Spatio-temporal Variability in Surface Ocean pCO_2 Inferred from Observations

A thesis submitted to the School of Environmental Sciences of the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

By Steve D Jones

November 2012

Abstract

The variability of surface ocean pCO_2 is examined on multiple spatial and temporal scales. Temporal autocorrelation analysis is used to examine pCO_2 variability over multiple years. Spatial autocorrelation analysis describes pCO_2 variability over multiple spatial scales. Spatial autocorrelation lengths range between $<$50 km in coastal regions and other areas of physical turbulence up to 3,000 km along major currents. Analysis of the drivers of pCO_2 shows that ocean currents are the primary driver of spatial variability. Autocorrelation lengths of air-sea CO_2 fluxes are approximately half as long as for pCO_2 due to the effects of highly variable wind speeds.

The influence of modes of climate variability on ocean pCO_2 and related air-sea CO_2 fluxes is examined through correlations of climate indices with interannual pCO_2 anomalies separated from the long-term trend and mean seasonal cycle. Changes in the El Niño Southern Oscillation alter pCO_2 levels by -6.6 $\pm$ 1.0 μatm per index unit (μatm iu^{-1}) in the Equatorial Pacific, leading to changes in air-sea flux of up to 0.40 $\pm$ 0.06 Pg C yr^{-1}. The Pacific Decadal Oscillation shows statistically significant correlations with pCO_2 across the Equatorial Pacific, North Pacific and North Atlantic. No statistically significant correlations are found with the North Atlantic Oscillation in the North Atlantic.

An important product of the analysis performed in this thesis is a spatially and temporally complete interpolated data set of surface ocean pCO_2 data over an extended period. This data product is the first of its kind, both in terms of its coverage and the fact that it does not rely on the derivation of empirical relationships between pCO_2 and other biogeochemical variables. The technique works as well as or better than previous regional interpolations, with 90% of values likely to be within 30 μatm of the actual pCO_2 value.

Acknowledgements

I am thoroughly indebted to my academic supervisors, Corinne Le Quéré, Christian Rödenbeck and Andrew Manning. They have provided endless good advice, ideas and encouragement whenever I have needed it throughout the past three and a half years. They allowed me the freedom to explore those topics that interested the most, which made for a thoroughly enjoyable PhD experience. The occasional push in the right direction when I drifted off course is equally appreciated.

No work in the academic world is performed in isolation. My work has benefitted from contact with other scientists, both in the UEA and at other institutions. Particular thanks go to Ute Schuster for many productive discussions of my ideas and results. Others too many to name have also contributed to my PhD in equally numerous ways, from scientists ready to answer any and all questions (despite the fact that we may have met only briefly at a conference) to office mates simply discussing things over a coffee.

This PhD would not have been possible without the efforts of the scientists who have willingly contributed their data and time to the LDEO and SOCAT databases of combined global ocean pCO_2 measurements, frequently without the benefit of funding to support the huge amount of work involved. Dorothee Bakker and Benjamin Pfiel in particular have given constant support and advice for my use of the SOCAT database. Similarly, the technical staff running the HPC Cluster have bent over backwards to accommodate my often less than reasonable computing needs - they have provided the best IT support I've ever encountered.

As everyone who has completed a PhD knows, the support of friends and family is an absolute necessity. They may not understand the sometimes esoteric work we do, but they will patiently listen while we try to explain it or whine about why it's all going wrong. Mum, Dad, Matt, and Helen in particular, but also everyone I know - thank you.

Contents

List of tables

List of figures

Chapter 1

Context and Background

1.1 The greenhouse effect

Without an atmosphere, the Earth's mean surface temperature would be approximately -18°C, over 30°C cooler than the observed temperature of approximately 15°C [*Seinfeld*, 2011]. The atmosphere acts as a filter on the incoming and outgoing radiation at Earth's surface, trapping some of the radiation and thereby maintaining its higher surface temperature (Figure 1.1, page 4). Certain gases present in the Earth's atmosphere do not readily absorb the high-frequency incoming radiation, but do absorb the low-frequency outgoing radiation [*Fourier*, 1827; *Tyndall*, 1861; *Kiehl and Trenberth*, 1997]. These are known as greenhouse gases. The extra energy thus captured leads to an energy imbalance at the top of the atmosphere. The imbalance is countered by an increase in the temperature of the atmosphere [*Arrhenius*, 1896; *Manabe and Wetherland*, 1967; *Kiehl and Trenberth*, 1997], which in turn increases the temperature at the Earth's surface.

The warming effect of the atmosphere on the Earth's surface is known as the greenhouse effect. The majority of the greenhouse effect is controlled by just a few greenhouse gases. Water vapour (H_2O) is the largest contributor to the greenhouse effect, accounting for 60% of the warming effect in clear skies. Carbon dioxide (CO_2) is the second biggest contributor (26%) [*Kiehl and Trenberth*, 1997], followed by methane (CH_4) with 9%, and nitrous oxyde (N_2O) with 3% [*Ramaswamy et al.*, 2001]. All other greenhouse gases combined contribute only 2% of the greenhouse effect.

