Ref", see Sect. A1.4 below):

$$C_{\mathrm{m}}^{\mathrm{DIC}} = \Delta C_{\mathrm{m}}^{\mathrm{DIC}} + C_{\mathrm{m}}^{\mathrm{DIC,\,Ref}} \tag{A7}$$
$$A = \Delta A + A^{\mathrm{Ref}} \tag{A8}$$
$$S = \Delta S + S^{\mathrm{Ref}} \tag{A9}$$

Linearization of Eq. (A5) around these references gives

$$p_{\mathrm{m}}^{CO_2} = \beta\left(T,\, T^{\mathrm{Ref}}\right) \times \tag{A10}$$
$$\left[p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}} + \frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}}\Delta C_{\mathrm{m}}^{\mathrm{DIC}} + \frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial A}\Delta A + \frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial S}\Delta S\right]$$

The DIC sensitivity has been calculated from the response factors γ^{DIC} by Egleston et al. (2010):

$$\frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}} = \frac{p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}}}{\gamma^{\mathrm{DIC}}}. \tag{A11}$$

The alkalinity sensitivity has been calculated using the approximation

$$\frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial A} = \frac{\partial p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}}}{A^{\mathrm{Ref}}} \times$$
$$\frac{-(A^{\mathrm{Ref}})^2}{\left(2C_{\mathrm{m}}^{\mathrm{DIC,\,Ref}} - A^{\mathrm{Ref}}\right)\left(A^{\mathrm{Ref}} - C_{\mathrm{m}}^{\mathrm{DIC,\,Ref}}\right)} \quad (A12)$$

(Eq. (8.3.17) of Sarmiento and Gruber, 2006, note missing minus sign). The salinity sensitivity (direct effect on chemical equilibrium only) has been set to

$$\frac{p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}}}{\partial S} = \frac{p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}}}{S^{\mathrm{Ref}}} \quad (A13)$$

(Eq. (8.3.5) of Sarmiento and Gruber, 2006). Any seasonal and interannual variations in the proportionalities $\partial p_{\mathrm{m}}^{CO_2}/\partial C_{\mathrm{m}}^{\mathrm{DIC}}$, etc., are neglected. The variations in alkalinity, A, and salinity, S, are fixed according to input data sets (Table 1). Their effect on $p_{\mathrm{m}}^{CO_2}$ 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. A19).

Equation (A10) provides the desired linear link between changes in $p_{\mathrm{m}}^{CO_2}$ and $C_{\mathrm{m}}^{\mathrm{DIC}}$ according to carbonate chemistry. For simplification in the further use, we combine the gas exchange and chemistry formulas: we first define an apparent DIC concentration in equilibrium with the atmosphere,

$$\Delta C_{\mathrm{a}}^{\mathrm{DIC}} = \left(\frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}}\right)^{-1}\left[\frac{p_{\mathrm{a}}^{CO_2}}{\beta\left(T, T^{\mathrm{Ref}}\right)} - p_{\mathrm{m}}^{CO_2,\,\mathrm{Ref}}\right.$$
$$\left. - \frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial A}\Delta A - \frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial S}\Delta S\right] \quad (A14)$$

Using this definition, Eq. (A10) becomes

$$p_{\mathrm{m}}^{CO_2} = p_{\mathrm{a}}^{CO_2} + \beta\left(T, T^{\mathrm{Ref}}\right)\frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}}\left(\Delta C_{\mathrm{m}}^{\mathrm{DIC}} - \Delta C_{\mathrm{a}}\right) \quad (A15)$$

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

$$k^{\mathrm{DIC}} = k^{CO_2} \cdot L^{CO_2}\beta\left(T, T^{\mathrm{Ref}}\right)\frac{\partial p_{\mathrm{m}}^{CO_2}}{\partial C_{\mathrm{m}}^{\mathrm{DIC}}}, \quad (A16)$$

