

Woolf D. K. (Orcid ID: 0000-0003-4469-553X)
Shutler J.D. (Orcid ID: 0000-0002-8018-123X)
Goddijn-Murphy L. (Orcid ID: 0000-0003-4820-8313)
Watson A. J. (Orcid ID: 0000-0002-9654-8147)
Nightingale P. D. (Orcid ID: 0000-0001-7177-5469)
Piskozub J. (Orcid ID: 0000-0003-3386-6604)
Yelland M. J (Orcid ID: 0000-0002-0936-4957)
Holding T. (Orcid ID: 0000-0002-9386-6414)
Girard-Ardhuin F. (Orcid ID: 0000-0001-7819-7665)
Grouazel A. (Orcid ID: 0000-0002-9113-9467)
Warren M. (Orcid ID: 0000-0002-7941-5254)
Land P. E. (Orcid ID: 0000-0001-7518-8683)
Torres R. (Orcid ID: 0000-0002-9035-2637)
Prytherch J. (Orcid ID: 0000-0003-1209-289X)
Ardhuin F. (Orcid ID: 0000-0002-9309-9681)

Key Uncertainties in the Recent Air-Sea Flux of CO₂

D. K. Woolf¹, J. D. Shutler², L. Goddijn-Murphy³, A. J. Watson², B. Chapron⁴, P. D. Nightingale⁵, C. J. Donlon⁶, J. Piskozub⁷, M. J. Yelland⁸, I. Ashton⁹, T. Holding², U. Schuster², F. Girard-Ardhuin⁴, A. Grouazel⁴, J-F. Piolle⁴, M. Warren⁵, I. Wrobel-Niedzwiecka⁷, P. E. Land⁵, R. Torres⁵, J. Prytherch^{8*}, B. Moat⁸, J. Hanafin^{4†}, F. Ardhuin⁴, and F. Paul⁴

¹International Centre for Island Technology, Heriot-Watt University, UK.

²College of Life and Environmental Sciences, University of Exeter, UK.

³University of the Highlands and Islands, UK.

⁴Ifremer, Univ. Brest, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, Brest, France.

⁵Plymouth Marine Laboratory, UK.

⁶European Space Agency, The Netherlands.

⁷Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland.

⁸National Oceanography Centre, Southampton, UK.

⁹College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK.

* now at Department of Meteorology, Stockholm University, Sweden.

† now at National University of Ireland Galway, Ireland.

Corresponding author: David Woolf (d.k.woolf@hw.ac.uk)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2018GB006041

Abstract

The contemporary air-sea flux of CO₂ is investigated by the use of an air-sea flux equation, with particular attention to the uncertainties in global values and their origin with respect to that equation. In particular, uncertainties deriving from the transfer velocity and from sparse upper ocean sampling are investigated. Eight formulations of air-sea gas transfer velocity are used to evaluate the combined standard uncertainty resulting from several sources of error. Depending on expert opinion, a standard uncertainty in transfer velocity of either ~5% or ~10% can be argued and that will contribute a proportional error in air-sea flux. The limited sampling of upper ocean fCO₂ is readily apparent in the Surface Ocean CO₂ Atlas (SOCAT) databases. The effect of sparse sampling on the calculated fluxes was investigated by a bootstrap method; i.e. treating each ship cruise to an oceanic region as a random episode and creating 10 synthetic datasets by randomly selecting episodes with replacement. Convincing values of global net air-sea flux can only be achieved using upper ocean data collected over several decades, but referenced to a standard year. The global annual referenced values are robust to sparse sampling, but seasonal and regional values exhibit more sampling uncertainty. Additional uncertainties are related to thermal and haline effects and to aspects of air-sea gas exchange not captured by standard models. An estimate of global net CO₂ exchange referenced to 2010 of $-3.0 \pm 0.6 \text{ Pg C yr}^{-1}$ is proposed, where the uncertainty derives primarily from uncertainty in the transfer velocity.

Plain Language Summary

The oceanic carbon sink reduces the rate of accumulation of CO₂ in the atmosphere, but is also responsible for the acidification of the ocean. One method of estimating the size of the oceanic carbon sink depends on a calculation of upward and downward flows of CO₂ at the sea surface. This study revisits this calculation using updated knowledge of the transfer processes at the sea surface and the results of a large international collaborative effort (SOCAT) to collect and compile measurements of CO₂ in the upper ocean. Greater sampling of the oceans improves estimates, but direct calculation in each year is not practical. Instead, we calculate fluxes in a recent year (2010) using upper ocean measurements of CO₂ over many years. The remaining uncertainty is dominated by limited knowledge of the efficiency of stirring of gas across the sea surface, the air-sea transfer velocity. The study suggests a relatively large downward flow of CO₂ into the ocean compared to previous applications of this method and other methods to estimate the oceanic carbon sink. Increased knowledge is rewarded by reduced uncertainty in the net global flux; that flux is estimated at $-3.0 \pm 0.6 \text{ Pg C yr}^{-1}$. Further understanding of transfer velocities and better sampling may reduce the uncertainty in the future.

Key Points:

- Increased understanding of air-sea gas transfer processes and better sampling of the upper ocean enables higher confidence in calculations of air-sea CO₂ fluxes.
- The calculations imply a relatively large global net air-to-sea flux of $-3.0 \text{ Pg C yr}^{-1}$ (referenced to 2010).
- This flux is known within 0.6 Pg C yr^{-1} , where uncertainty in air-sea transfer velocity is the largest contribution to the combined uncertainty.

1 Introduction

The air-sea flux of carbon dioxide is a critical part of the climate system (Ciais et al., 2013; Rhein et al., 2013) and a major factor in the development of the ocean (e.g. ocean acidification). An ocean sink is an important component of the anthropogenic perturbation of the global carbon cycle. Several methods are recognised to give fairly consistent estimates of the global time-averaged ocean sink or uptake rate (Rhein et al., 2013). In particular, an inventory of anthropogenic carbon in the ocean implies an uptake rate of $-2.3 \text{ Pg C yr}^{-1}$ (in the range of -1.7 to -2.9) from 2000 to 2010, while an estimate based on atmospheric O_2/N_2 measurements implies an uptake rate for the same decade of $-2.5 \pm 0.6 \text{ Pg C yr}^{-1}$. Also, Takahashi et al. (2009) estimate a sink of $-2.0 \pm 1.0 \text{ Pg C yr}^{-1}$ “normalized to the year 2000” based on a calculated contemporary air-sea flux of $-1.6 \pm 0.9 \text{ Pg C yr}^{-1}$ corrected by an estimated pre-industrial flux of $+0.4 \pm 0.2 \text{ Pg C yr}^{-1}$. From these three estimates, there is a high level of confidence in a global uptake between -1.0 and -3.2 PgC yr^{-1} (Rhein et al., 2013). Estimates for the preceding decades are generally lower and a trend of an increasing sink with time is implied. Note that we will always represent a net downward flux as a negative value, for more consistency with the convention for uptake.

The estimates by Takahashi et al. (2009) are based on a method of estimating the contemporary air-sea flux of carbon dioxide through an air-sea flux equation that calculates the local air-sea gas flux as the product of an exchange coefficient and a partial pressure difference. (This is broadly equivalent to calculating the flux as a product of transfer velocity and an effective concentration difference - see Eqn 1 below - but cannot deal satisfactorily with vertical temperature gradients (Woolf et al., 2016)). The method depends on both estimates of exchange coefficient (the product of transfer velocity and solubility) and sufficient measurements of CO_2 in and above the ocean. It follows that improved knowledge of key factors could reduce the uncertainty of the air-sea flux, both globally and more locally. It is generally understood (e.g. Takahashi et al., 2009) that both the parameterisation of the transfer velocity and the relative sparsity of measurements of upper ocean dissolved CO_2 contribute substantially to the uncertainty of the global flux. The air-sea transfer velocity remains an active research topic (Garbe et al., 2014), while the collection, quality control and release of CO_2 data has grown and improved with the formation and activity of the SOCAT community (Bakker et al., 2014). In this paper, we investigate if the growing knowledge of both CO_2 concentrations in the upper ocean and the air-sea gas transfer velocity are yet sufficient to reduce the uncertainty in the global flux.

The results were produced using FluxEngine, which is described in detail by Shutler et al. (2016). Raw data on atmospheric and oceanic CO_2 concentrations and other environmental data (particularly wind speed, temperature and salinity) are ingested into FluxEngine, which then calculates fluxes based on parameterisations. Details of the data and parameterisations used in this study are explained in the methods section. One importance of the “air-sea flux method” is that it can be applied to estimate far more local fluxes than other methods. The calculations made with FluxEngine are at a monthly and one degree (latitude \times longitude) resolution, but in this paper we will restrict the presentation of results to integrations over four large oceanic regions and their total (“global values”).

We use a standard equation for the net air-sea flux, F , of a gas:

$$F = -K_w (C_a/H - C_w) \quad (1)$$

Here H is Henry’s law constant and K_w is the transfer velocity. We specify that the calculations are for poorly soluble gases. In this case, we can satisfactorily substitute k_w for K_w , where k_w is a water-side transfer velocity (Garbe et al., 2014). In the next section, we will

briefly summarise the available options for k_w and the sources of data required for the calculations. In addition, the concentration difference is to a good approximation only across a marine microlayer, the mass boundary layer (MBL) at the sea surface, with a concentration at the top of the MBL, $C_I = C_a/H$, and a concentration at the base, $C_M = C_w$. The concentration anywhere within the ocean is the product of a solubility and a fugacity, fCO_2 , but as solubility varies with temperature and salinity, all calculations of flux should ultimately be in terms of concentrations.

2 Methodology

2.1 Parameterisation of transfer velocity

A traditional wind-speed-dependent transfer velocity is parameterized as

$$k_w = (Sc/660)^{-1/2} (c_0 + c_1U + c_2U^2 + c_3U^3) \quad (2)$$

where U is the instantaneous ten-metre-elevation wind speed and Sc is the Schmidt number of the dissolved gas. (Since the effect of the wind at the surface is sensitive to atmospheric stability, wind speeds, U , are corrected to an equivalent value in a neutrally stable atmosphere). These transfer velocities can be calculated in a computationally efficient method as described in the Supplementary Information.