Although H_2O is the most influential greenhouse gas [*Kiehl and Trenberth*, 1997], direct changes to its concentration are very small and have an insignificant effect on surface temperatures. The largest changes in H_2O concentration result indirectly from changing concentrations of other greenhouse gases such as hydrogen, methane, ozone and chlorine [*Le Texier et al.*, 1988; *Röckmann et al.*, 2004]. As they cause surface temperatures to increase, more water evaporates from the Earth's surface, causing a positive feedback. The effects of this feedback are relatively small compared to the temperature increase incurred by the initial release of greenhouse gases [*Held and Soden*, 2000; *Philipona et al.*, 2005; *Schneider et al.*, 2010]. In contrast, the lifetime of CO_2 in the atmosphere following a perturbation has multiple time scales: about 50% will be removed within 30 years, a further 30% within a century, and the final 20% may remain for thousands of years [*Archer*,

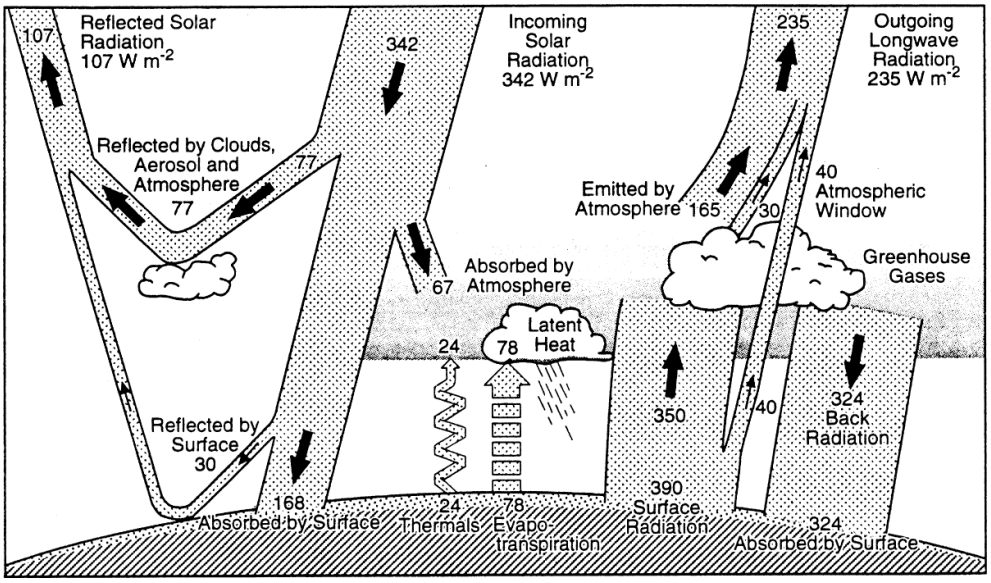

Figure 1.1: The Earth's annual global mean energy budget, in Watts m^{-2}. From *Kiehl and Trenberth* [1997].

2005; *Denman et al.*, 2007]. CO_2 absorbs and re-emits radiation at portions of the radiative spectrum not covered by H_2O [*Ahrens*, 2007] and thereby has a significant additional impact over H_2O. Thus CO_2 is the most influential greenhouse gas that is sensitive to anthropogenic activity and its increase has caused concerns over humanity's effect on the atmosphere and the resulting changes in the global climate. The Greenhouse Gas Index [*Hofmann et al.*, 2006] indicates that CO_2 has accounted for almost 80% of the warming observed in the period 1990-2010.

1.1.1 Anthropogenic changes to the greenhouse effect

Before the start of the industrial revolution in the 18th Century, variations in the Earth's mean surface temperature were due primarily to changes in natural forcings [*Crowley*, 2000], mainly changes in solar irradiance [*Lean et al.*, 1995] and volcanic activity [*Briffa et al.*, 1998]. During the 20th Century, Earth's temperature increased by 0.74°C, with the majority of that warming (0.55°C) occurring after 1970 [*Trenberth et al.*, 2007]. The HadCRUT3 temperature record [*Brohan et al.*, 2006] shows that warming has continued into the 21st Century, with the 2001-2010 period approximately 0.2°C warmer than the preceding decade. Variations in the above-mentioned natural climate forcings cannot explain the warming observed during the 20th Century; there is evidence that on the

contrary, natural forcings of the Earth's climate would have produced a cooling effect on climate in the absence of anthropogenic greenhouse gases [*Hegerl et al.*, 2007].

The increasing levels of greenhouse gases in the atmosphere observed throughout the 20th Century from sources such as fossil fuel burning, land use, land use change and deforestation [e.g. *Keeling et al.*, 2001; *Khalil et al.*, 2002; *Karakurt et al.*, 2012] and their corresponding influence on the greenhouse effect are the key to understanding the source of the recent warming. Atmospheric CO_2 concentrations varied between 260 and 280 ppm for the 10,000 years preceding the industrial revolution [*Indermühle et al.*, 1999]. Since that time approximately 365 petagrams of carbon (Pg C) have been released into the atmosphere from fossil fuel combustion [*Peters et al.*, 2012] and 150 Pg C from land use change [*Houghton*, 2007]. The global mean atmospheric concentration computed from measurements [*Masarie and Tans*, 1995] has increased to in excess of 390 ppm, a rate of change unprecedented in the last 60,000 years [*Indermühle et al.*, 2000].

The likely impact of the changes in greenhouse gas levels on the earth's surface temperature (termed the climate sensitivity) is typically expressed in terms of the projected equilibrium increase in the global average surface temperature in response to a doubling of atmospheric CO_2 concentration. The estimated temperature increase for a doubling of CO_2 is between 2°C and 4.5°C [*Meehl et al.*, 2007]. There is a large uncertainty about the Earth's climate sensitivity due to the complexity of the climate system and the many positive and negative feedbacks that occur in response to temperature changes [*Roe and Baker*, 2007]. The observed warming realised up to now is fully consistent with both the theoretical understanding of greenhouses gases and with the observed increase of atmospheric CO_2 and other greenhouse gases [*Hegerl et al.*, 2007].

1.2 The carbon budget

Not all of the emitted CO_2 remains in the atmosphere. A proportion of the emitted CO_2 is absorbed by the natural carbon reservoirs via the oceanic and terrestrial carbon 'sinks' [Figure 1.2, page 6; *Le Quéré et al.*, 2009]. The fraction of emitted CO_2 that remains in the atmosphere is termed the airborne fraction. Quantifying the airborne fraction and the magnitude of the sinks is essential for predicting how anthropogenic emissions will affect the Earth's future climate [*Houghton*, 2007].