Eqs. (A1) and (A15) can be combined and simplified into

$$f_{\mathrm{ma}}^{CO_2} = k^{\mathrm{DIC}}\varrho\left(\Delta C_{\mathrm{m}}^{\mathrm{DIC}} - \Delta C_{\mathrm{a}}\right), \quad (A17)$$

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

A1.3   Mixed-layer DIC budget

To finally establish the link between the mixed-layer DIC concentration and the ocean-interior sources and sinks, we consider that changes in the spatio-temporal field of dissolved carbon in the mixed layer need to be balanced by the sum of fluxes:

$$\frac{\mathrm{d}}{\mathrm{d}t}C_{\mathrm{m}}^{\mathrm{DIC}} = \frac{1}{h\varrho}\left(-f_{\mathrm{ma}}^{CO_2} + f_{\mathrm{hist}}^{\mathrm{DIC}} + f_{\mathrm{frw}}^{\mathrm{DIC}} + f_{\mathrm{int}}^{\mathrm{DIC}}\right). \quad (A18)$$

Fluxes (carbon exchange per unit horizontal area and unit time) comprise loss through sea–air exchange $f_{\mathrm{ma}}^{CO_2}$, 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 processes (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 MLD product directly based on data that also represents 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^{-1}), while fluxes are volume/area based.

Two specific ocean-internal processes are not considered part of $f_{\mathrm{int}}^{\mathrm{DIC}}$ but are parameterized explicitly[5] in the budget (Eq. A18):

- 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 sea-ice formation/melting. Analogously to the often used salinity normalization (Sarmiento and Gruber, 2006), we assume that freshwater does not contain either carbon or salt (and that all salinity variations are related to freshwater fluxes), and calculate the freshwater effect on DIC from salinity changes:

$$f_{\mathrm{frw}}^{\mathrm{DIC}} = h\varrho\frac{C_{\mathrm{m}}^{\mathrm{DIC,\,Ref}}}{S^{\mathrm{Ref}}}\frac{\mathrm{d}S}{\mathrm{d}t}. \quad (A19)$$

  This will tend to overestimate the effect to the extent that freshwater input does carry carbon.

- The "history" flux $f_{\mathrm{hist}}^{\mathrm{DIC}}$ is part of the carbon exchange between the mixed layer and the underlying ocean. Consider the hypothetical case that no sources or sinks are acting below the mixed layer. Then, as soon as the mixed layer is deepening, it re-entrains water that was left behind previously during mixed-layer shoaling. As mixed-layer concentrations will generally have changed

---

[5] The explicit treatment of these processes is not strictly needed if the only purpose of the scheme was mapping of p^{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. 3.7), and also prepares for the envisaged link to oxygen.

meanwhile, this re-entrainment is equivalent to a carbon flux which is proportional to the difference between the current concentration $C_m^{DIC}(t)$ and that at time t_{prev} when the mixed layer was as deep for the last time. We call this the "history" flux. It can be written as

$$f_{hist}^{DIC}(t) = \varrho \cdot \left(C_m^{DIC}\left(t_{prev}\right) - C_m^{DIC}(t) \right) \cdot \Theta\left(\frac{dh}{dt}\right) \quad (A20)$$

The Heaviside step function Θ ensures that f_{hist} only acts during mixed-layer deepening ($dh/dt > 0$). 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_{int}^{DIC} with the mean sea–air exchange $f_{ma}^{CO_2}$ (plus the mean DIC accumulation in the mixed layer itself). This has been verified numerically, and Sect. S3 proves this balance analytically.

Any changes of the DIC concentration below the mixed layer will lead to additional entrainment fluxes, forming an important part of f_{int}^{DIC}.