The polynomial expressions considered here for k_w are summarised in Table 1. The expressions are shown graphically in the Supplementary Information (Figure S1). These expressions are only a few of those found experimentally, but provide a reasonable test of the uncertainties in flux resulting from uncertainties in transfer velocity. We consider mainly parameterisations based on “dual tracer experiments” (DTEs). Estimates based on DTEs depend on measurements of the relatively rapid decline in dissolved ^3He relative to SF_6 as a result of the much higher transfer velocity of ^3He . Calculation of the transfer velocity for any single gas, including CO_2 , relies on the Schmidt-number dependence implicit in Eqn. 2. A large number of DTEs have been conducted (e.g. Nightingale et al. (2000) and Ho et al. (2006, 2007 and 2011)). Parameterisations of transfer velocity based on DTEs are widely used, but are not the only ones available. A wider spectrum of expert opinion on transfer velocity follows from including parameterisations based on different methods. Here, we include a single parameterisation (McGillis et al., 2001) based on direct micrometeorological measurements of CO_2 flux.

Table 1. Polynomial relationship of k_w to wind speed as described by Eqn. 2 for various studies.

Model Number	Authors	Identifier	Methodology	c_0	c_1	c_2	c_3
1	Ho et al., 2006	quad06	DTE	0	0	0.254	0
2	Ho et al., 2007	cubic07	DTE	0	0	0	0.0162
3	Ho et al., 2007	quadpc07	DTE	3.6	0	0.231	0
4	Ho et al., 2007	cubicpc07	DTE	9.2	0	0	0.0124
5	Smith et al., 2011	AFCquad	DTE	0	0	0.299	0
6	Wanninkhof et al., 2009	W2009	DTE	3	0.1	0.064	0.011
7	Nightingale et al., 2000	N2000	DTE	0	0.318	0.212	0
8	McGillis et al., 2001	M2001	Direct Flux	3.3	0	0	0.026

2.2 Data Sources and Processing

As described by Shutler et al. (2016), we have developed a flexible processing system, FluxEngine, for the calculation of air-sea gas fluxes. The system can ingest a large number of data sets and the results will depend for example on the choice of sea surface temperature product. FluxEngine also includes a number of switches that modify the calculations performed. The first global flux products with a $1^\circ \times 1^\circ$ spatial resolution produced by FluxEngine used the same data inputs (based on approximately 3 million surface ocean values of partial pressure) as used by Takahashi et al. (2009), but interpolated to a new grid. That option allowed a verification against the fields published by Takahashi et al. (2009), which showed that the net global flux differed by $<1\%$ from the original publication (Shutler et al., 2016). In parallel, the development of the SOCAT community has provided uniform, quality controlled datasets for surface ocean fCO_2 (Pfeil et al., 2013) containing progressively more data values. Recent papers by Woolf et al. (2016) and Goddijn-Murphy et al. (2015) describe the principles and practical steps involved in reanalysing upper-ocean CO_2 data from SOCAT synthesis data files (Pfeil et al., 2011; Bakker et al., 2014; Bakker et al., 2016) to prepare them for synoptic scale flux calculations. The version of FluxEngine (version 2) and data used in this study are included in Holding et al. (2018), which also provides links to the software. In this paper, we present calculations referenced to 2010, using an enlarged dataset on upper ocean CO_2 (Version 4 of SOCAT; 18.5 million values, 1957-2015) and various other datasets specifically for each calendar month of 2010.

A major difference in principle between the calculations presented by Takahashi et al. (2009) and those presented here is that the difference in concentration is carefully calculated with due regard to gradients of temperature and salinity (Woolf et al., 2016). Since many years of upper ocean CO_2 data are used and secular changes in temperature are significant to the calculations at the base of the MBL, temperatures are required for all months and years corresponding to CO_2 data, in addition to the 12 months of the reference year, 2010. In this study, we use National Climatic Data Center Optimum Interpolation Sea Surface Temperature (NCDC OISST) values, which are available from 1981 to present (<https://www.ncdc.noaa.gov/oisst>). OISST are a “Reynolds SST” corrected to buoy measurements and thus a “sub-skin temperature”. These temperatures were used directly for calculations at the base of the MBL, while 0.17 K was subtracted to give an appropriate skin temperature based on the estimates of Donlon et al. (2002).

In principle, salinity gradients introduce similar complications to temperature gradients. However, the sensitivity of CO_2 concentrations to salinity is much weaker than to temperature and the effects are correspondingly weaker (Woolf et al., 2016). Additionally, while trends in oceanic temperature are substantial, those in salinity are relatively weak. These facts justify the relatively simple approach to salinity in this study. Unlike temperature, we do not re-compute in-situ SOCAT fCO_2 using gridded values of salinity (Goddijn-Murphy et al., 2015). Climatological values of surface salinity from Takahashi et al. (2009) are used for calculations of solubility and transfer velocity in the reference year. Uncertainty in the haline skin effect is large (Yu, 2010; Zhang & Zhang, 2012), but the effect is of order of 0.1 PSU. Climatological surface salinities are assumed to be “sub-skin” values, while 0.1 PSU is added for calculations at the top of the MBL.

As described by Goddijn-Murphy et al. (2015), we reprocess SOCAT values in two steps to give individual values of fugacity in the reference year. First, the equivalent fugacity is calculated at an accepted value of sub-skin SST in the month and location (on a 1 degree \times 1 degree grid of the surface ocean). Secondly, a secular trend is applied to fugacities to give a value in the reference year in the same calendar month and location. In many cases, values in

a particular calendar month and location of the reference year may originate from several different years and various cruises. In the original version (Goddijn-Murphy et al., 2015), these values were averaged in each grid cell without weighting, but in this study contributions from each year were weighted by the number of cruises contributing to that cell, which is consistent with the weighting method used in Sabine et al., (2013). A key uncertainty relates to the application of a secular trend. A study of SOCAT Version 2 supports a steady trend of $1.5 \mu\text{atm yr}^{-1}$ (Zeng et al., 2014), but the trend will introduce uncertainty and a part of our study (Section 2.3d) addressed this issue.

In spite of the use of multi-year data, there are many permutations of calendar month, latitude and longitude ($1^\circ \times 1^\circ$) cells where no data on upper ocean fugacity is available. Data coverage is adequate for some oceanic areas (e.g. Wróbel & Piskozub, 2016), but a lack of data in some cells is common in regional and global studies of CO_2 flux. Some method of mapping (interpolation and extrapolation of fragmented data) is necessary. The SOCOM study (Rodenbeck et al., 2015) compared many approaches to this mapping, including the use of various numerical and statistical models. That study showed that while there are some differences, the integrated regional CO_2 fluxes are not highly sensitive to the mapping method. This context is important in relation to the absolute values of global and regional fluxes reported here since our method is based on that described by Goddijn-Murphy et al. (2015), which was included in the SOCOM study. Our mapping is based on simple geospatial methods, that is, interpolation is based on proximity only, not on any inferred relationship with other data types, since the robustness of such relationships is unknown. The standard method is “ordinary block kriging”, but another method must be used where this method fails because a variogram cannot be computed (Section 4 of Goddijn-Murphy et al., 2015). Kriging methods have the merit that they are statistically unbiased and do not require a “baseline” or “prior” value that might skew the results, but instead are entirely based on data. In this study, the greater number of data in SOCAT Version 4 made the kriging method more reliable, but it could not be used for the month of November and inverse distance interpolation had to be used instead for that month.

Flux calculations also require values of wind speed and atmospheric CO_2 values. This study takes wind speeds from the European Space Agency multi-sensor merged and calibrated GlobWave dataset and values of the molar fraction of CO_2 in the dry atmosphere from the NOAA Marine Boundary Layer reference dataset (<https://www.esrl.noaa.gov/gmd/ccgg/mbl/index.html>). These were converted to $p\text{CO}_2(\text{air})$ using the formulation of Weiss (1974), as implemented in Shutler et al., (2016), and sea level atmospheric pressure from the NCEP/NCAR Reanalysis dataset (Kalnay et al., 1996). Atmospheric “ C_a/H ” (from Eqn. 1) is then calculated as described in Section 2 of Woolf et al. (2016), taking care to use “skin” values of temperature and salinity.

The focus of this study is the uncertainty of fluxes integrated over regions and globally, but for illustration we show a typical global map of the $1^\circ \times 1^\circ$ resolution fluxes in Figure 1 below. The specific example shown is using the transfer velocity parameterisation of Nightingale et al. (2000) - Model 7 in Table 1 – and defaults for all other processing described below. The general pattern of fluxes is quite similar to many other studies (see for example, Roobaert et al., 2018) but with some local differences and a subtle general shift towards more predominantly downward fluxes. That shift can be traced to the treatment of temperature in the processing method of Goddijn-Murphy et al. (2015). Calculations at each grid cell require the reprocessed SOCAT values described above and other data. The figure shows the point flux calculations in each grid cell where there is sufficient coincident data available (i.e. atmospheric and aqueous CO_2 , air pressure, salinity, SST and wind speed data all available for the same grid cell). The details of the standard calculation method are

described in the Appendix of Shutler et al. (2016), including corrections where direct calculation is not possible. White areas near the coast result from missing values of one or more necessary variable and the integrated fluxes reported below include the first-order correction explained by Shutler et al. (2016).