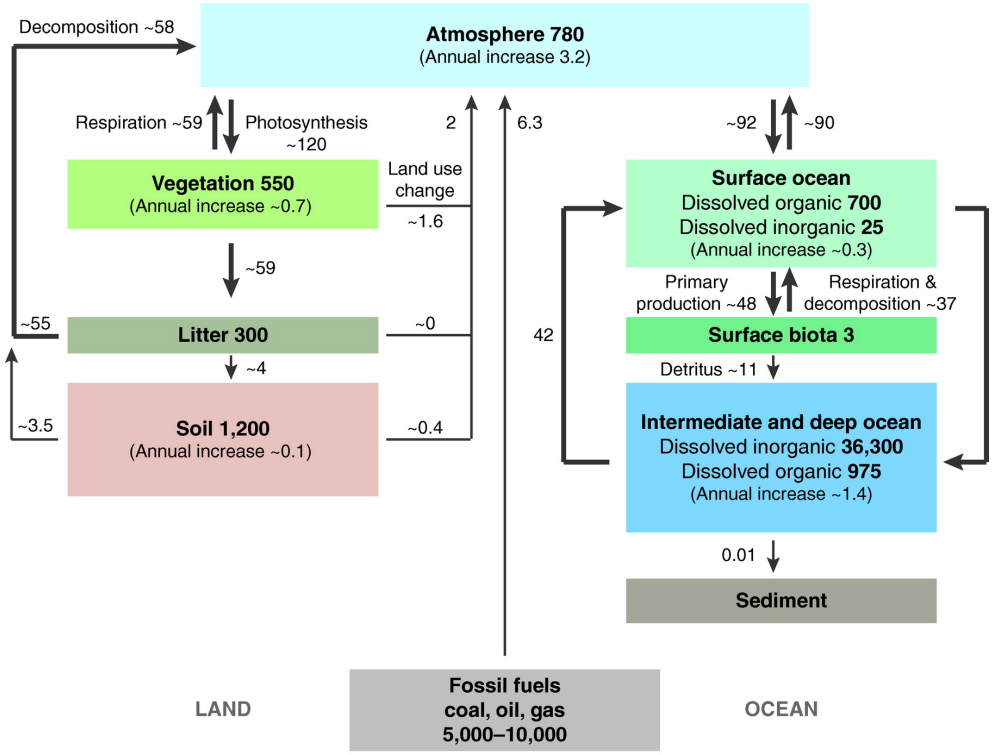

Figure 1.2: The global carbon cycle, including estimates of the capacity of carbon reservoirs (in Pg C) and transfers between them (in Pg C yr^{-1}) during the 1990s. From *Houghton* [2007].

CO_2 emissions, atmospheric growth rate and CO_2 sinks have been combined to form an overall picture of the anthropogenic emissions and their fate, into what is called a 'carbon budget' by the research community (http://www.globalcarbonproject/budget). The latest estimate of the global carbon budget indicates that approximately 10 Pg C were released into the atmosphere in 2010, continuing an upward trend that has persisted throughout most of the 20th Century [*Peters et al.*, 2012]. The terrestrial and oceanic sinks absorbed just over half of these emissions within sub-decadal timescales, leaving an airborne fraction of between 40% and 50% [*Sabine et al.*, 2004; *Jones and Cox*, 2005; *Le Quéré et al.*, 2009]. The amount of carbon absorbed by the sinks has been increasing in line with the increasing atmospheric growth rate (Figure 1.3, page 7). There are concerns however that the efficiency of the sinks may have decreased in recent decades [*Canadell et al.*, 2007; *Le Quéré et al.*, 2007; *Schuster and Watson*, 2007; *Raupach et al.*, 2008] leading to an increase in the airborne fraction of CO_2 emissions and possibly to an amplification of climate change [*Friedlingstein et al.*, 2006]. However, this finding is debated both because of the high uncertainty in airborne fraction [*Knorr*, 2009], and of the importance

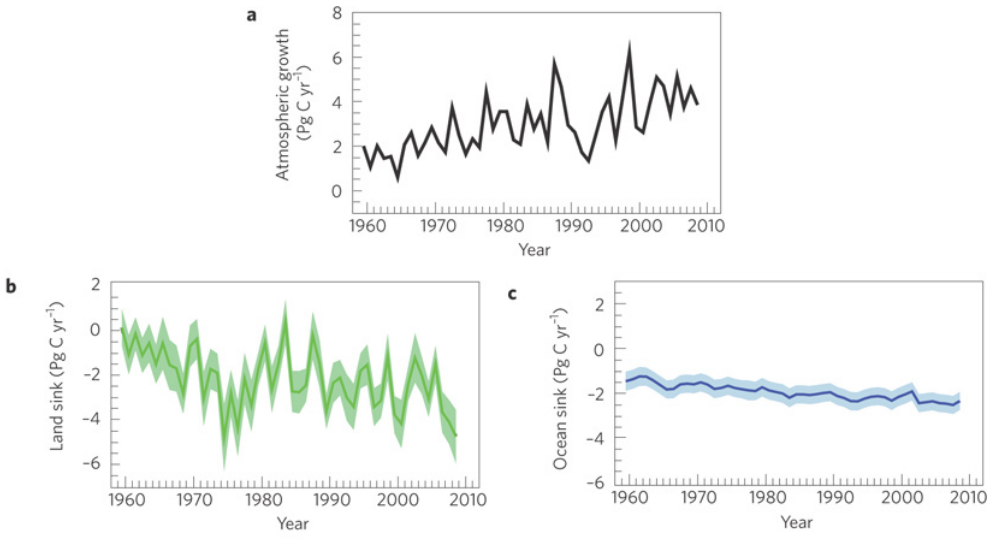

Figure 1.3: Trends in the rates of growth in (a) atmospheric carbon, (b) the land carbon sink, and (c) the oceanic carbon sink. Sinks are negative to show their mitigating effect on the growth in atmospheric carbon. From *Le Quéré et al.* [2009].

of emissions trajectories and natural variability for airborne fraction trends [*Gloor et al.*, 2010]. Calculating the individual components of the carbon cycle is subject to large uncertainties due to high natural variability of the carbon cycle on interannual time scales and gaps in the knowledge of transfers of carbon between the different parts of the carbon cycle [*Le Quéré et al.*, 2009]. This translates into large uncertainties in estimates of the magnitude and long-term trend of the airborne fraction of anthropogenic CO_2, and consequent difficulties in creating highly detailed predictions of the likely future impacts of emissions on the Earth's climate.

