AGU PUBLICATIONS

Global Biogeochemical Cycles

Supporting Information for

Key Uncertainties in the Recent Air-Sea Flux of CO_2

D. K. Woolf[1], J. D. Shutler[2], L. Goddijn-Murphy[3], A. J. Watson[2], B. Chapron[4], P. D. Nightingale[5], C. J. Donlon[6], J. Piskozub[7], M. J. Yelland[8], I. Ashton[9], T.Holding[2], U. Schuster[2], F. Girard-Ardhuin[4], A. Grouazel[4], J-F. Piolle[4], M. Warren[5], I. Wrobel-Niedzwiecka[7], P. E. Land[5], R. Torres[5], J. Prytherch[8*], B. Moat[8], J. Hanafin[4†], F. Ardhuin[4], and F. Paul[4]

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

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

[3]University of the Highlands and Islands, UK.

[4]Institut Francais de Recherché pour l'Exploitation de la Mer, France.

[5]Plymouth Marine Laboratory, UK.

[6]European Space Agency, The Netherlands.

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

[8]National Oceanography Centre, Southampton, UK.

[9]College 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.

Contents of this file

Introduction

Most of the data supporting this study are available from a source cited in the main paper (Holding et al., 2018): Additional information is provided here on the models of transfer velocities, the SOCAT v4 data and definitions of oceanic regions used in this study. The computation of transfer velocities is described. The standard calculation of confidence intervals is described.

Text S1. Computation of transfer velocities

In this study, parameterizations are limited to a simple polynomial form (as described in the main paper):

$$k_w = (Sc/660)^{-1/2} (c_0 + c_1U + c_2U^2 + c_3U^3) \quad \text{(S1)}$$

where U is the instantaneous ten-metre-elevation wind speed and Sc is the Schmidt number of the dissolved gas. The coefficients, [c_n], for an established model can be taken from the literature, or a new set can be proposed and tried. Measurements of wind speed, temperature and salinity can be used to calculate a set of values, [b_n], defined by

$$b_i = (Sc/660)^{-1/2}\, U^i \quad \text{(S2)}$$

for each grid cell and it follows that

$$k_w = c_0b_0 + c_1b_1 + c_2b_2 + c_3b_3 \quad \text{(S3)}$$

Note that this approach reduces the computation load and time. The full set of [b_n] needs to be computed only once for each grid cell and – having been saved as a look-up table – can be reused for calculations with different coefficients, [c_n].

A calculation of b_i on a "grid cell" implies an average over time and an area of sea surface. This procedure introduces some questions about appropriate averaging scales. These questions have been investigated for CO_2 fluxes by Fangohr et al. (2008) among others, and it is known that temporal variability is a much greater issue than spatial variability. An instantaneous wind speed (more practically, a wind speed averaged over several minutes, or over a satellite footprint) is specified in Eqn. 2, since it is expected that transfer velocities will respond quite rapidly to changes in wind forcing. Thus, in Eqn. 3 it is necessary to use the "average of the *i*th power of U" rather than the "*i*th power of the average of U". The moments of wind speed are readily available up to third order (e.g. the average of the cube of wind speed in a given month, or the related skewness of the wind speed distribution) and thus grid calculations of b_i are practical for I = 0, 1, 2, 3. Temporal and spatial co-variances (for example, between U and Sc) are neglected in our study and fluxes are calculated as a simple product of averages.

Further, if we consider calculating an air-sea flux integrated regionally or globally, then we can again take advantage of the summative nature of the transfer velocities for the simple polynomial expression. If we define a contribution to the flux (gross or net, depending on the multiplier *X*) as

$$F_i = c_i \Sigma\, b_i X \quad \text{(S4)}$$

where $\Sigma\, b_i X$ provides the required numerical integration, we can then define a normalized flux contribution G_i, such that

$$G_i = F_i/c_i = \Sigma\, b_i X \quad \text{(S5)}$$

in which case, the corresponding flux for any polynomial can be calculated as

$$F = c_0 G_0 + c_1 G_1 + c_2 G_2 + c_3 G_3 \quad \text{(S6)}$$

This architecture is designed to enable an infinite set of flux calculations by a small number of intensive, basic computations. Here, we will present results for a few [c_n] and a few *X*, but it is trivial to calculate [G_n] from a few simultaneous equations and then calculate the implied integrated flux for any other polynomial expression of k_w in wind speed.

Text S2. Calculation of confidence intervals

Type A estimates of a measurand are calculated by classical methods as follows: The expected value of the quantity, q, will be the simple average of the measured values, q_k:

$$\overline{q} = \frac{1}{n} \sum_{k=1}^{n} q_k \quad \text{(S7)}$$

It is assumed that every reasonable effort has been made to eliminate systematic errors in the quantity by this stage and the variation of q_k about the true value is random. The experimental variance of the observations is given by:

$$s^2(q_k) = \frac{1}{n-1} \sum_{j=1}^{n} (q_j - \overline{q})^2 \quad \text{(S8)}$$

The experimental variance of the mean is then given by

$$s^2(\overline{q}) = \frac{s^2(q_k)}{n} \quad \text{(S9)}$$

A Type A "standard uncertainty" of the quantity is estimated as the experimental standard deviation of the mean:

$$u(q) = s(\overline{q}) \quad \text{(S10)}$$

The values, $\overline{q}$ and u provide a central estimate and an expression of uncertainty ("standard uncertainty) respectively of the quantity of interest. The value of *n* must be sufficient to achieve a satisfactory estimate of q and u.