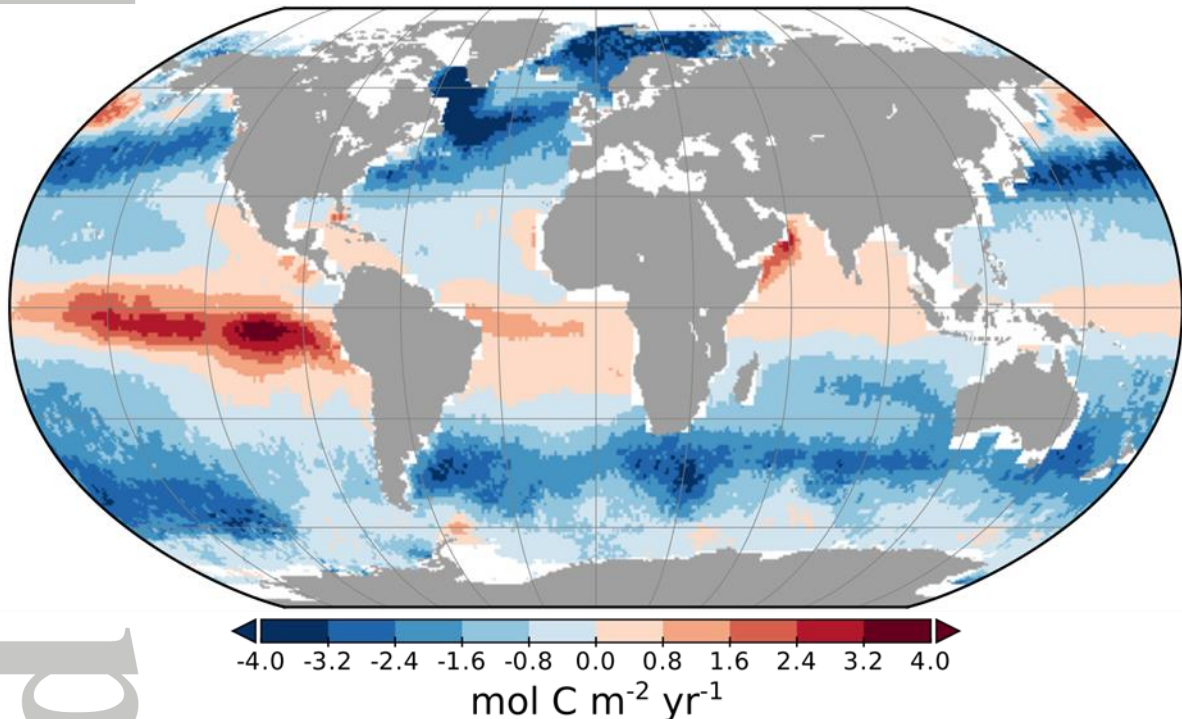


Figure 1. Illustrative map of annual CO₂ flux in 2010. Upward fluxes are defined as positive. Details are described in the main text.

Integrated fluxes are calculated as an integration of Equation 1 over the surface of a spheroidal Earth. Where a grid cell includes land or ice, the integrated flux is adjusted by the default scheme described in the Appendix of Shutler et al. (2016). Ice fractions were drawn from the OSI-SAF/SSMI sea ice data record. As noted already, analysis in this study is restricted to “global” air-sea fluxes and to four oceanic areas. The four regions are illustrated in the Supplementary Information (Figure S2). A relatively small region south of 60°S is defined as “Southern”, while the main oceanic expanse and marginal seas are assigned as “Pacific”, “Atlantic” or “Indian” according to International Hydrographic Office (IHO) definitions. The sum of these four regions provides a global value. This global value excludes inland seas such as the Mediterranean Sea, but includes the deep ocean and their marginal seas. Generally, through widespread international effort and collaboration the sampling of upper ocean CO₂ measurements has increased over the last few years, as reflected in Figure 2. However, the increase has not been steady and the geographical distribution is not even. The northern hemisphere is sampled better than the southern hemisphere. Political instability and piracy has deterred measurements in large parts of the Indian Ocean for several years, while some locations, e.g. within the seas of Southeast Asia, lack any data.

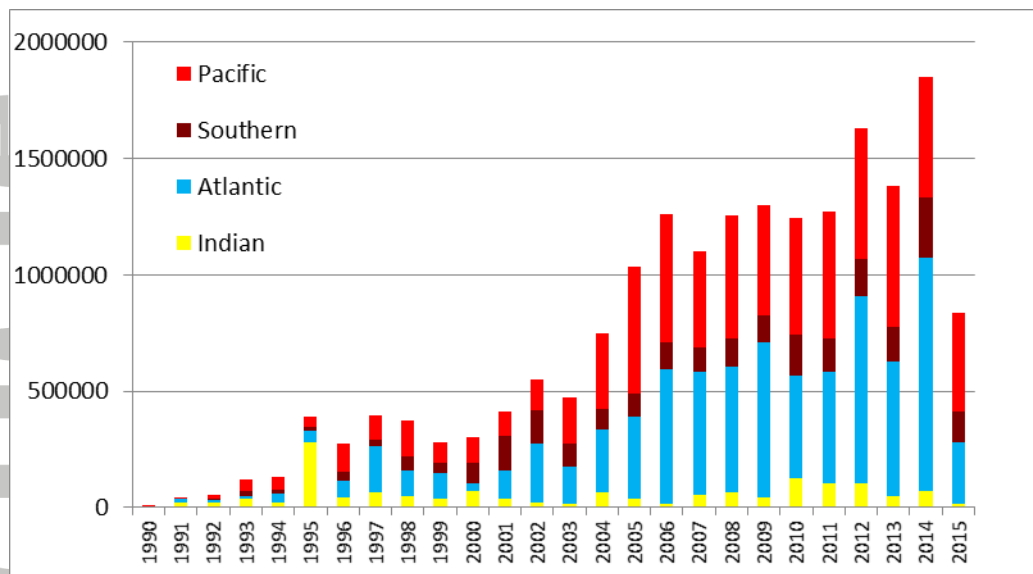


Figure 2. Distribution of the number of data in SOCAT Version 4 (and this study) by calendar year and IHO region.

2.3 Experiments

Experiments are designed to explore sensitivities of global and regional fluxes to several factors. The first two sets of experiments are anticipated in the preceding subsections, but in this subsection we will summarise all the experiments.

2.3.1 Temperature and salinity gradients

As described in Section 2.2, skin effects of -0.17 K and $+0.1$ PSU are assumed as standard, but one set of calculations was conducted with no skin effect (i.e. using sub-skin temperatures and salinities throughout the calculation).

2.3.2 Transfer velocities

As described in Section 2.1, several parameterizations of transfer velocity are available. We made calculations using the 8 polynomial relationships to wind speed described in Table 1. Standard temperature and salinity gradients were applied.

2.3.3 Sparse sampling

As described in Section 5.4 of Goddijn-Murphy et al. (2015), the influence of the fortuitous timing and track of specific measurement cruises can be evaluated by a bootstrap experiment. We constructed an ensemble of ten synthetic datasets by selecting randomly with replacement from actual cruises (using the cruise ID within SOCAT). Each synthetic dataset was processed in exactly the same way as the standard dataset. Atmospheric datasets, wind speeds, sea surface temperatures and salinities are not afflicted by sparse sampling to the same degree as surface ocean fugacities of CO_2 and we used a standard dataset for these other variables throughout all the calculations.

2.3.4 Year of sampling

Atmospheric partial pressures of CO_2 rose at slightly more than $1.5 \mu\text{atm yr}^{-1}$ on average through the period 1980 -2000, but accelerated more recently. A fairly steady increase of $\sim 2.0 \mu\text{atm yr}^{-1}$ in the period 2000-2010 was followed by a further acceleration in

recent years (Ed Dlugokencky and Pieter Tans, NOAA/ESRL; www.esrl.noaa.gov/gmd/ccgg/trends/). Over a similar time period, analyses of surface ocean have not shown clear evidence of an acceleration in the rise of fugacity (Zeng et al., 2014), which underpins our decision to apply a steady $1.5 \mu\text{atm yr}^{-1}$ increase throughout the central calculations. Part of the problem is the need to extrapolate old measurements forward to 2010. For example, if we extrapolate forward fugacities from 1990 to 2010, then an error of $0.5 \mu\text{atm yr}^{-1}$ in the trend leads to an error of $10 \mu\text{atm}$ in the 2010 value. If extrapolated globally, that amounts to an error of $> 2 \text{ PgC yr}^{-1}$ in the net flux, which is clearly unacceptable. Since only $\sim 24\%$ of the data precede 2005, the situation is less serious, but there are regions and months where older data and the assumed trend may substantially affect the calculated result. A few key issues are apparent from Figure 2; for example more than 50% of data in the Indian sector precedes 2004, including almost 20% in 1995 alone. There is an issue with applying a simple and uncertain trend and the procedure and experiments described in the following two paragraphs are designed respectively to minimise the effect and to assess the uncertainty.

A bias in the value of a steady trend should be less critical if the reference year is centred within the period of data, since the error for data extrapolated forward should be balanced by the error for data extrapolated backward (assuming approximately the same number of data before and after the reference year). A poor value of trend will result in local errors and increased uncertainty, but the central estimate of the net global flux should be affected only slightly, unless the trend accelerates or decelerates. Local and regional patterns of sampling (e.g. if a particular oceanic region is mainly sampled before the reference year) will introduce regional biases.

Our study included simulations where only SOCAT data from 2005-2014 (roughly centred on 2010) was used. We present data for only a single transfer velocity parameterisation (Nightingale et al., 2000), but calculations are presented for two applied trends, $1.5 \mu\text{atm yr}^{-1}$ and $2.0 \mu\text{atm yr}^{-1}$.

3 Results

3.1 Temperature and salinity gradients

Calculations were conducted with and without skin effects using the transfer velocity parameterisation of Nightingale et al. (2000); Model 7 in Table 1. This pair of calculations show that the combined effect of the thermal and haline skins is an increase in global net flux of $-0.35 \text{ PgC yr}^{-1}$. Note that this amount is proportional to the atmospheric concentration and to the transfer velocity, but is insensitive to upper ocean concentrations, and should be reliable given reasonable estimates of the global mean skin differences (-0.17 K and $+0.1$ salinity) and global mean transfer velocity.

3.2 Transfer velocities

Fluxes were calculated globally and regionally using each of the eight models described in Table 1. The results for the annual and global totals are summarised in Figure 3. (Some seasonal and regional results are included in subsections 3.3 and 3.4 below).