The controversy over the trend in airborne fraction highlights the difficulties in accurately tracking the complex dynamics of the carbon cycle and the subsequent effects on future climate change. Increasing our understanding of the oceanic and terrestrial carbon sinks and quantifying their response to global and regional climate change and variability is therefore fundamental to current climate change research.

1.3 The oceanic carbon sink

1.3.1 Overview

The ocean is the Earth's largest reservoir of carbon outside the solid earth, containing approximately 38,000 Pg C or 50 times as much as the atmosphere (Figure 1.2, page 6). CO_2 is exchanged between the atmosphere and ocean as their relative concentrations tend towards equilibrium. Projections of long-term carbon cycle estimate that the airborne fraction of anthropogenic CO_2 after emissions cease will reduce to 20-35% over a time period of two centuries to thousands of years as the atmosphere and oceans reach equilibrium [*Archer et al.*, 2009].

When CO_2 enters the ocean from the atmosphere, it reacts with the water to form carbonic acid (Equation 1.1). The carbonic acid then forms a bicarbonate anion and a proton (Equation 1.2), which further dissociates giving a carbonate ion and another proton (Equation 1.3).

$$CO_2 + H_2O \Leftrightarrow H_2CO_3 \tag{1.1}$$

$$H_2CO_3 \Leftrightarrow H^+ + HCO_3^- \tag{1.2}$$

$$HCO_3^- \Leftrightarrow H^+ + CO_3^{2-} \tag{1.3}$$

The equilibrium form of dissolved CO_2 is therefore:

$$CO_2 + H_2O + CO_3^{2-} \Leftrightarrow 2HCO_3^- \tag{1.4}$$

The combined total of CO_2, CO_3^{2-} and HCO_3^- is known as dissolved inorganic carbon (DIC), 91% of which is in the form of HCO_3^-. Only 1% of DIC remains as CO_2 which can be exchanged with the atmosphere [*Le Quéré and Metzl*, 2004]. CO_2 dissolves more readily in cold water than warm water, and as water is heated it loses some of its capacity to hold CO_2. In equatorial regions the ocean acts as a net source of atmospheric CO_2 (Figure 1.4, page 9). Carbon-rich waters are brought to the surface through upwelling, where the concentration exceeds that of the atmosphere and CO_2 is outgassed. The outgassing is enhanced as the water is warmed. In high latitudes, where the surface waters are cooler, CO_2 is dissolved more readily. This water is then drawn down by ocean transport to form

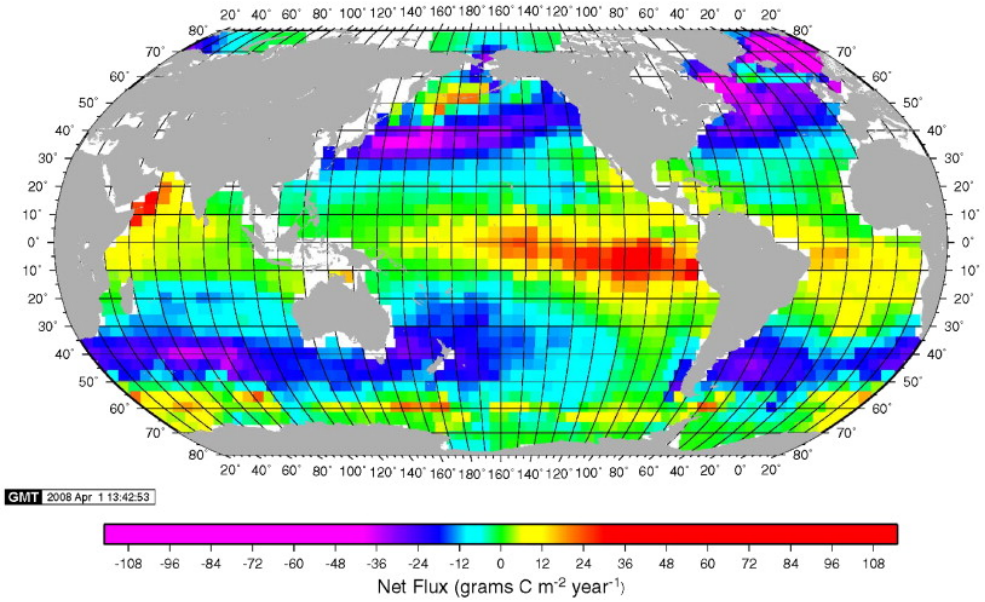

Figure 1.4: Climatological mean annual air-sea CO_2 fluxes for the global ocean. From *Takahashi et al.* [2009].

intermediate and bottom water, to be replaced at the surface by water flowing from the low- and mid-latitudes which cools as it travels polewards. Intermediate and deep waters are therefore richer in carbon than surface waters.

CO_2 transport into the deep ocean in high latitudes occurs as part of the thermohaline circulation, contributing to the ocean's store of DIC. Warm, saline water from the tropics is transported polewards by the Gulf Stream in the Atlantic, where it cools and becomes denser causing it to sink [*McCartney and Talley*, 1984; *Dickson and Brown*, 1994]. The North Atlantic is the largest net sink of atmospheric CO_2 (Figure 1.4) due to factors enhancing the effect of the sinking surface waters. pCO_2 concentrations are relatively low, resulting in a larger air-sea CO_2 gradient and therefore increased ocean uptake. Fluxes are increased still further by the persistent high winds during winter months [*Takahashi et al.*, 2009]. The drawn down water forms North Atlantic Deep Water, which is then transported around the world by deep water currents [*Schmitz*, 1995]. DIC can remain isolated from the surface in this manner for periods ranging from a few decades to over 2,000 years [*Primeau and Holzer*, 2006; *Holzer and Primeau*, 2008].