Substituting Eq. (A17) into Eq. (A18) gives

$$\frac{d}{dt}\Delta C_m^{DIC} = -\frac{k^{DIC}}{h}\Delta C_m^{DIC} + \frac{k^{DIC}}{h}\Delta C_a^{DIC} \quad (A21)$$
$$+ \frac{1}{h\varrho}f_{hist}^{DIC}\{\Delta C_m^{DIC}\} + \frac{1}{h\varrho}f_{frw}^{DIC} + \frac{1}{h\varrho}f_{int}^{DIC},$$

where we further exploited that the temporally constant reference $C_m^{DIC, Ref}$ disappears from the left-hand side. The reference also cancels out in the history flux (Eq. A20), which is thus a linear functional of $\Delta C_m^{DIC}(t)$. Equation (A21) thus represents a 1st order differential equation in ΔC_m^{DIC} that can be solved (numerically) for given "forcing" f_{int}^{DIC}. At the beginning t_i of the time period, we use the initial condition

$$\Delta C_m^{DIC}(t_i) = \Delta C_a(t_i) + \Delta C_m^{DIC, ini}. \quad (A22)$$

Up to a constant offset (determined by $C_m^{DIC, Ref}$) and an initial transient (determined by $\Delta C_m^{DIC, ini}$, see Supplement Sect. S1.2), Eq. (A21) thus provides a linear dependence of C_m^{DIC} on f_{int}^{DIC} for every pixel.

A1.4 Reference values

For the reference values $C_m^{DIC, Ref}$, A^{Ref}, T^{Ref}, and S^{Ref}, we use the long-term averages of the respective driver data sets (Table 1) for each pixel. The partial pressure reference value $p_m^{CO_2, Ref}$ in Eq. (A10) formally follows from the functional relation

$$p_m^{CO_2, Ref} = p_m^{CO_2}\left(C_m^{DIC, Ref}, A^{Ref}, T^{Ref}, S^{Ref}\right). \quad (A23)$$

Solving this equation would be numerically involved but fortunately is not needed in practice: Firstly, $C_m^{DIC, Ref}$ only determines the absolute levels of DIC (Eq. A7) not considered here, but does not directly affect our target quantities $p_m^{CO_2}$ or sea–air fluxes (as it cancels out both from Eqs. (A15), (A17), and (A21); we nevertheless use a data-based field because a minor influence of $C_m^{DIC, Ref}$ arises through Eq. (A12) (alkalinity sensitivity) and Eq. (A19) (freshwater effect)). Secondly, the partial pressure reference value $p_m^{CO_2, Ref}$ has no effect either: Though changes in $p_m^{CO_2, Ref}$ shift the long-term mean of ΔC_a through Eq. (A14), the same shift then happens to ΔC_m^{DIC} through Eq. (A21) with Eq. (A22), such that it again cancels out from Eqs. (A15) and (A17). The choice of $p_m^{CO_2, Ref}$ is thus arbitrary (we use $p_a^{CO_2}(t_i)$ as for numerical reasons a value on the order of the actual $p_m^{CO_2}$ is desirable).

A2 Inversion

The estimation procedure of this paper is derived from the linear Bayesian atmospheric 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 parameterizations (Sect. A2.2) and by using the p^{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 c_{obs} (time series at observation sites, see Table 1). The inversion calculation seeks fluxes f that lead to the best match between c_{obs} and corresponding modelled mixing ratios $c_{mod}(f)$, in the sense that the value of the cost function

$$J_c = \frac{1}{2}(c_{obs} - c_{mod})^T Q_c^{-1}(c_{obs} - c_{mod}) \quad (A24)$$

is minimal (superscript T means vector transpose). The (diagonal) matrix Q_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 simulated by a numerical transport model (here, the TM3 model, Heimann and Körner, 2003), which can formally be written as

$$c_{mod} = A f + c_0 \quad (A25)$$

with the transport matrix A and the initial concentration 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 $\boldsymbol{x}$:

$$f(\boldsymbol{x}) = f_{\text{pri}} + \mathbf{F}\boldsymbol{x}. \qquad (A26)$$

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

$$J = J_c + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}. \qquad (A27)$$

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

The specification of the flux model elements f_{pri} and $\mathbf{F}$ is detailed in Rödenbeck (2005). The flux is first written as a sum of contributions from several components:

$$f = f_{\text{ma}} + f_{\text{nee}} + f_{\text{fos}} + f_{\text{ini}} \qquad (A28)$$

(f_{ma} = sea–air flux, f_{nee} = terrestrial net ecosystem exchange (NEE), f_{fos} = fossil fuel burning emissions, f_{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. (A26), i.e.

$$f(\boldsymbol{x}) = \sum_i \left(f_{\text{pri},i} + \mathbf{F}_i \boldsymbol{x}_i \right), \qquad (A29)$$

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 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 f_{ma}) while suppressing flux adjustments elsewhere.