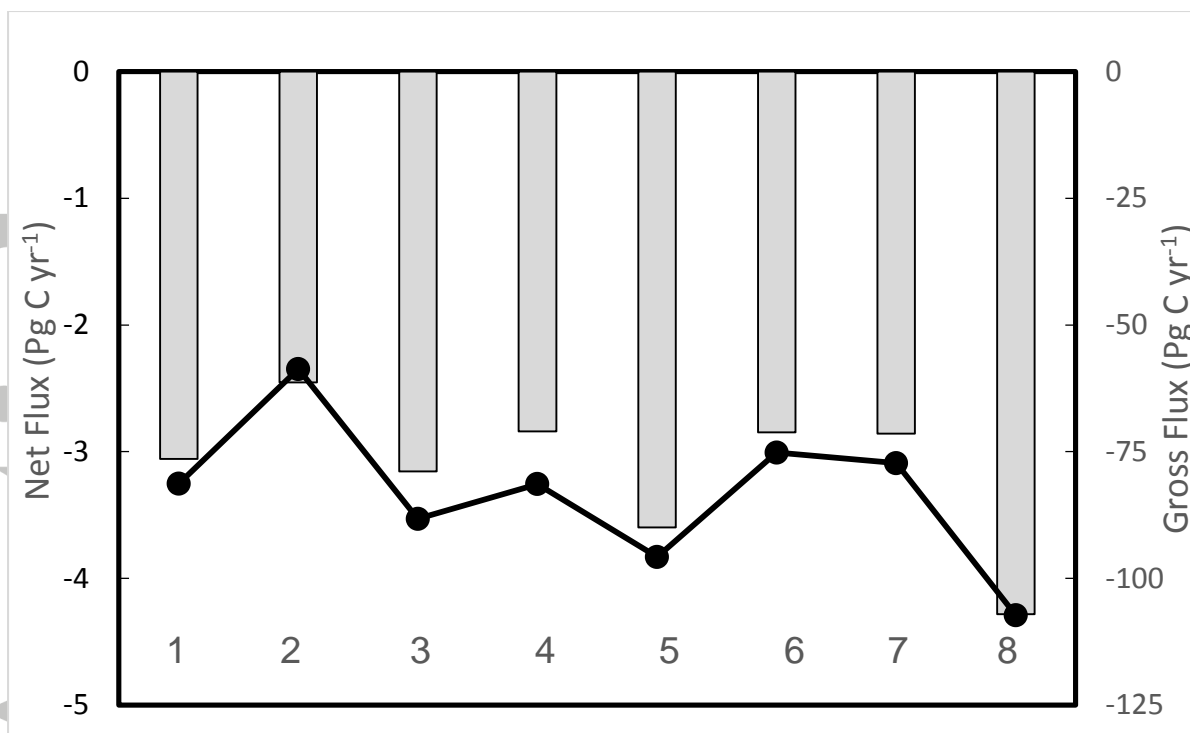


Figure 3. Net and gross downward fluxes of CO₂ calculated for the eight transfer velocity models described in Table 1. Net fluxes are the grey bars, gross fluxes are the points. Note the different scales on the two y-axes (coincidence implies a net flux that is 4% of the gross flux).

The presentation in Figure 3 illustrates the variation in both gross and net downward fluxes among the eight models described in Table 1. An increase in a transfer velocity will always increase the gross downward flux, but the situation is more complicated for net fluxes. There is a net upward flux in CO₂ in some locations and months. An increase in transfer velocity at those locations will decrease the total net downward flux. It is useful to compare the effect of a transfer velocity model on both net and gross fluxes, since both comparisons can provide insight. The net fluxes vary between ~3.6% and ~4.2% of the gross flux among the 8 models.

Among all eight models, the annual net global flux ranges between -2.4 and -4.3 PgC yr⁻¹, but most values cluster around a net flux of around -3 PgC yr⁻¹ on a gross downward flux of ~-80 PgC yr⁻¹. Analyses of subsets of this data are described in Section 4.2.

3.3 Sparse sampling

The influence of sparse sampling is illustrated by presenting values from the standard calculations along with those of the ten members of the ensemble from the bootstrap experiment. Results are calculated using Model 1 of the transfer velocity (Ho et al., 2006), but effects of sparse sampling will be similar for other models. The annual net global CO₂ flux varies with a standard deviation of 2% among the ensemble members about a mean of -2.99 PgC yr⁻¹, while no member varies by more than 4% from the standard value (-3.06 PgC yr⁻¹).

Annual net values are shown in Figure 4. The results suggest that the values for each oceanic region are reasonably robust. The greatest variation is seen among the values for the Indian region, but the Pacific and Atlantic also contribute significantly to the global uncertainty.

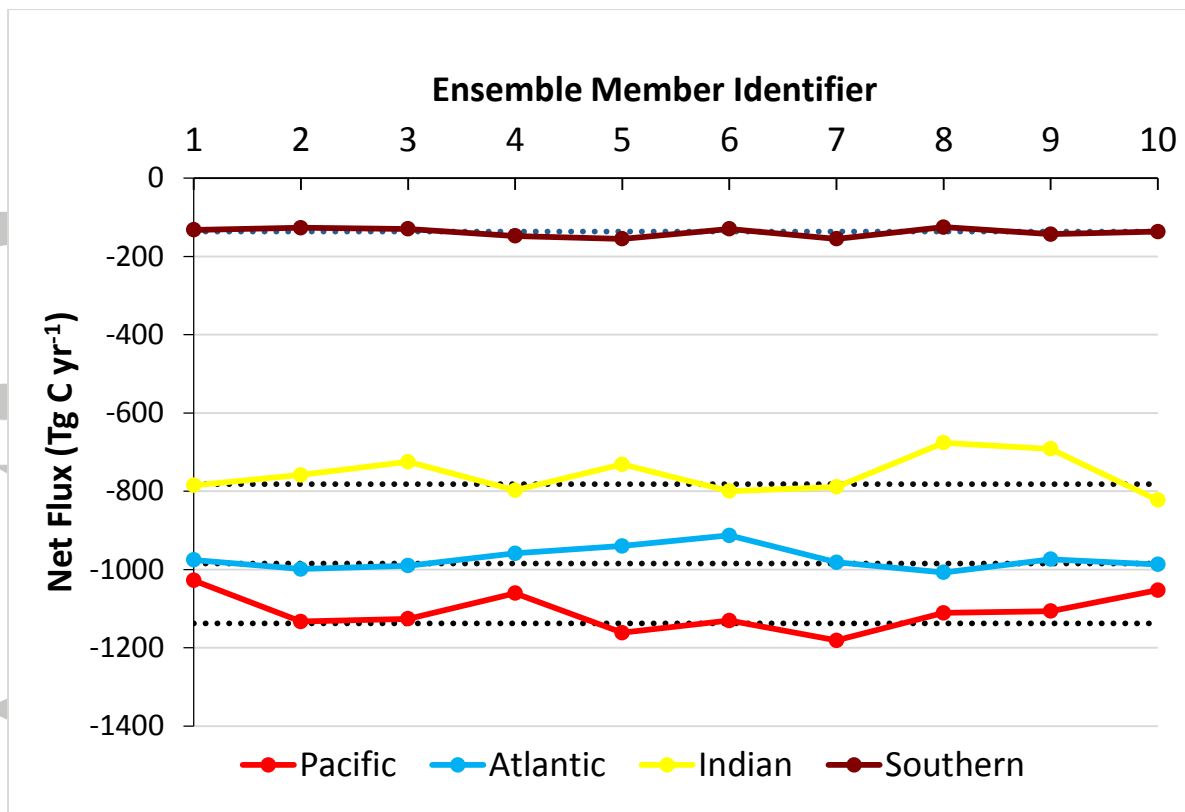


Figure 4. Variation of the net global CO₂ flux among the 10 members of the ensemble by region. For each region, the standard calculation is shown by a horizontal dotted line.

Sparse sampling is a more significant issue for seasonal variations. Standard and ensemble values are presented by calendar month and region in Figure 5. These variations amongst ensemble members are larger than those of annual values, however values for most regions and months seem to be fairly robust. There are several combinations of month and region that show high variation and these can be identified as significant gaps in sampling. The Southern Ocean seasonal cycle appears robust, with a substantial downward flux in the austral summer (peaking in February) and a near-balance in winter and spring. However, examination of the spatial distribution of the data indicates that the great majority of the data in winter comes from a restricted region around Drake Passage, so the bootstrap technique may not adequately reflect the true uncertainty of the fluxes in the Southern Ocean as a whole (see below, section 4.2.2.2). In the Pacific Ocean, variation is substantial in all months, but peaks in September to November. Among the Atlantic sector results, June and August appear to be the most problematic. The net downward flux in the Indian sector is small from December to March and appears to be well constrained by sampling in those months, while the net downward flux is larger in the rest of the year, but less adequately constrained by data.

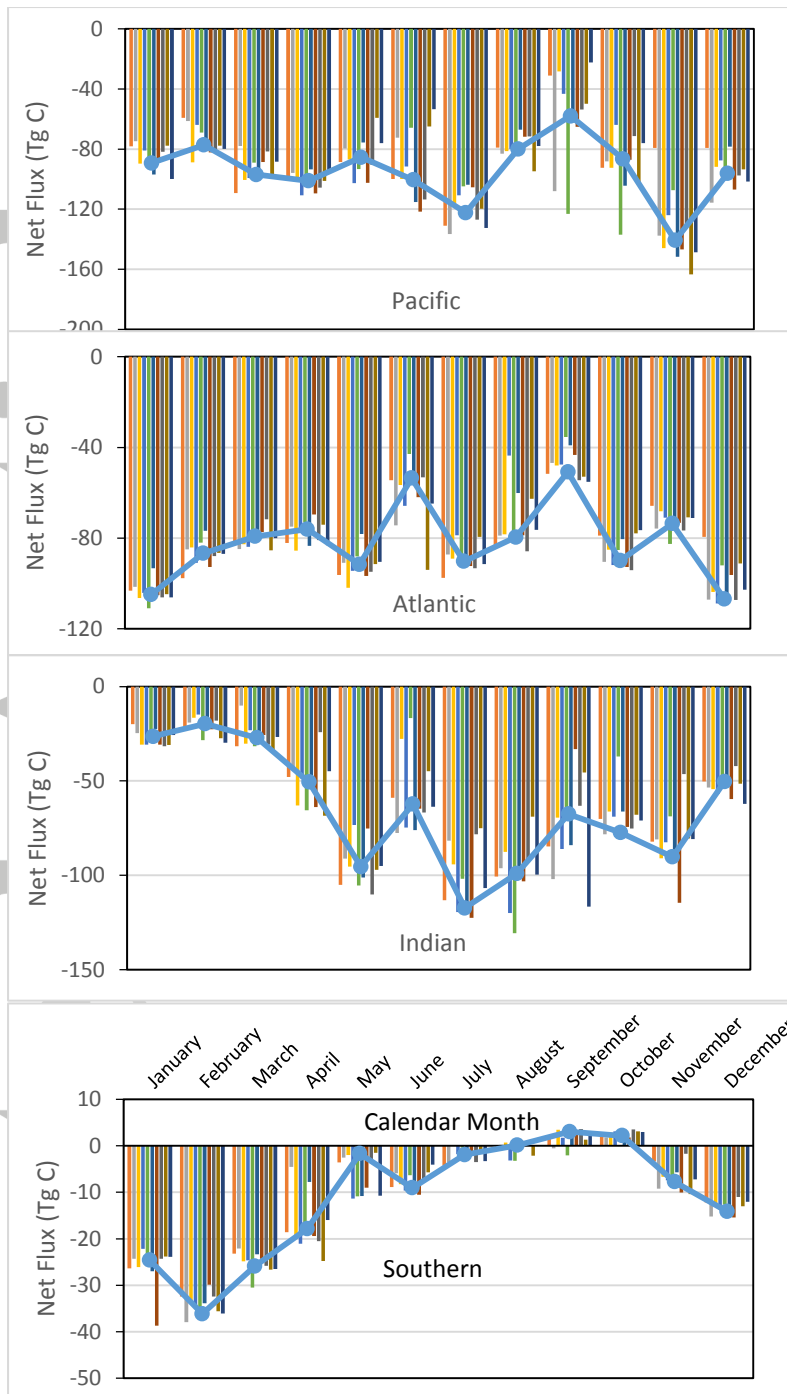


Figure 5. Net fluxes of CO₂ in each oceanic region by calendar month. Calendar months are ordered left to right. The standard calculation is shown by a line and each bootstrap experiment by colored bars, ordered left to right within each calendar month.