Approximately 48 Pg C is removed from the surface waters through biological processes each year [Figure 1.2, page 6; *Longhurst et al.*, 1995]. The primary mechanism of export is photosynthetic activity of phytoplankton, converting the dissolved CO_2 into

organic matter which then becomes part of the oceanic food chain. This is termed primary production. 37 Pg C is returned to the surface through respiration and excreta, as well as decomposition of dead organisms. Therefore, approximately 11 Pg C remains and is exported below the surface layers of the ocean each year through biological activity. As these organisms excrete detritus or die, some of the entrained carbon re-dissolves into the water as dissolved organic carbon (DOC) where it is subjected to the same physical processes as DIC [*Hansell et al.*, 2009]. The remainder sinks as particulate matter. The majority of DOC is remineralised to DIC over a period of decades [*Jiao et al.*, 2010], and thus has a similar residence time to DIC.

1.3.2 Anthropogenic changes to the oceanic carbon sink

Much of the change in the oceanic carbon sink will be directly related to increases in the atmospheric CO_2 concentration. As this increases, so will the gradient between the atmospheric and oceanic CO_2 concentration, leading to increased air-sea flux rates. Recent studies show that these fluxes have increased to absorb approximately 25% of the anthropogenic carbon emissions of CO_2 released into the atmosphere each year [*Mikaloff Fletcher et al.*, 2006; *Le Quéré et al.*, 2009]. This increase will continue until the oceanic concentration equilibrates with the atmospheric concentration, which may already be happening in some regions [*Le Quéré et al.*, 2007; *Schuster and Watson*, 2007; *Raupach et al.*, 2008].

Human-induced climate change is likely to have some effect on ocean circulation, which in turn will affect the uptake and subsequent transport of CO_2 from the atmosphere. For example, a hypothetical reduction in salinity in the Southern Ocean caused by increased ice melt [*Broecker et al.*, 1999; *Aiken and England*, 2008] could result in increased stratification and reduced transport of carbon to deep waters [*Caldeira and Duffy*, 2000]. Similarly, climate change effects may reduce the efficiency of the Atlantic Meridional Overturning Circulation with similar reductions in transport of carbon between the surface and the deep ocean [*Schuster and Watson*, 2007; *Ullman et al.*, 2009; *Tjiputra et al.*, 2010].

The rate of biological uptake of CO_2 is not determined by the availability of CO_2, but instead by light, nutrients and minerals such as iron [*Martin et al.*, 1990; *Falkowski et al.*,

1992]. Changes in atmospheric CO_2 and air-sea fluxes caused by anthropogenic activity consequently have little direct impact on the biological portion of the oceanic carbon cycle. However, increasing atmospheric CO_2 can have significant effects on biological production via indirect impacts. Increasing levels of carbon in the ocean lower the pH of the water [*Raven et al.*, 2005; *Bindoff et al.*, 2007; *Doney et al.*, 2009], while burning of fossil fuels increases the direct input of acids into the ocean [*Doney et al.*, 2007]. Ocean acidification leads to reduced growth mass in calcifying organisms and therefore decreased biological uptake of CO_2 from surface waters [*Fabry et al.*, 2008; *Doney et al.*, 2009]. Climate-induced changes in physical ocean properties such as circulation and mixing can also affect the nutrient supply and related biological activity in some ocean regions [*Behrenfeld et al.*, 2006].

1.4 Calculation of the air-sea CO_2 flux

The ability to quantify accurately the air-sea flux of CO_2 is essential if we are to understand the likely long-term effects of anthropogenic CO_2 emissions. The rate at which CO_2 is transferred between the atmosphere and ocean (the flux) is calculated using a simple formula [*Fangohr and Woolf*, 2007]:

$$F = k\ s\ \Delta pCO_2 \tag{1.5}$$

where k is the gas transfer velocity (the rate at which CO_2 is exchanged between the air and water), s is the solubility of CO_2 in the water, and ΔpCO_2 is the difference in partial pressure of CO_2 between the ocean and atmosphere (sea - air). The solubility s is computed from a simple formula based on the salinity and temperature [*Weiss*, 1974]. ΔpCO_2 can be calculated from direct measurements of surface ocean and atmospheric CO_2.

The gas transfer velocity k is less certain. Factors affecting the gas transfer velocity include wind speed [*Liss and Merlivat*, 1986; *Wanninkhof*, 1992], characteristics of surface films [*Frew et al.*, 1990; *Tsai and Liu*, 2003], bubbles [*Asher et al.*, 1996; *McNeil and D'Asaro*, 2007] and precipitation [*Ho et al.*, 1997; *Takagaki and Komori*, 2007]. The

exact contribution of these on the total gas transfer velocity is still the subject of investigation [*Wanninkhof et al.*, 2009]. Direct measurements of gas transfer velocity have been made using a number of techniques. The most common technique involves examining the covariance of eddies in the density of CO_2 and vertical wind velocity [*Smith and Jones*, 1985; *Edson et al.*, 1998; *McGillis et al.*, 2001; *Kondo and Tsukamoto*, 2007; *Yelland et al.*, 2009]. Other approaches such as measuring isotopic fluxes of CO_2 [e.g. *Broecker et al.*, 1986] have seemingly fallen out of favour. None of the direct measurment methods account for all the factors described above, and can give significantly different results [*Broecker et al.*, 1986; *Wesely*, 1986]. Additionally, direct measurements are not widely avaialble; the vast majority of studies record only sea surface pCO_2, and fluxes must be calculated using other data sources.