In the case that several inputs contribute to the final uncertainty in the quantity of interest, it is necessary to calculate a combined uncertainty of the quantity. Often the quantity of interest, q will be calculated as a function of many input values

$$q = f(x_1, x_2, \ldots\ldots, x_N) \tag{S11}$$

If each input value is independent of all the others then a combined standard uncertainty, u_c, can be calculated from:

$$u_c^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2 (x_i) \tag{S12}$$

If instead of measuring the uncertainty in the input, we measure the uncertainty in the output, *q*, associated with each input, u_i, then a simpler form of the same equation can be written:

$$u_c^2 = \sum_{i=1}^{N} u_i^2 (q) \tag{S13}$$

The combined standard uncertainty, u_c, is the preferred method of expressing the uncertainty of a measured quantity. However, the average and combined standard uncertainties do not alone express an interval within which the true value should fall. In order to define that interval, one must also introduce the expanded uncertainty, *U*, such that if $\overline{q}$ is the measured sample mean then the true value, *Q*, is likely to fall in the interval:

$$\overline{q} - U \leq Q \leq \overline{q} + U \tag{S14}$$

The expanded uncertainty, *U*, is defined very simply by introducing a coverage factor, *e*, such that:

$$U = e\, u_c \tag{S15}$$

Clear definitions of *p* and *e* follow from the Central Limit Theorem, which implies measured $\overline{q}$ should be drawn from a normal distribution of calculable standard deviation (Annex G of BIPM, 2008). A "confidence interval" can be calculated by recognising that the probability distribution is defined by a t-distribution characterised by the measured average and standard deviation and by the effective number of degrees of freedom. For a fairly large number of degrees of freedom, an interval defined by *e = 2* should encompass more than 95% of the probability distribution, while *e= 3* yields a level of confidence of approximately 99%.

The treatment of Type A uncertainties described above is extended to combined uncertainties calculated partly or entirely from Type B uncertainties as described in Section 4.1 of the main text.

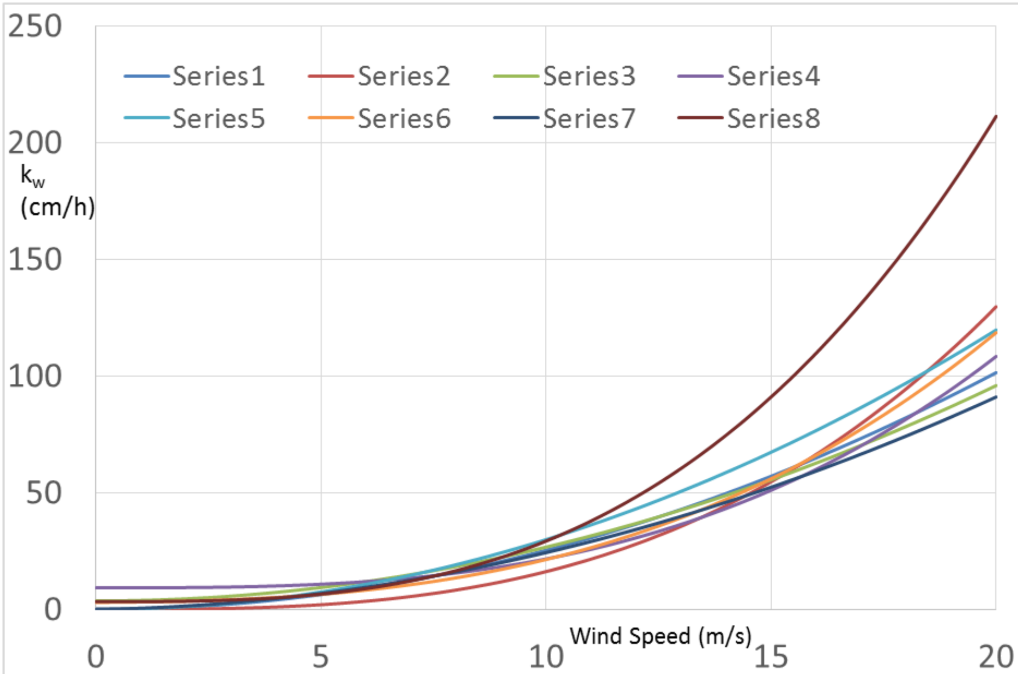

Figure S1. Transfer velocity plotted against wind speed. Transfer velocities are calculated for a a Schmidt number of 660. Series numbers correspond to the Model numbers described in the main paper and in Table S1.