3.4 Year of sampling

In addition to the standard results based on upper ocean data since 1990 (as illustrated in Figure 2), we have two sets of calculations restricting data to the period 2005 to 2014. As expected, since the data for the latter two sets are distributed around the reference year of 2010, the sensitivity to the assumed trend was almost removed. Interestingly, for both trends,

the annual net downward CO₂ flux was decreased by more than 0.2 PgC yr⁻¹ (-2.86 PgC yr⁻¹ for all data and 1.5 μatm yr⁻¹ trend; -2.65 PgC yr⁻¹ for data since 2005 and 1.5 μatm yr⁻¹ trend; -2.63 PgC yr⁻¹ for data since 2005 and 2.0 μatm yr⁻¹ trend). However, on closer examination, we determined that the results when restricting the period of data were less credible than the standard results, as we will explain below.

Examining the results of calculations by region and calendar month is revealing (Figure 6). It is clear that in most regions and months there is very little difference in results among the three sets, but the standard values diverge from the other two sets in several key cases. In the global values of net flux, the restriction by year greatly reduces the net downward flux in June, November and December, but this is compensated partly by higher values in August. Breaking down the analysis by region reveals more: The result in June is related to a dataset contributing to values in the Atlantic and Indian sectors. The difference in August relates to Indian Ocean data. The differences in November and December are related to early datasets for the Pacific Ocean. Crucially, in every case that we have examined, the large difference is a result of excluding data in regions and seasons where no data since 2005 are available. That exclusion places an even greater burden on interpolation and the results of geospatial interpolation are manifestly poor when early data are excluded. The global and regional seasonality is affected by excluding data prior to 2005 and in our judgment, the result appears far less credible. That finding is unfortunate, since it is clear (from the almost complete coincidence of these two sets of results in Figure 6) that excluding earlier data and centring the data can greatly reduce sensitivity to secular trends. Nonetheless, by using all data from SOCAT version 4, and centring on 2010, we judge that the confidence intervals can be narrow; specifically, the global interval could exclude the “2005 to 2014 only” value. A substantial challenge remains to sample CO₂ in the upper ocean sufficiently, but we are encouraged that only a limited number of region and month combinations exhibit major issues. In most cases, the seasonal dependence can be inferred with reasonable confidence by informing assessment of poorly sampled months with neighbouring better-sampled months; but the last few calendar months in the Pacific are a major concern. Sampling efforts could be prioritized further using maps of uncertainty. Such maps are not included here, but are a direct product of the interpolation method (Goddijn-Murphy et al., 2015).

Using multi-year data remains a necessary compromise for global calculations and calculations of many regions, including using data from the twentieth century for evaluations of the recent oceanic sink.

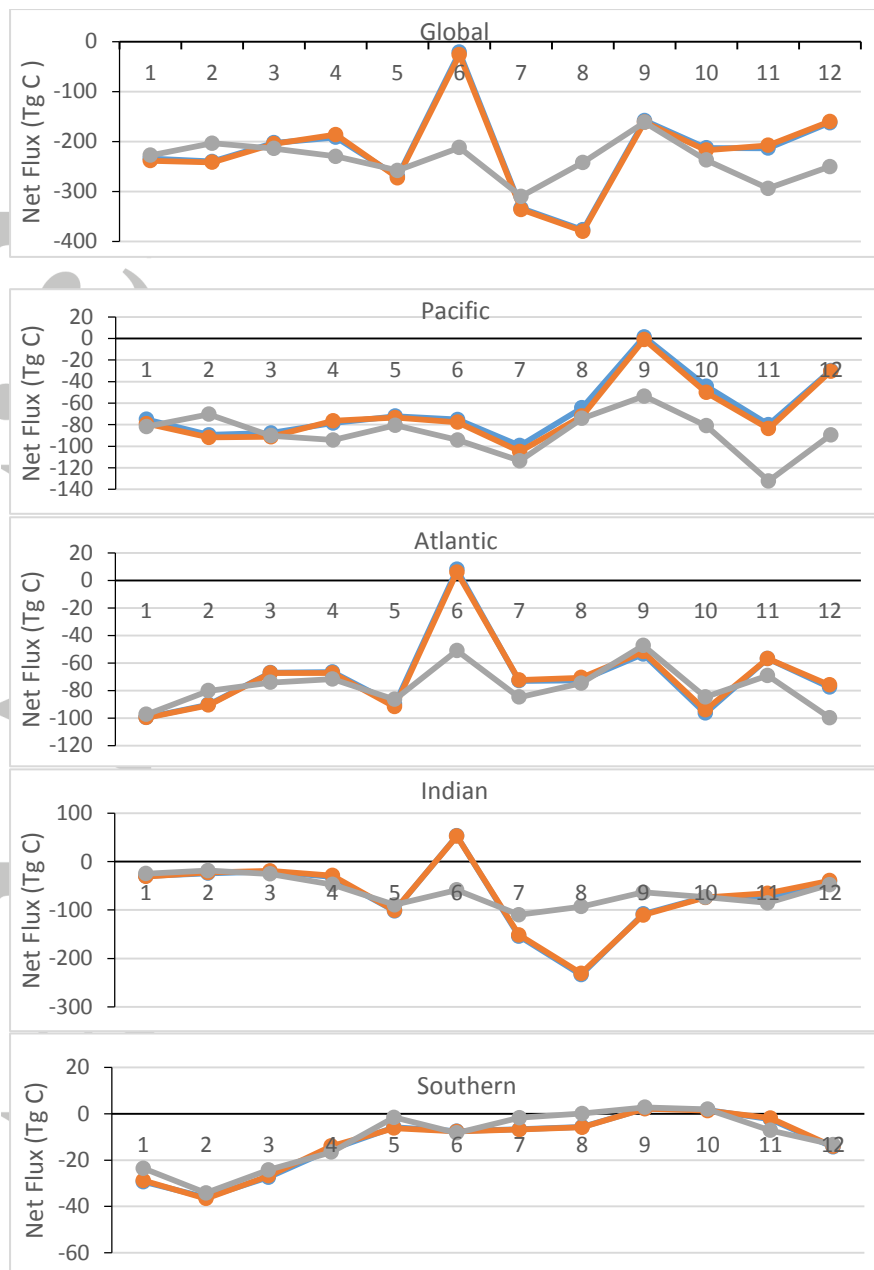


Figure 6. Net downward fluxes of CO₂ in each oceanic region by calendar month. Calculations from the full SOCAT version 4 dataset are compared to calculations using only data from 2005 to 2014. Complete set and 1.5 $\mu\text{atm yr}^{-1}$ trend in gray, limited and 1.5 $\mu\text{atm yr}^{-1}$ trend in orange, limited and 2.0 $\mu\text{atm yr}^{-1}$ trend in blue (almost entirely obscured by orange).

4 Analysis and Discussion

4.1 Expression of uncertainty

In this study, uncertainties are expressed following the guidelines set out by the Joint Committee for Guides in Metrology (BIPM, 2008) as far as is practical. Those guidelines seek a “realistic” rather than a “safe” estimate of uncertainty (see Annex E in BIPM, 2008). That objective leads to combining uncertainties irrespective of whether individual

components are random or systematic in nature. The logic and mathematics of the method are summarised in this sub-section and then their application to the global net flux of CO₂ is described in the following sub-section.

The simplest statistical analyses of uncertainty are based on “*n*” repeated, independent measurements of a quantity (a “Type A” evaluation, in the terminology of BIPM (2008)). Since this method is standard, we summarise it very briefly here, but provide a more thorough treatment in the Supplementary Information for reference.

A simple analysis of a set of repeated measurements yields a central estimate of the true value, \bar{q} , and its standard uncertainty, *u*. The value of interest might be inferred from the measurement of one or more related quantities (“inputs”). If not all substantial uncertainties are included in a single measurement process then a combined uncertainty, *u_c*, may need to be calculated from uncertainties arising from individual inputs. In the case of independent errors, the combined uncertainty can be calculated as the square root of the sum of uncertainties squared. Confidence intervals are placed on the quantity, $Q = q \pm e u_c$, where *e* = 2 or *e* = 3 are commonly used, the choice depending on how conservative an estimate is required.

It is not always practical to evaluate an uncertainty by a standard Type A method. A key philosophy of the BIPM methodology is that all substantial uncertainties should be considered irrespective of whether a Type A estimate is practical. Type B methods are diverse and include any method of evaluation that is not Type A (BIPM, 2008). Type B evaluation requires more expert judgement, but can be as reliable as Type A given adequate information. A combined uncertainty is calculated ignoring the method of evaluation.

This study requires Type B evaluations to be included. For key uncertainties, we will use the same equations for sample mean and standard error as a Type A evaluation, but typically using a only small number of estimates. In this case, the experimental average, standard deviation and standard uncertainty will be poor estimates of the true values. Perhaps more importantly, since the effective number of degrees of freedom will be small, the level of confidence associated with a defined extended uncertainty will be weaker than for Type A uncertainties and poorly defined. Therefore, while an interval can be reported, its definition is inexact. We will calculate an interval using *e* = 2, thus quoting a value in the form:

$$Q = \bar{q} \pm 2 u_c \quad (3)$$

We can reasonably state that the value is “likely” to fall within the defined interval, but we cannot be more specific.