The majority of air-sea CO_2 flux calculations use gas transfer velocities calculated from a parametrisation of wind speed alone, since it is relatively well understood and wind speed data are available across the entire ocean via remote sensing. Several such paramtrisations have been developed using both laboratory experiments and in situ measurements [*Liss and Merlivat*, 1986; *Wanninkhof*, 1992; *Wanninkhof and McGillis*, 1999; *Nightingale et al.*, 2000; *Ho et al.*, 2006; *Sweeney et al.*, 2007]. Although the choice of a single parameterisation from the available options is in itself a source of uncertainty in global flux estimates [*Ho et al.*, 2006], and a single wind speed relationship is unlikely to be representative across the globe [*Zhang*, 2007], this is the most practical approach given the limits of current knowledge.

1.4.1 Limitations on knowledge of the oceanic CO_2 flux

The main limiting factor in our ability to calculate air-sea CO_2 fluxes is the availability of surface ocean pCO_2 measurements. Other data required to calculate air-sea CO_2 fluxes are readily available at all temporal and spatial scales. Atmospheric CO_2 can be measured using land-based air sampling stations [e.g. *Conway et al.*, 1994]. Measurements do not need to be taken at spatially close locations since the atmosphere is well mixed and measurements can be accurately extrapolated to produce globally complete coverage [*Masarie and Tans*, 1995]. Wind speeds can be derived via remote sensing [*Atlas et al.*, 2011], as

can sea surface temperatures (SST) [*Minnett*, 2001]. Sea surface salinity (SSS) can be obtained from interpolated data sets such as the World Ocean Atlas [*Antonov et al.*, 2006]. SSS has only a small effect on air-sea CO_2 fluxes so using interpolated data is sufficient to calculate flux values. SST and SSS are also frequently measured alongside surface ocean pCO_2 measurements.

Surface ocean pCO_2 can only be measured using under-way sensing equipment installed on ships [e.g. *Cooper et al.*, 1998; *Pierrot et al.*, 2009] or fixed measuring stations [e.g. *Karl and Lukas*, 1996; *Michaels and Knap*, 1996; *González-Dávila et al.*, 2003; *Nemoto et al.*, 2009]. Although pCO_2 sensors are being developed that will be suitable for deployment on autonomous profiling floats and gliders, these have yet to come to fruition [*Johnson et al.*, 2009]. To date, therefore, pCO_2 measurements across the global ocean are only available where ships travel equipped with appropriate devices.

Databases of collected pCO_2 measurements are the basis of the majority of large-scale estimates of the oceanic CO_2 flux. Historically this data has only been released as a calculated climatology [e.g. *Takahashi et al.*, 2002, 2009]. More recently, complete versions of the measurement database have been published, each providing several million individual measurements for analysis [*Takahashi and Sutherland*, 2009; *Pfeil et al.*, submitted]. These databases have allowed scientists to drill deeply into the available surface ocean pCO_2 data for the first time, and also highlighted the limitations of the data collected to date.

There are large ocean regions with very few or no pCO_2 measurements (Figure 1.5, page 14). The lack of available measurements is a significant problem, because unlike atmospheric concentrations oceanic pCO_2 is highly variable on temporal and spatial scales [*Feely et al.*, 1997; *Bates et al.*, 1998; *Cooper et al.*, 1998; *Sarma*, 2003; *Lüger et al.*, 2004; *Li et al.*, 2005; *Midorikawa et al.*, 2006; *Santana-Casiano et al.*, 2007; *Ishii et al.*, 2009; *Tortell and Long*, 2009; *Shadwick et al.*, 2011].

There are regions where there are sufficient pCO_2 measurements to make some assessment of interannual air-sea flux variability. These regions are mostly in the North Atlantic, where equipment has been installed on commercial ships travelling between the United Kingdom and the Caribbean [*Cooper et al.*, 1998; *Schuster and Watson*, 2007] and between Iceland and Newfoundland [*Corbière et al.*, 2007]; the North Pacific [e.g. *Inoue*

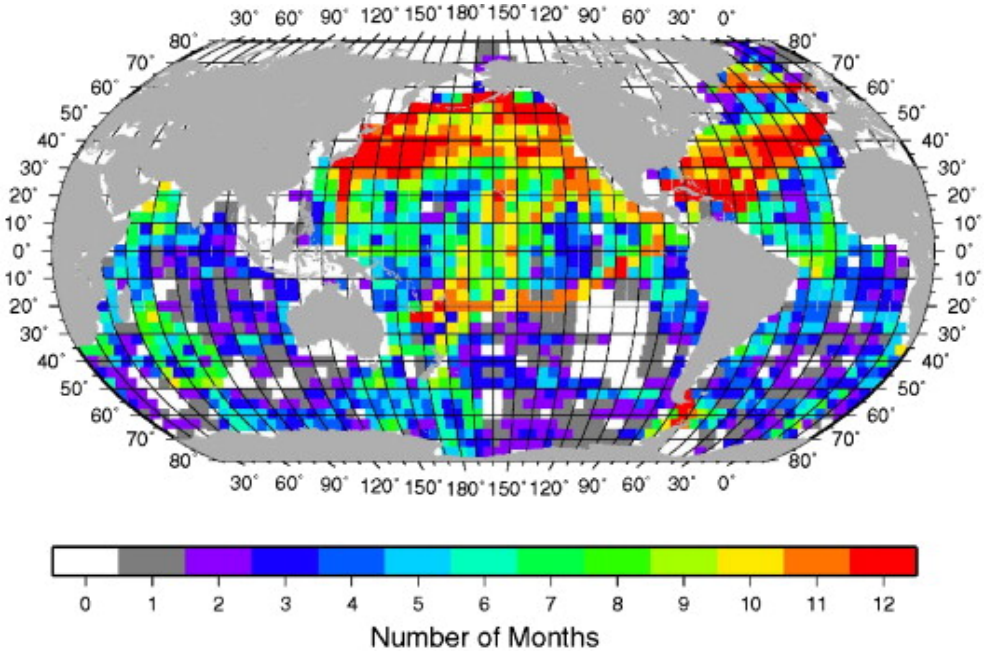

Figure 1.5: The number of months in the calendar year for which measurements are available in the LDEO database of pCO_2 measurements between 1970 and 2008. From *Takahashi et al.* [2009].