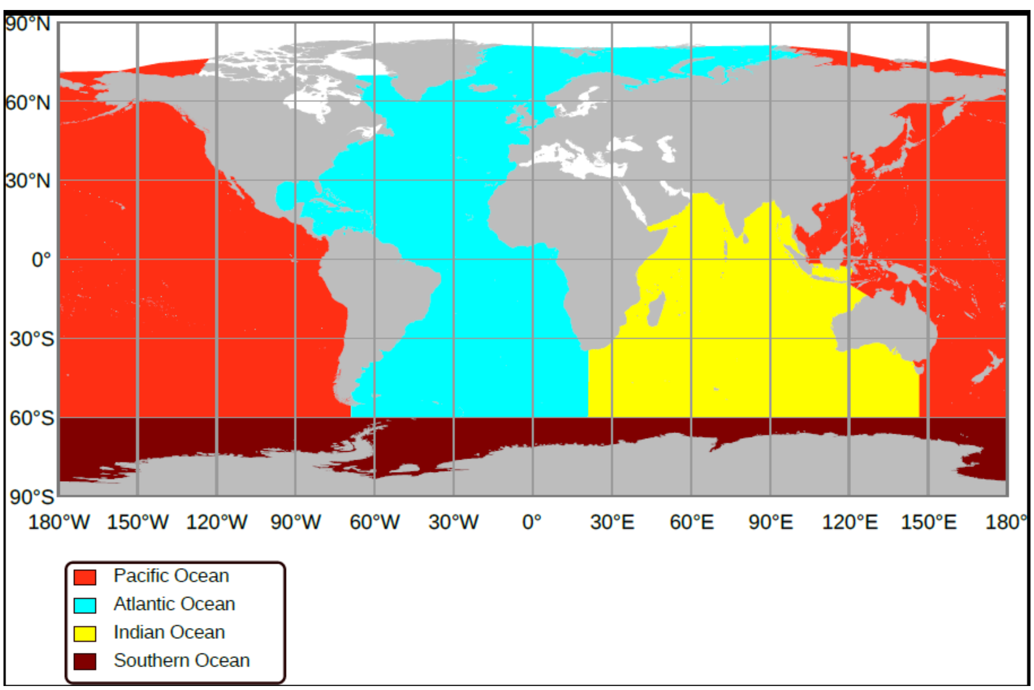

Figure S2. The definition of ocean regions as defined by IHO and applied in this study.

Schmidt number	600	6002	6003	6004	660	6605	6606	6607
Reference	**a0**	**a1**	**a2**	**a3**	**a0**	**a1**	**a2**	**a3**
Ho 2006 "quad"	**0**	0	0.266	0	0	0	0.253621	0
Ho 2007 "cubic"	**0**	0	0	0.017	0	0	0	0.016209
Ho 2007 "quad + c"	**3.8**	0	0.242	0	3.623158	0	0.230738	0
Ho 2007 "cubic + c"	**9.7**	0	0	0.013	9.248587	0	0	0.012395
Smith et al. 2011 "AFC quad"	**0**	0	0.314	0	0	0	0.299387	0
Wanninkhof et al. 2009	**3.146427**	0.104881	0.067124	0.011537	3	0.1	0.064	0.011
Nightingale et al. 2000	**0**	0.333	0.222	0	0	0.317503	0.211669	0
McG 2001	**3.461069**	0	0	0.027269	3.3	0	0	0.026

Table S1. Coefficients of the polynomial models calculated for Schmidt numbers of 600 and 660. Each of the original papers describe the coefficients for one or other standard Schmidt number and have been calculated in this study for the other number and rounded to 4 decimal places.

Year	Indian	Atlantic	Pacific	Southern	ocean total
1990	271	306	6951	3	7531
1991	18692	17220	7237	2137	45286
1992	18598	15553	16113	2471	52735
1993	37706	13520	47360	18925	117511
1994	18620	38243	53979	18714	129556
1995	278484	53112	45964	11813	389373
1996	43635	72118	123942	35592	275287
1997	65978	196183	105107	28596	395864
1998	47854	109999	155731	59881	373465
1999	37210	110129	91521	42262	281122
2000	71074	31018	111574	87290	300956
2001	36357	124460	108296	145316	414429
2002	19013	255923	131551	142107	548594
2003	16985	158455	198450	101350	475240
2004	66134	269037	322310	89793	747274
2005	36175	353759	545951	97926	1033811
2006	14408	577725	552271	118437	1262841
2007	51371	529593	414670	105503	1101137
2008	65898	538555	530675	120691	1255819

2009	43867	666427	475598	115826	1301718
2010	123558	444239	503222	172928	1243947
2011	102239	479449	546895	144690	1273273
2012	103926	802946	558538	162380	1627790
2013	47441	581533	601605	149645	1380224
2014	70236	1003668	521779	257375	1853058
2015	14360	265711	426300	129527	835898
sum 1990-2015	1450090	7708881	7203590	2361178	18723739
sum 1990-2004	776611	1465276	1526086	786250	4554223
% 1990-2004	53.55606	19.00764	21.18508	33.2990567	24.32325616

Table S2. Breakdown of SOCAT v4 data frequency by year and oceanic region. A tabulation of the data shown graphically in Figure 1 of the main paper.