4.2 Global net flux

4.2.1 Central value

The main purpose of this study was to investigate key uncertainties, but first we comment on the central estimate of the global net air-sea flux of CO₂, about -3.0 Pg C yr⁻¹ in 2010. This value is much larger than estimated by Takahashi et al. (2009), but that should be considered in the context both of other studies suggesting a rapid growth of this flux after 2000 and a contribution of -0.35 Pg C yr⁻¹ from skin effects that were omitted by Takahashi et al. A more useful comparison is with an estimate of -2.3 ± 0.6 PgC yr⁻¹ in 2011 by Landschützer et al. (2016), who omitted skin effects, used a simple quadratic relationship to wind speed and scaled their results to a fairly low estimate of the ocean uptake in 2000 (-1.7 ± 0.4 PgC yr⁻¹ in 2000; Gruber et al., 2009).

Combining our estimate of $-3.0 \text{ Pg C yr}^{-1}$ with a value of the annual pre-industrial flux of $+0.4 \pm 0.2 \text{ PgC yr}^{-1}$ (Takahashi et al., 2009), implies an annual anthropogenic sink of -3.4 Pg C , which is high compared to the accepted range of values for the period 2000-2010 of between -1.0 and $-3.2 \text{ Pg C yr}^{-1}$ (Rhein et al., 2013). That seems reasonable, since the sink probably had an upward trend in 2000-2010 (a trend of $\sim -0.5 \text{ PgC}$ per decade is implied by Le Quéré et al. (2014; see especially Figure 7). For the carbon budget to be balanced, a larger oceanic sink requires a smaller land sink, which requires further investigation.

4.2.2 Uncertainties

Quantitative estimates of uncertainty will be limited to the annual global net flux of CO_2 (referenced to 2010) and will be calculated by the methods outlined in the previous subsection. We begin with considering uncertainty related to the transfer velocity and then consider upper-ocean sampling and thermal and haline effects. Our conclusions are contingent on having included all major sources of uncertainty in the global value.

4.2.2.1 Transfer velocity

Models of transfer velocity were chosen to elucidate various sources of uncertainty in the net fluxes (Sections 2.1, 2.3.2 and 3.2). Models 1 to 7 are based on estimates of gas transfer velocity by a single experimental method, i.e. DTEs, while Model 8 is a highly cited model based on micrometeorological fluxes of CO_2 (McGillis et al., 2001). By choosing different subsets of these models, we will reach four estimates of uncertainty. Since all fluxes are simply proportional to the transfer velocity, a particular percentage uncertainty in transfer velocity translates to an identical percentage uncertainty in the flux.

Two estimates (*A* and *B*) of uncertainty will be based on Models 1 to 4, which are different statistical fits to a single data set. Model 1 (Ho et al., 2006) defines a simple square-law dependence on wind speed with reported confidence limits of the single fitting coefficient, “ c_2 ”, of 7%. Models 2, 3 and 4 are taken from a response by Ho et al. (2007) to a comment on the previous paper. Ho et al. argue that while Models 2, 3 and 4 also fit the data, they can be excluded. However, in our opinion there is no theoretical reason to favour a simple square law and only Model 2 is poorer empirically. The uncertainty identified here is that associated with the statistical model and scatter within the data set; both a structural uncertainty (the choice of model) and a parameter uncertainty associated with the inability of any simple model to describe all the variance in the data. The average flux and the structural uncertainty can be estimated by applying Eqn.s S7-S10 to a subset of the models, but note that the number of models is small and they are not independent, therefore these are not Type A uncertainties. The parameter uncertainty can be estimated approximately from the reported confidence limits of $\pm 7\%$ in Model 1. For the purpose of completing calculations similar to Eqn.s S13 and 3, those limits are interpreted as a standard uncertainty of 3.5% and assumed to be independent of the structural uncertainty. Two estimates of structural uncertainty alternately exclude (*Estimate A*) or include (*Estimate B*) the least credible model, Model 2. The results are included in Table 2. Note that in each case, the combined standard uncertainty is calculated from the parameter uncertainty and the structural uncertainty using Eqn. S13. The application of that equation is questionable, since the uncertainties are not independent, and has a substantial effect as the structural and parameter uncertainty are comparable. For *Estimate A*, final limits of $\pm 9\%$ are estimated, while $\pm 13\%$ is calculated for *Estimate B*.

Table 2. Estimates of combined standard uncertainty in transfer velocity (Eqn. 14) and the inferred limits of confidence (Eqn. 17) in the net global air-sea CO₂ flux

Estimate	k_w models	$u_{\text{structural}}$ (%)	u_c (%)	flux (Pg C yr ⁻¹) ± Δflux(%)
A	1, 3 and 4	3.1	4.7	-3.02 ± 9%
B	1, 2, 3 and 4	5.4	6.4	-2.88 ± 13%
C	1, 5, 6 and 7	N/A	5.7	-3.09 ± 11%
D	1, 5, 6, 7 and 8	N/A	8.3	-3.32 ± 17%

A different path to estimating uncertainties related to transfer velocities is to take a number of “expert opinions” and analyse the transfer velocities and fluxes resulting from those. This method is similar in principle to sending samples to a number of laboratories and analysing the variation among the laboratory reports. (An opinion may be accompanied by expressions of uncertainty, but if we include those uncertainties in a calculation of combined uncertainty then that is likely to be “double-counting” (see Section 4.3.10 of BIPM, 2008) since a large part of both the “difference in opinion” and “limits of confidence” will share common sources). The method is hazardous, firstly since there are subjective judgments about what opinions should be included and secondly since the more credible opinions are not really independent. In the case of transfer velocities, most of the models considered are based on a single type of experiment - DTEs (see explanation at end of Section 2.1) – but Model 8 has been deliberately chosen to include a contrary opinion. The remaining models, Models 1 to 7, can be used to investigate uncertainty internal to the DTE method, but they cannot shed light on biases common to all applications of the DTE method (e.g. Jacobs et al., 2002; Goddijn-Murphy et al., 2016). Thus, any estimate based on any subset of Models 1 to 7 is likely to underestimate the true uncertainty. On the other hand, by including Model 8, we are introducing a more controversial opinion and might exaggerate the true uncertainty. Generally including methods other than DTEs will increase the calculated uncertainty, since the results of other methods often differ substantially from those of DTEs. Specifically, Model 8, based on direct flux measurements of CO₂ (McGillis et al., 2001), implies relatively high transfer velocities in strong winds. More recent direct flux measurements of CO₂ report a diversity of results, ranging from further evidence of very high transfer velocities in strong winds (e.g., Blomquist et al., 2017) to values fairly consistent with the DTE-based models (e.g., Bell et al., 2017).

Models 1 to 4 have already been analysed (*Estimates A and B* above and Table 2) and share the same data and authorship. Model 5 also shares the same transfer velocity data with Models 1 to 4, but Smith et al. (2011) combined these with different wind speed data and Model 5 predicts the highest fluxes among Models 1 to 7. Models 6 and 7 share experiments with the other models, but provide a different - though not independent – interpretation. Therefore, we use Models 1, 5, 6 and 7 to analyse variation in expert opinion, limited to interpretation of DTEs (*Estimate C*). We add Model 8 to form a larger subset and a broader variation in expert opinion (*Estimate D*). In each case, the variation among the model outputs is interpreted through Eqn.s 9-11 to a complete value of the standard uncertainty.

This study assesses transfer velocity models based on local process studies and their extrapolation globally. An inventory of Bomb ^{14}C in the ocean can provide a global constraint (Naegler et al., 2006; Sweeney et al., 2007); essentially a single-point calibration with estimated confidence limits of $\sim\pm 25\%$. Applying the constraint to a square law dependence on wind speed yields a coefficient $c_2 = 0.27$ (Takahashi et al., 2009) or 0.251 (Wanninkhof, 2014), values which are satisfyingly close to Model 1. Other polynomial dependencies on wind speed could meet the same single constraint and it is not possible to find a unique solution. Since the Bomb ^{14}C flux is largely downward, it is more directly a constraint on the gross downward flux (see Figure 3). An approximate understanding of the effect of this constraint can be derived through normalising the net fluxes, by shifting all gross downward fluxes to a single value. It is apparent that the Bomb ^{14}C constraint could recalibrate Models 2, 5 and 8 to give net fluxes in a similar range to the other models (Model 1 and Model 5 would then be identical). However, while the values of Models 1, 3, 4, 6 and 7 would be rearranged, their range would not be greatly affected. In conclusion, while global inventories provide a useful check on models of transfer based on tracer studies, they are insufficiently accurate to narrow confidence intervals.

Table 2 summarises the four *Estimates*. The mean of the four flux values together imply a net air-sea flux of greater than $-3.0 \text{ Pg C yr}^{-1}$. On the other hand, Models 4, 6 and 7 are among the most credible models and all yield a net air-sea flux close to $-2.85 \text{ Pg C yr}^{-1}$. On balance, we think that $-3.0 \text{ Pg C yr}^{-1}$ represents a reasonable “consensus” value of the net flux. Among the estimates of uncertainty, two suggest a standard uncertainty of approximately 5% (limits of confidence, $\pm 10\%$), while the other two suggest a relatively high standard uncertainty of up to 10% (limits of confidence, $\pm 20\%$). A rational case can be made for either conclusion and we will consider both possibilities in the following. We continue now to consider other sources of error in the estimated flux.

4.2.2.2 Ocean sampling

The results from the experiments on sparse sampling (Section 3.3) and year of sampling (Section 3.4) are revealing. It is important here to distinguish between “sparse sampling” (requiring interpolation between distant cruise tracks, but within coherent water bodies) and “absence of data” (requiring extrapolation into oceanographically distinct regions). The experiments reveal the uncertainty related to interpolation, but can only hint at issues related to unsampled oceanographic regions.

Applying Eqn.s S8-S10 to the members of the ensemble described in Section 3.3, we find an average of $-2.99 \text{ Pg C yr}^{-1}$ with a standard uncertainty of only 0.6%. The standard uncertainty is quite low compared to a $\sim 2\%$ difference between the standard value ($-3.06 \text{ Pg C yr}^{-1}$) and the average of the ensemble, but the standard value would not be an outlier among the ensemble.

Uncertainties are larger when we resolve particular oceanic regions (Figure 4) and specific calendar months (Figure 5), but most results remain fairly robust. The accumulation and quality-controlling of data in the SOCAT databases contributes greatly to reducing uncertainties to a useful level even in many remote regions. One noteworthy result is an apparently robust estimate of the net annual downward flux south of 60°S (the IHO definition of the Southern Ocean) of $\sim -0.13 \text{ Pg C yr}^{-1}$ (see Figures 4 and 5). That result contrasts with an upward flux calculated by Takahashi et al. (2009) south of 58°S , but is far more consistent with other analyses of the global carbon cycle (e.g. Gruber et al., 2009).