et al., 1996; *Midorikawa et al.*, 2003; *Takahashi et al.*, 2006; *Takamura et al.*, 2010]; and the Equatorial Pacific where a network of fixed buoys is deployed [*Feely et al.*, 2002, 2006; *Ishii et al.*, 2009]. However, even in these regions of greatest data coverage a full picture of pCO_2 variability on multiple temporal and spatial scales is missing. In the rest of the global ocean, pCO_2 and air-sea fluxes are largely limited to monthly climatologies [*Takahashi et al.*, 1997, 2002, 2009] and long-term trends in specific regions [*Lefèvre et al.*, 2004; *Takahashi et al.*, 2006; *Metzl*, 2009; *Schuster et al.*, 2009]. While these data products can be combined into global estimates of trends in air-sea fluxes, they show large interannual variability [*Le Quéré et al.*, 2009] that is poorly constrained by the lack of ocean-wide flux data.

1.4.2 Estimating surface ocean pCO_2

A globally and temporally complete data set of interpolated and extrapolated oceanic pCO_2 values would be an invaluable resource for a number of studies. First, it would allow comprehensive air-sea flux rates to be calculated for all sectors of the ocean for an

extended time period, and would provide a much greater insight into the temporal variability of fluxes compared to existing climatologies [e.g. *Takahashi et al.*, 2009]. Second, it could act as a valuable input for model validation studies [e.g. *Le Quéré et al.*, 2009] and as an input of prior estimates for atmospheric inverse methods [*Gurney et al.*, 2002]. There will also be countless opportunities to drill into such a comprehensive data set to examine many characteristics of surface ocean pCO_2 that may have been impossible previously. Advanced interpolation and extrapolation techniques have been developed to estimate temperature [*Locarnini et al.*, 2006], salinity [*Antonov et al.*, 2006] and oxygen [*Garcia et al.*, 2006] anomalies for the world's ocean, but they have not yet been adapted to carbon variables.

1.4.2.1 Interpolation and Extrapolation Techniques

Data interpolated in both space and time are preferable to climatological estimates for many studies since they can encompass interannual and longer term variability. The most common approach to extrapolating pCO_2 data used so far has been to establish relationships between pCO_2 and other oceanic variables and use these relationships to reconstruct pCO_2 for those regions and times where no measurements are available. At its simplest, this approach can establish a simple linear relationship between changes in sea surface temperature (SST) and changes in pCO_2 [*Lee et al.*, 1998; *Boutin et al.*, 1999; *Lefèvre and Taylor*, 2002]. Such simple relationships can be useful if SST is the dominant factor affecting pCO_2 [*Boutin et al.*, 1999; *Shim et al.*, 2007], or if the SST/pCO_2 relationship is known to represent closely the myriad other factors affecting pCO_2 levels, such as biological activity or mixing, although the latter is unlikely [*Park et al.*, 2006]. Whether or not a single parameter can be used as an adequate proxy for pCO_2 depends largely on the characteristics of highly localised regions [*Shim et al.*, 2007]; it is unlikely that this approach will be effective over large spatial scales.

More sophisticated interpolation approaches have been used to establish relationships between pCO_2 and multiple other variables representing aspects of the ocean's biogeochemical cycle. Such approaches typically use a combination of SST, chlorophyll (Chl) and mixed layer depth (MLD) [e.g. *Ono et al.*, 2004; *Jamet et al.*, 2007; *Watson et al.*, 2009], although other indicators such as NO_3 are sometimes included [*Wanninkhof et al.*,

1996]. A multi-variate relationship between all these variables is constructed using multi-linear regression techniques. This provides a more complete representation of the factors affecting pCO_2 levels than a single relationship, which is important where no single factor has a dominant influence. A similar technique uses neural networks to construct a self-organising map of relationships between pCO_2 and related oceanic variables, which can then be used as a reference for inferring missing pCO_2 values from the known values of the related variables [*Telszewski et al.*, 2009; *Watson et al.*, 2009; *Valsala et al.*, 2011]. Still others have constructed relationships utilising principal components analysis [*Lohrenz and Cai*, 2006].

Such multi-variate approaches to estimating pCO_2 values are very useful in studies of limited spatial and temporal extent. Accuracy tests show that, where there are plenty of pCO_2 measurements, the interpolated pCO_2 values can be very accurate: mean errors of 10-15 μatm are typical [*Jamet et al.*, 2007; *Telszewski et al.*, 2009], though errors less than 2 μatm have been reported in some instances [*Watson et al.*, 2009]. However, the relationships established may not be representative across large regions or extended time periods. This is particularly evident in *Wanninkhof et al.* [1996], where the errors in the interpolated pCO_2 values vary between 1 μatm and 48 μatm. Multiple relationships at different spatial locations and times would be required to achieve consistently accurate results using regression techniques. Some studies have begun to break down the relationships for seasonal variations [e.g. *Lee et al.*, 1998; *Jamet et al.*, 2007], but extending the analysis to encompass geographical and interannual variations [e.g. *Park et al.*, 2006] is much less common.

1.5 Interannual variability of the climate and the carbon cycle

Changes in the Earth's climate system over periods of one to a few years are frequently discussed in terms of large scale modes of climate variability. These are typically defined in relatively simple terms, such as teleconnections in atmospheric pressure systems or changes in sea surface temperature in specific regions.