The detailed flux model settings for the components f_i of the surface-to-atmosphere CO_2 flux in the "Jena inversion" are

- f_{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.

- $f_{\text{nee,lt}} + f_{\text{nee,seas}} + f_{\text{nee,var}}$: 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).

- f_{fos}: fossil fuel burning. Only a fixed (prior) term, as in the Jena inversion v3.2 (yearly emissions from the EDGAR v4.0 database, update of Olivier and Berdowski, 2001).

- f_{ini}: purely technical component. Flux pulse at beginning of inversion period to adjust initial atmospheric mixing ratio field (see Rödenbeck, 2005).

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

A2.2 Extension: using ocean parameterization

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 f_{ma} is not estimated directly but implemented as a function of the ocean-internal flux f_{int} using the parameterizations of Sect. A1. Then, instead of f_{ma}, the ocean-internal flux f_{int} is adjusted to match the data constraint. The detailed flux-model specifications for f_{int} are similar to those for the sea–air flux in the atmospheric inversion:

- The a-priori state has been defined by $f_{\text{int,pri}}^{\text{DIC}} = 0$ (no internal sources or sinks). The "error" of this prior model $f_{\text{int,pri}}^{\text{DIC}} = 0$ is then 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 (see Supplement S1.3). The a-priori uncertainties are scaled such that the implied a-priori uncertainty of the globally integrated sea–air flux is 1200 Tmol yr^{-1} (with respect to 3-monthly anomalies).

- Bayesian a-priori spatial and temporal correlations have been implemented that dampen variability in $f_{\text{int}}^{\text{DIC}}$ on scales smaller than around 640 km (longitude), 320 km (latitude), and about 2 weeks (time), Fig. A1. The chosen e-folding lengths are one-third of the ocean flux correlations in the standard Jena Inversion v3.4 (update of
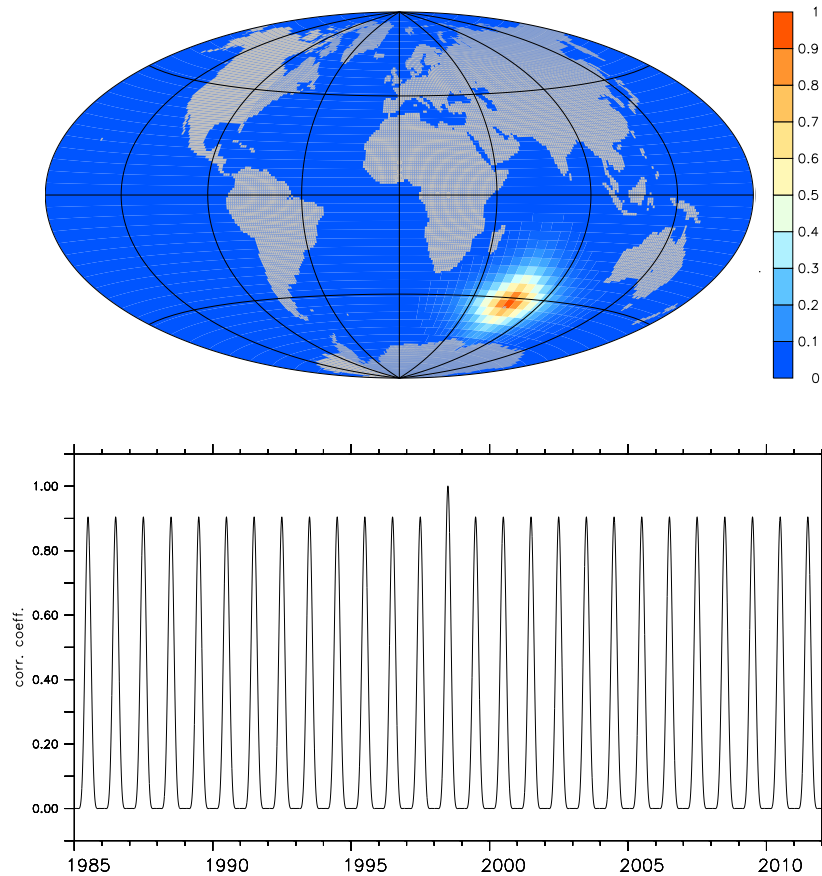
Fig. A1. Top: map of a-priori correlation coefficients with respect to an example pixel at (70° E, 49° S), illustrating how inverse adjustments to f_{int}^{DIC} de-correlate with distance. Bottom: a-priori correlation coefficients with respect to an example time step in the middle of the inversion period; in addition to the de-correlation away from the example time step, correlations rise again each year due to the dominance of the seasonal cycle.