Excluding older data would be desirable, but we find that excluding data prior to 2005 makes the results less credible (Section 3d and Figure 6). By choosing a fairly central

reference year, 2010, we have avoided a strong sensitivity of the global flux to the assumed trend in upper ocean CO₂. However, we should consider the three values of net global flux from these experiments (see Section 3d) in reaching an estimate of the combined standard uncertainty resulting from upper ocean sampling (including the necessary choice of a trend). In conclusion, a standard uncertainty related to sampling of 3% seems reasonable for the global net flux.

The bootstrap method used in this study randomly resamples individual cruises irrespective of their age. Since the method does not resample periods, it cannot test for effects of decadal or multi-decadal variability. It is possible that our global and regional estimates for 2010 are distorted by samples from earlier decades, but that cannot be evaluated.

4.2.2.3 Thermal effects

The standard runs in this study assume a cool skin effect of -0.17 K and a salty skin effect of +0.1 PSU giving a contribution of -0.35 PgC yr⁻¹ to the global net flux (approximately -0.40 PgC yr⁻¹ for the thermal effect and + 0.05 PgC yr⁻¹ for the haline effect). While a correction is certainly necessary (Woolf et al., 2016), there is uncertainty in the magnitude of both skin anomalies and in the effect of a thermal skin effect on CO₂ flux, while the effect of warm layers is not explicitly included. We will make a “Type B” estimate of the standard uncertainty associated with thermal and haline effects separately.

Several factors need to be considered in estimating the flux uncertainty related to thermal effects. There is a distribution in the values of the measured thermal skin difference (Donlon et al., 2002) and both systematic and varying differences between SST products from different instrumentation (Merchant et al., 2014). Our value of -0.17 K is taken from the analysis of Donlon et al. (2002), who report an average cool skin effect of -0.17 K for a wind speed greater than 6 ms⁻¹ with a root-mean-square variation of 0.07K. Random variations are relatively unimportant to a calculation since the resulting standard uncertainty is very small given the large number of measurements. Differences of the average values among datasets (e.g. Table 1 of Donlon *et al.* (2002) shows averages ranging from -0.14 K to -0.20 K) are more significant, since they may imply either small methodological biases or regional and seasonal differences. Nevertheless, the challenge of measuring the thermal skin effect is no longer a major impediment. Two competing systematic errors may be more significant (Woolf et al., 2016). Firstly, the calculation of the effect on CO₂ flux depends on choosing a model of how the temperature profile within the thermal skin affects the CO₂ profile (Sections 4 and 5 of Woolf et al. (2016)). The calculations in this study assume the “rapid model”, while the alternative “equilibrium model” would predict a greater effect. Secondly, the presence of warm layers (Section 6 of Woolf et al., 2016) implies an opposing effect on CO₂ fluxes. Since there are two opposing effects of similar magnitude, we do not apply a systematic correction, but apply a combined standard uncertainty of 0.04 K to the thermal skin effect, which implies a standard uncertainty of 3% in the net global flux.

4.2.2.4 Haline effects

The size of the “salty skin anomaly” is highly uncertain (Yu, 2010; Zhang & Zhang, 2012) and the action of the anomaly on CO₂ fluxes is complex (Section 7 of Woolf et al., 2016). We suggest a standard uncertainty of 0.1 PSU on the best estimate of 0.1 PSU is reasonable. Even with that large relative uncertainty, the effect on the net global flux is relatively small, 1.7%.

4.2.2.5 Combined uncertainty

We can now combine our four sources of uncertainty (transfer velocity, sampling, thermal and haline) into a combined standard uncertainty of the flux. As explained earlier, we retain two representative values of the standard uncertainty related to transfer velocity (5% and 10%). Both of these values are large enough to dominate over the other three values and we calculate combined standard uncertainties of $(5^2 + 3^2 + 3^2 + 1.7^2)^{1/2} = 7\%$ and $(10^2 + 3^2 + 3^2 + 1.7^2)^{1/2} = 11\%$. Using Eqn. 3, we can then calculate two alternative limits of confidence in the net global flux, $-3.0 \pm 0.4 \text{ Pg C yr}^{-1}$ or $-3.0 \pm 0.7 \text{ Pg C yr}^{-1}$, which probably span the reasonable range.

4.2.2.6 Other sources of uncertainty

This study is based on key experiments rather than an exhaustive model of all uncertainties. We have argued *a priori* that uncertainty in our understanding of the transfer velocity and sparse sampling of upper ocean CO₂ are critical, with temperature and salinity sampling also important. Some uncertainties are not considered explicitly, for example those relating to different datasets on wind speed, temperature and salinity, since we have adopted a preferred single dataset for each of these variables. We assume that basic understanding rather than faults in the wind speed data are primarily responsible for uncertainties in transfer velocity. We acknowledge that measurement uncertainty in the wind speed is sufficient to imply substantial uncertainties in air-sea heat flux calculations (Kent et al., 2013), and can affect gas flux calculations (Roobaert et al., 2018), but stand by the statement that model uncertainty is more important to gas transfer velocities. Other uncertainties of interest – but neglected in our calculations – include those related to the covariance of variables (these are ignored when we use averages for each cell) and the interaction of wind, waves and currents (Villas Bôas et al., 2019).

4.2.3 Limitations of the standard model

The models of gas transfer used in this study are simple in principle and easy to apply to climatological calculations and within Earth-system models. However, these models are probably too simple, since they ignore mechanistic processes (Goddijn-Murphy et al., 2016). Simplifying air-sea gas transfer to a simple case of wind-driven stirring excludes processes including convection, modification by surface-active materials (Pereira et al., 2018) and the full complexity of wind-wave-current interaction, wave breaking and bubbles. The effect of rain (Ashton et al., 2016) is also neglected. All of these complications could introduce regional and seasonal biases that probably have little effect on the global gross flux, but may introduce substantial biases in all seasonal and regional fluxes and in the calculated global net flux.

Modelling of carbon isotopes by Krakauer et al. (2006) suggest a regional dependence of transfer velocities contradictory to the simple models tested here. Altering the latitudinal distribution of transfer velocities to lower transfer velocities at high latitudes, while raising them at low latitudes may be compatible with tracer data and would lower the net downward CO₂ flux. On the other hand, the asymmetry of bubble-mediated transfer may also introduce an additional downward flux of $\sim 0.1 \text{ PgC yr}^{-1}$ (Zhang, 2012) or more (Leighton et al., 2018). In summary, the central estimate of $-3.0 \text{ Pg C yr}^{-1}$ may be reasonable, but we should be cautious in our estimates of confidence, veering towards the more conservative of the two limits of confidence already discussed. Limits of confidence of $\pm 0.6 \text{ Pg C yr}^{-1}$ are proposed.

5 Conclusions

The uncertainties in air-sea CO₂ flux estimates based on applying the air-sea flux equation are reducing as the most pernicious sources of error are eroded. Fairly consistent models of air-sea gas transfer velocity by different authors and a far greater accumulation and consistent quality controlling of upper ocean data, all contribute to greater confidence in estimates. However, the process is ongoing and both gaps in ocean data and weaknesses in gas transfer models require further attention. As knowledge of transfer velocity and sampling of the upper oceans improves, some other sources of uncertainty will gain priority.

A contemporary global annual net air-sea CO₂ flux of -3.0 ± 0.6 Pg C yr⁻¹ (referenced to 2010) is likely, though a value outside of those bounds cannot be ruled out. The primary cause of uncertainty in this global annual value is uncertainty in transfer velocities, but if the standard uncertainty of transfer velocity is reduced to <5%, then uncertainties in thermal effects and sampling become important. For the “referenced, annual and global” value of net air-sea CO₂ flux, the sampling of upper ocean CO₂ appears to be adequate, but sparsity of sampling is a much greater issue for regional and seasonal fluxes and is insufficient for annually-resolved fluxes. Confidence limits will necessarily be broader for reference years distant from 2010, since uncertainty in the trend of upper-ocean CO₂ is a major issue, especially if the reference year is not centred within the available data.

Acknowledgments, Samples, and Data

None of the authors has a conflict of interest with the conclusions of this study. We are grateful for the contributions of Steve Jones and Sylvain Herlédan to the study. The study was funded by the European Space Agency Support to Science Element (STSE) OceanFlux Greenhouse Gases and Oceanflux Greenhouse Gases Evolution projects (contract numbers 4000104762/11/I-AM and 4000112091/14/I-LG). MJY was co-funded by the Natural Environment Research Council [ORCHESTRA, grant number NE/N018095/1]. AJW was funded by a Royal Society Research Professorship. The Surface Ocean CO₂ Atlas (SOCAT) is an international effort, endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS) and the Integrated Marine Biosphere Research (IMBeR) program, to deliver a uniformly quality-controlled surface ocean CO₂ database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT. The data used in this study are described by Holding et al. (2018) which also provides a link to the processing software. Further data are described in the Supporting Information.

References

- Ashton, I. G., Shutler, J. D., Land, P. E., Woolf, D. K., & Quartly, G. D. (2016). A sensitivity analysis of the impact of rain on regional and global sea-air fluxes of CO₂, *PLoS ONE*, 11, e0161105. doi:10.1371/journal.pone.0161105.
- Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., et al. (2016). A multi-decade record of high quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT), *Earth System Science Data*, 8, 383-413. doi:10.5194/essd-8-383-2016.
- Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., et al. (2014). An update to the Surface Ocean CO₂ Atlas (SOCAT version 2), *Earth System Science Data*, 6, 69-90. doi:10.5194/essd-6-69-2014.
- Bell, T. G., Landwehr, S., Miller, S. D., de Bruyn, W. J., Callaghan, A. H., Scanlon, B., et al. (2017). Estimation of bubble-mediated air-sea gas exchange from concurrent DMS and CO₂

transfer velocities at intermediate–high wind speeds, *Atmospheric Chemistry and Physics*, 17, 9019-9033. doi.org/10.5194/acp-17-9019-2017.

BIPM (2008). Evaluation of measurement data — Guide to the expression of uncertainty in measurement. JCGM, 100.

http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

Blomquist, B. W., Brumer, S. E., Fairall, C. W., Huebert, B. J., Zappa, C. J., Brooks, I. M., et al. (2017) Wind speed and sea state dependencies of air-sea gas transfer: Results from the High Wind speed Gas exchange Study (HiWinGS). *Journal of Geophysical Research: Oceans*, 122 (10), 8034-8062. <https://doi.org/10.1002/2017JC013181>.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., et al. (2013). Carbon and Other Biogeochemical Cycles. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 465-570). New York, NY: Cambridge University Press.