Rödenbeck, 2005)[6], and similar to those of p^{CO_2} found from correlation analysis along cruise tracks by Jones et al. (2012a). The correlations are seen as a mathematical device to stabilize the estimation by suppressing noise which results, for example, from unequal sampling (they represent the main mechanism by which the information from the discrete atmospheric or oceanic data points is extrapolated into space and time, see Fig. 5 for illustration). The correlation scales need to be large enough to accomplish this extrapolation, but small enough not to unduly dampen actual signals. Correlations shorter than in the standard atmospheric inversion have been chosen to reflect the higher spatial detail in the p^{CO_2} data. For example, this is important in the North Atlantic where shorter correlations are needed to prevent a large influence of coastal data on

the regional average (test not shown). Shorter correlations also prevent implausible leakage of interannual variability of the tropical Pacific northward and southward (see Rödenbeck et al., 2013). On the other hand, longer correlations may have been beneficial in the datasparse Southern Ocean as they lead to slightly better performance in synthetic-data tests (analogously to Appendix B, test not shown). In any case, simple distance-based de-correlation necessarily represents a crude approximation to the unknown real correlation structure of the ocean-internal sources/sinks determined by complex processes. To ensure that this approximation does not critically determine the result, runs with different de-correlation lengths are included in the set of sensitivity cases in our robustness test (Sect. 3.3).

In the time domain, we additionally dampen all non-seasonal variability by a factor 16 with respect to the seasonal variability (this is accomplished by implementing the time dependence of f_{int}^{DIC} as a Fourier series and reducing the a-priori uncertainty of the non-seasonal

---

[6] As an additional minor technical change, the basic spatial grid of the adjustable degrees of freedom is every pixel of the TM3 transport model grid, rather than aggregates of 4 neighbouring pixels, reflecting the higher spatial resolution of the data.

modes, see Rödenbeck, 2005). Such tighter a-priori uncertainties for the interannual variations are consistent with the expectation that their amplitudes are smaller than those of the seasonal variations. Moreover, they dampen spurious interannual variations that may otherwise arise from the very different coverage of the SO-CAT data in the individual years, and extrapolate the mean seasonal cycle of f_{int}^{DIC} throughout the inversion period. As the chosen dampening factor is largely arbitrary, it has also been varied in sensitivity tests.

A further degree of freedom is $\Delta C_m^{DIC, ini}$, needed only for technical reasons: Through adjusting $\Delta C_m^{DIC, ini}$, the initial conditions of the carbon budget equation can become consistent with the data (see Supplement Sect. S1.2). Adjustments to $\Delta C_m^{DIC, ini}$ are spatially resolved with the same spatial correlations as f_{int}, but constant in time; the a-priori value is zero.