Donlon, C. J., Minnett, P. J., Gentemann, C. Nightingale, T. J., Barton, I. J., Ward, B., & Murray, M. J. (2002). Toward improved validation of satellite sea surface skin temperature measurements for climate research, *Journal of Climate*, 15 (4), 353-369.

Fangohr, S., Woolf, D. K., Jeffery, C. D., & Robinson I. S. (2008). Calculating long-term global air-sea flux of carbon dioxide using scatterometer, passive microwave, and model reanalysis wind data, *Journal of Geophysical Research: Oceans*, 113, C09032. doi:10.1029/2005JC003376.

Garbe, C.S., Rutgersson, A., Boutin, J., Delille, B., Fairall, C. W., Gruber, N., et al. (2014). Transfer across the air-sea interface. Chapter 2 in P.S. Liss & M.T. Johnson (Eds.), *Ocean-Atmosphere Interactions of Gases and Particles* (pp. 55-144). New York, NY: Springer Earth System Sciences. doi:10.1007/978-3-642-25643-1_2.

Goddijn-Murphy, L., Woolf, D. K., Callaghan, A. H., Nightingale, P. D. & Shutler, J. D. (2016). A reconciliation of empirical and mechanistic models of the air-sea gas transfer velocity, *Journal of Geophysical Research: Oceans*, 121, 818–835, doi:10.1002/2015JC011096.

Goddijn-Murphy, L. M., Woolf, D. K., Land, P. E., Shutler, J. D., & Donlon, C. (2015). The OceanFlux Greenhouse Gases methodology for deriving a sea surface climatology of CO₂ fugacity in support of air–sea gas flux studies. *Ocean Science*, 11(4), 519-541.

Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., et al. (2009). Oceanic sources, sinks, and transport of atmospheric CO₂, *Global Biogeochemical Cycles*, 23, GB1005. doi:10.1029/2008GB003349

Ho, D. T., Law, C. S., Smith, M. J., Schlosser, P., Harvey, M., & Hill P. (2006), Measurements of air-sea gas exchange at high speed winds in the Southern Ocean: Implications for global parameterizations, *Geophysical Research Letters*, 33, 1-6. <https://doi.org/10.1029/2006GL026817>

Ho, D. T., Law, C. S., Smith, M. J., Schlosser, P., Harvey, M., & Hill P. (2007), Reply to comment by X. Zhang on “Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations”, *Geophysical Research Letters*, 34. doi:10.1029/2007GL030943

Ho, D. T., Wanninkhof, R., Schlosser, P., Ullman, D. S., Hebert, D. & Sullivan, K. F. (2011), Toward a universal relationship between wind speed and gas exchange: Gas transfer

- velocities measured with $^3\text{He}/\text{SF}_6$ during the Southern Ocean Gas Exchange Experiment. *Journal of Geophysical Research: Oceans*, 116, C00F04. doi:10.1029/2010JC006854
- Holding, T., Ashton, I., Woolf, D K. & Shutler., J. D. (2018) FluxEngine v2.0 and v3.0 reference and verification data. *Pangaea*, doi:10.1594/PANGAEA.890118
- Jacobs, C., Kjeld, J. F., Nightingale, P., Upstill-Goddard, R., Larsen, S. & Oost, W. (2002). Possible errors in CO_2 air-sea transfer velocity from deliberate tracer releases and eddy covariance measurements due to near-surface concentration gradients. *Journal of Geophysical Research: Oceans*, 107, C9, 3128. doi:10.1029/2001JC000983
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W. Deavan, D. Gandin, L., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, 77, 437-470. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)
- Kent, E. C., Fangohr, S. and Berry, D. I. (2013). A comparative assessment of monthly mean wind speed products over the global ocean. *International Journal of Climatology*, 33(11), 2520-2541. doi:10.1002/joc.3606.
- Krakauer, N. Y., Randerson, J. T., Primeau, F. W., Gruber, N. & Menemenlis, D. (2006). Carbon isotope evidence for the latitudinal distribution and wind speed dependence of the air-sea gas transfer velocity. *Tellus, Series B*, 58, 390-417.
- Landschützer, P., Gruber, N. & Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, 30, 1396-1417. <https://doi.org/10.1002/2015GB005359>.
- Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., et al. (2016) Global carbon budget 2016. *Earth Systems Science Data*, 8, 605–649. <https://doi.org/10.5194/essd-8-605-2016>
- Leighton, T. G., Coles, D. G. H., White, P.R, Srokosz, M. & Woolf, D. K. (2018). Asymmetric transfer of CO_2 across a broken sea surface. *Scientific Reports*, 8301. <https://doi.org/10.1038/s41598-018-25818-6>
- McGillis, W. R., Edson, J. B., Ware, J. D., Dacey, J. W. H., Hare, J. E., Fairall, C. W. & Wanninkhof, R. (2001). Carbon dioxide flux techniques performed during GasEx 98. *Marine Chemistry*, 75, 267-280.
- Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI), *Geoscience Data Journal*, 1, 179-191. doi: 10.1002/gdj3.20
- Naegler, T., Ciais, P., Rodgers, K., & Levin, I. (2006). Excess radiocarbon constraints on air-sea gas exchange and the uptake of CO_2 by the oceans. *Geophysical Research Letters*, 33(11),
- Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., & Liss, P. S. (2000). In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, *Global Biogeochemical Cycles*, 14, 373-387.
- Pereira, R., Ashton, I., Sabbaghzadeh, B., Shutler, J. D. & Upstill-Goddard, R. C. (2018). Reduced air-sea CO_2 exchange in the Atlantic Ocean due to biological surfactants. *Nature Geoscience*. doi: 10.1038/s41561-018-0136-2.

- Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., et al. (2013) A uniform, quality controlled Surface Ocean CO₂ Atlas (SOCAT), *Earth System Science Data*, 5, 125-143. doi:10.5194/essd-5-125-2013.
- Rhein, M., Rintoul, S.R., Aoki, S., Campos, E., Chambers, D., Feely, R.A. et al. (2013). Observations: Ocean. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 255-315). New York, NY: Cambridge University Press.
- Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., et al. (2015). Data-based estimates of the ocean carbon sink variability - First results of the Surface Ocean pCO₂ Mapping intercomparison (SOCOM). *Biogeosciences*, 12(23), 7251-7278.
- Roobaert, A., Laruelle, G. G., Landschützer, P., & Regnier, P. (2018). Uncertainty in the global oceanic CO₂ uptake induced by wind forcing: quantification and spatial analysis. *Biogeosciences*, 15, 1701-1720. <https://doi.org/10.5194/bg-15-1701-2018>.
- Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., et al. (2013). Surface Ocean CO₂ Atlas (SOCAT) gridded data products, *Earth System Science Data*, 5, 145-153, <https://doi.org/10.5194/essd-5-145-2013>
- Shutler, J. D., Land, P., Piolle, J-F., Woolf, D. K., Goddijn-Murphy, L. M., Paul, F., et al. (2016). FluxEngine: A flexible processing system for calculating atmosphere-ocean carbon dioxide gas fluxes and climatologies. *Journal of Atmospheric and Oceanic Technology*, 33(4), 741-756.
- Smith, M. J., Ho, D. T. Law, C. S. McGregor, J., Popinet, S., & Schlosser, P. (2011). Uncertainties in gas exchange parameterization during the SAGE dual-tracer experiment, *Deep-Sea Research II*, 58.
- Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L. & Wanninkhof, R.. (2007). Constraining global air-sea gas exchange for CO₂ with recent bomb ¹⁴C measurements. *Global Biogeochemical Cycles*, 21, GB2015. doi:10.1029/2006GB002784
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO₂, and net sea-air CO₂ flux over the global oceans. *Deep-Sea Research II*, 56, 554-577. doi:10.1016/j.dsr2.2008.12.009
- Villas Bôas, A. B., Arduin, F., Ayet, A., Bourassa, M. A., Brandt, P., Chapron, B. et al. (2019). Integrated observations of global surface winds, currents, and waves: requirements and challenges for the next decade. *Frontiers in Marine Science*, 6, 425. doi:10.3389/fmars.2019.00425.
- Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography Methods*, 12, 351-362. <https://doi.org/10.4319/lom.2014.12.351>
- Wanninkhof, R., Asher, W. E., Ho, D. T., Sweeney, C. & McGillis, W. R. (2009). Advances in quantifying air-sea gas exchange and environmental forcing, *Annual Review of Marine Science*, 1, 213-244. <https://doi.org/10.1146/annurev.marine.010908.163742>
- Weiss, R. F. (1974) Carbon dioxide in water and seawater: the solubility of a non-ideal gas, *Marine Chemistry*, 2(3), 203-215. doi: 10.1016/0304-4203(74)90015-2

Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L. M., & Donlon, C. J. (2016). On the calculation of air-sea fluxes of CO₂ in the presence of temperature and salinity gradients. *Journal of Geophysical Research: Oceans*, 121, 1229–1248. doi:10.1002/2015JC011427

Wróbel, I. & Piskozub, J. (2016). Effect of gas-transfer velocity parameterization choice on air-sea CO₂ fluxes in the North Atlantic Ocean and the European Arctic. *Ocean Science*, 12, 1091-1103. <https://doi.org/10.5194/os-12-1091-2016>.

Yu, L. (2010). On sea surface salinity skin effect induced by evaporation and implications for remote sensing of ocean salinity, *Journal of Physical Oceanography*, 40, 85–102. doi:10.1175/2009JPO4168.1

Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., & Nakaoka, S. (2014). A global surface ocean fCO₂ climatology based on a feed-forward neural network. *Journal of Atmospheric and Oceanic Technology*, 31(8), 1838-1849. <https://doi.org/10.1175/JTECH-D-13-00137.1>

Zhang, X. (2012). Contribution to the global air-sea CO₂ exchange budget from asymmetric bubble-mediated gas transfer. *Tellus, Series B*, 64. doi:10.3402/tellusb.v64i0.17260

Zhang, Y., & Zhang, X. (2012), Ocean haline skin layer and turbulent surface convections, *Journal of Geophysical Research: Oceans*, 117, C04017. doi:10.1029/2011JC007464

Accepted Article