In run **ATM**, there are further degrees of freedom not needed in run **SFC**: as in the atmospheric inversion, land fluxes are adjustable as specified in Sect. A2.1, as well as the adjustable atmospheric initial condition.

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 $\boldsymbol{c}_{obs}$, all selected partial pressure values from the SOCAT database (Table 1) are collected into a vector $\boldsymbol{p}_{obs}$. A corresponding vector $\boldsymbol{p}_{mod}$ of modelled partial pressures is formed by sampling the gridded $p_m^{CO_2}$ field of the model at the locations and times of the measurements (i.e. the modelled value is that found in the gridcell/timestep which encloses the location/time of the corresponding observation). If several SO-CAT data points fall into the same gridcell and timestep, they are averaged together and only form a single element in $\boldsymbol{p}_{obs}$ and $\boldsymbol{p}_{mod}$.

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

$$ J_p = \frac{1}{2} \left(\boldsymbol{p}_{obs} - \boldsymbol{p}_{mod} \right)^T \mathbf{Q}_p^{-1} \left(\boldsymbol{p}_{obs} - \boldsymbol{p}_{mod} \right). \tag{A30} $$

The covariance matrix $\mathbf{Q}_p$ is chosen diagonal as well; the uncertainty for every individual pixel with two or more SOCAT data points is set to $10 \, \mu atm$ (for pixels with a single data point to $\sqrt{2} \times 10 \, \mu atm$). The fit to the SOCAT data is then done in the same way as to the atmospheric data, just with J_c replaced by J_p (those degrees of freedom that do not depend on $p_m^{CO_2}$ – land fluxes and atmospheric initial condition – are unconstrained in this inversion and just remain at their a-priori values).

For the case combining atmospheric and oceanic constraints (Sect. 3.6), 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" on the basis of a biogeochemical process model simulation (NEMOv2.3 with PlankTOM5, Buitenhuis et al., 2010), mapped to the grid of our scheme. The seasonal cycle of this field is similar to the observations at most locations except for the high latitudes, which is why we subtract a mean seasonal cycle and replace it by that of Takahashi et al. (2009). This combined field then contains variations similar to reality on all relevant timescales (though variations on the fastest (weekly to daily) timescales are smaller than observed). (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.

In terms of mean seasonality considered here, the retrieval (blue) fits the "known truth" (violet) closely (Fig. B1). Most region-to-region differences are reproduced. The largest differences are found in the North Pacific, and in parts of the temperate Southern Hemisphere. If p^{CO_2} is averaged over open ocean only (not shown), Northern Hemisphere and tropical mismatches further reduce, in particular with the essentially perfect match in the temperate North Atlantic.

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. 2.3). 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.
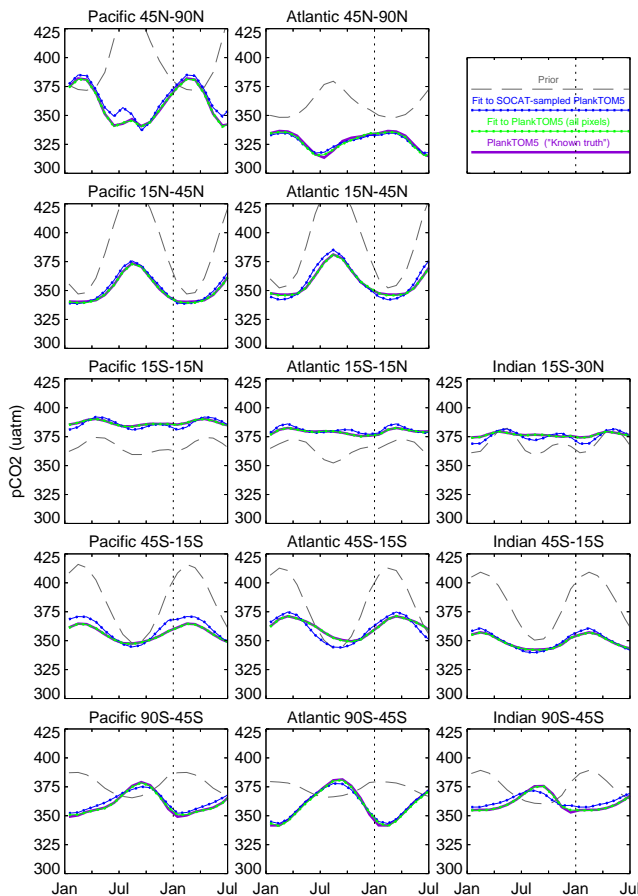
Fig. B1. Testing the capacity to retrieve the p^{CO_2} field from the SOCAT data set: fit of the diagnostic scheme to the p^{CO_2} field simulated by the PlankTOM5 biogeochemical model (Buitenhuis et al., 2010) (combined with seasonality from Takahashi et al. (2009)) sub-sampled at locations/times of the SOCAT measurements (blue), compared to the simulated p^{CO_2} field itself ("known truth", violet). For further comparison, a fit of the scheme to the simulation at every grid cell and time step ("complete information", green) and the Bayesian prior ("no data information", thin dashed grey, partially off-scale) is given. The p^{CO_2} fields have been averaged over the regions as in Fig. 7.

– 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 real-

ity but not contained in our 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 (see Supplement Fig. S7.4) 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 the high seasonal cycle from the pixels south of this area.

Seasonal data coverage is also limiting in the tropical Indian, the other problematic region (see Supplement Fig. S7.4).

Appendix C

Calculating phosphate concentrations from f_{int}^{DIC}

Considering the ocean-internal processes affecting mixed-layer concentrations of biogeochemical tracers, 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.

The biological contribution to the sources and sinks of carbon (both in the mixed layer and the ocean interior) is directly linked to sources and sinks of phosphate, and assumed to proceed in Redfield proportions $r_{C:P} = 106$. The remainder is proportional to the differences between the pre-formed carbon (or nutrient) concentrations of arriving water parcels and those of the destination mixed-layer. Considering only their variability as relevant here, these differences are assumed to be either small or close to Redfield proportions as well (this is equivalent to assuming that the internal sources/sinks of the tracer $C_m^{DIC*} = C_m^{DIC} - r_{C:P} C_m^{PO_4}$ (Gloor et al., 2003) have no variability). Then the variations of ocean-internal phosphate sources and sinks ($f_{int}^{PO_4}$) are proportional to the variations of f_{int}^{DIC}. We thus calculated the potential seasonal changes in the mixed-layer PO_4 concentration using a PO_4 budget equation as Eq. (A18) (including history and freshwater fluxes), except that (1) the ocean-internal flux was reduced by the Redfield ratio ($f_{int}^{PO_4} = f_{int}^{DIC}/r_{C:P}$), (2) the long-term mean was subtracted from $f_{int}^{PO_4}$, and (3) there was no sea–air flux of phosphate ($f_{ma}^{PO_4} = 0$). As the initial condition for the budget equation is unknown and phosphate does not relax towards an atmospheric value, calculated $C_m^{PO_4}$ contains an ar-

bitrary offset, such that we only compare the deviation from
the temporal mean.

**Supplementary material related to this article is
available online at:**

atmospheric CO_2, Science, 284, 464–467, 1999.

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, Bull. Am. Met. Soc., 77, 437–471, 1996.

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.

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.

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 B, 58, 390–417, 2006.

Law, R. M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I., Bergmann, D. J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P. J., Christensen, J. H., Delage, F., Denning, A. S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S. R., Kleist, J., Krol, M. C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P. K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO_2: Experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cy., 22, GB3009, doi:10.1029/2007GB003050, 2008.

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

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. Cy., 24, GB4007, doi:10.1029/2009GB003599, 2010.

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.

Lefèvre, N., Watson, A. J., and Watson, A. R.: A comparison of multiple regression and neural network techniques for mapping in situ p^{CO_2} data, Tellus B, 75, 375–384, 2005.

Lourantou, A. and Metzl, N.: Decadal evolution of carbon sink within a strong bloom area in the subantarctic zone, Geophys. Res. Lett., 38, L23608, doi:10.1029/2011GL049614, 2011.

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

Metzl, N., Corbière, A., Reverdin, G., Lenton, A., Takahashi, T., Olsen, A., Johannessen, T., Pierrot, D., Wanninkhof, R., Ólafsdóttir, S. R., Olafsson, J., and Ramonet, M.: Recent acceleration of the sea surface fCO2 growth rate in the North Atlantic subpolar gyre (1993–2008) revealed by winter observations, Global Biogeochem. Cy., 24, GB4004,

doi:10.1029/2009GB003658, 2010.

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.

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 Biogeochem. Cy., 22, GB3011, doi:10.1029/2007GB003065, 2008.

Naegler, T.: Reconciliation of excess ^{14}C-constrained global CO_2 piston velocity estimates, Tellus B, 61, 372–384, 2009.

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

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.

Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in: The Climate System, Berdowski, J., Guicherit, R., and Heij, B. J., A. A. Balkema Publishers/Swets and Zeitlinger Publishers, Lisse, The Netherlands, ISBN 9058092550, 33–78, 2001.

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.

Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C. L., Akl, J., Alin, S. R., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., Chen, A., Cosca, C., Fassbender, A. J., Feely, R. A., González-Dávila, M., Goyet, C., Hardman-Mountford, N., Heinze, C., Hood, M., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Jones, S. D., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G.-H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Santana-Casiano, J. M., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Tjiputra, J., Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J., Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT), Earth Syst. Sci. Data Discuss., 5, 735–780, doi:10.5194/essdd-5-735-2012, 2012.

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 B, 51, 213–232, 1999.

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.

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.

Sabine, C., Maenner, S. and Sutton, A.: High-resolution ocean and atmosphere pCO2 time-series measurements from mooring TAO140W. http://cdiac.esd.ornl.gov/ftp/oceans/Moorings/TAO140W/. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, 2010.

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

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

Sarmiento, J. L., Gloor, E., Gruber, N., Beaulieu, C., Jacobson, A. R., Fletcher, S. E. M., Pacala, S., and Rodgers, K.: Trends and Regional Distributions of Land and Ocean Carbon Sinks, Biogeosciences, 7, 2351–2367, doi:10.5194/bg-7-2351-2010, 2010.

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

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, Science, 316, 1732–1735, 2007.

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 Biogeoch. Cycl., 21, GB2015, doi:10.1029/2006GB002784, 2007.

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 p^{CO_2} and net sea-air CO_2 flux over the global oceans, Deep-Sea Res. II, 56, 554–577, 2009.

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., Moulin, C., Bakker, D., González-Dávila, M., Johannessen, T., Körtzinger, A., Lüger, H., Olsen, A., Omar, A., Padin, X., Ríos, A., Steinhoff, T., Santana-Casiano, M., Wallace, D. W. R., and Wanninkhof, R. H.: Estimating the monthly p^{CO_2} distribution in the North Atlantic using a self-organizing neural network, Biogeosciences, 6, 1405–1421, doi:10.5194/bg-6-1405-2009, 2009.

Valsala, K. V. and Maksyutov, S.: Simulation and assimilation of global ocean p^{CO_2} and air-sea CO_2 fluxes using ship observations of surface ocean p^{CO_2} in a simplified Biogeochemical offline model, Tellus B, 62, 821–840, doi:10.1111/j.1600-0889.2010.00495.x, 2010.

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

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeosciences Discuss., 9, 10961–11012, doi:10.5194/bgd-9-10961-2012, 2012.

Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas., Mar. Chem., 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, Mar. Chem., 8, 347–359, 1980.

While, J., Totterdell, I., and Martin, M.: Assimilation of p^{CO_2} Data Into a Global Coupled Physical-Biogeochemical Ocean Model, J. Geophys. Res., 117, C03037, doi:10.1029/2010JC006815, 2012.

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

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