The mode of climate variability which has the largest global effect is the El Niño Southern Oscillation (ENSO). ENSO indices are a measure of sea surface temperature

(SST) anomalies in the Equatorial Pacific [*Trenberth*, 1997]. The indices are complemented by the Southern Oscillation Index (SOI), a measure of the difference in pressure between Darwin and Tahiti [*Ropelewski and Jones*, 1987]. This index is representative of the atmospheric portion of ENSO [*Trenberth*, 1997]. SSTs in the Equatorial Pacific are typically characterised by cold waters upwelling from the eastern Equatorial Pacific that propagate across the Pacific to form what is known as the 'cold tongue' [*Yulaeva and Wallace*, 1994]. El Niño events occur at periods of two to seven years and last usually 6-12 months. They are characterised by a weakening of the trade winds, allowing warm waters to extend from the western Pacific and suppress the upwelling in the east [*Feely et al.*, 1987]. The shift from cold to warm surface waters during El Niño events in turn causes significant changes in the overlying atmosphere and can affect weather systems around much of the globe [e.g. *Joseph et al.*, 1994; *Chiew et al.*, 1998].

The North Atlantic Oscillation (NAO) is defined as an atmospheric teleconnection similar to the SOI, measured as the difference in sea level pressure between Iceland and the Azores. This variation influences surface temperatures, winds and precipitation, although the effects are much more geographically limited than those of ENSO and the SOI [*Hurrell*, 1995].

Large-scale climate variability in the North Pacific cannot be measured in terms of a single climatic phenomenon as in other regions. However, there is a consistent variation in climate that can be detected in a number of measured properties such as SST, sea level pressure, air temperatures and precipitation known as the Pacific Decadal Oscillation (PDO) [*Mantua et al.*, 1997].

1.5.1 Effects on surface pCO_2

The modes of climate variability described above are most commonly associated with their meteorological effects, but also have an influence on the carbon cycle. The ENSO indices are closely related to surface pCO_2 in the Equatorial Pacific as upwelling of carbon-rich water is suppressed during El Niño events and enhanced during La Niña events [*Feely et al.*, 1987; *Inoue and Sugimura*, 1992]. The SOI is closely related to ENSO but the two indices are opposite in sign. Changes in the SOI index therefore lead to an opposing relationship to pCO_2, but the strength of the correlation is nonetheless similar to that of ENSO

[*Inoue et al.*, 1996; *Etcheto et al.*, 1999; *Feely et al.*, 1999, 2002, 2006; *Sheu et al.*, 2010].

The PDO's influence over pCO_2 in the North Pacific is not so clearly defined as ENSO's effects on the Equatorial Pacific. However, the changes in several climatic variables related to the PDO are expected to have some effect on pCO_2. The changes in climate are related to variations in vertical mixing rates and biological activity, both of which are known drivers of surface pCO_2 variability [*McKinley et al.*, 2006; *Valsala et al.*, 2011].

The link between the NAO and pCO_2 levels in the North Atlantic is less certain and the subject of considerable debate. Model-based analyses tend to show a relationship caused by increasing MLD and decreasing SST in positive phases of the NAO [*Thomas et al.*, 2008; *Ullman et al.*, 2009]. The link tends to vary across the basin, with a dipole between the sub-tropics and sub-polar regions [*Le Quéré et al.*, 2000]. Analysis of the NAO based on observational data give differing results. Some have found links between the NAO and pCO_2 in at least some areas of the North Atlantic, again related to changes in MLD and SST [*Gruber et al.*, 2002; *Corbière et al.*, 2007; *Schuster et al.*, 2009; *McKinley et al.*, 2011]. Other studies have concluded that no such link exists [*Schuster and Watson*, 2007; *Watson et al.*, 2009; *Padin et al.*, 2011].

Quantifying the effects of climate variability on surface pCO_2 is a critical step in understanding how the processes represented by the indices described above affect air-sea CO_2 fluxes. Such links may be either direct, through changes in sea surface temperature and hence the solubility of CO_2 in sea water, or indirect through variations in ocean circulation, biological processes or wind speed and thus the gas transfer velocity. The systematic collection and collation of global pCO_2 measurement data will ease efforts to understand such links.

1.6 Aim of this research

The aim of this research is to extend the current knowledge of surface ocean pCO_2 variability in both spatial and temporal dimensions. While this topic has been studied in the past, the scope has generally been limited in time or space. The recent release of large databases of pCO_2 measurements from multiple data sets spanning the entire globe over more than 30 years [*Takahashi and Sutherland*, 2009; *Pfeil et al.*, submitted] provide an

invaluable opportunity to extend the analysis of pCO_2 variability to a global extent and multiple time scales. Specifically, this thesis aims to identify:

- Over what spatial and temporal scales do surface pCO_2 levels vary?
- Which underlying drivers have the greatest influence of spatial pCO_2 variability?
- How does spatial variability of air-sea fluxes compare with that of sea surface sea surface pCO_2?
- How much does surface pCO_2 change in response to large-scale modes of interannual climate variability?
- In which regions are the relationships between pCO_2 and modes of climate variability most evident?
- Which areas of the global ocean lack sufficient measurements to accurately assess pCO_2 and air-sea CO_2 flux variability?

Chapter 2 presents an analysis of the spatial and temporal variability of the available pCO_2 measurements in the form of autocorrelations. This information can be used to inform strategies for determining pCO_2 sampling rates in space and time [*Sweeney et al.*, 2002; *Lenton et al.*, 2009], and as a prior estimate of variability for inverse modelling techniques based on data assimilation [e.g. *Rödenbeck et al.*, 2003]. The chapter also contains an analysis of other oceanic variables to determine which aspects of the ocean system dominate the spatial pCO_2 variability. Finally it examines the spatial variability of air-sea fluxes.

Chapter 3 describes a new method to construct a spatially and temporally complete data set of interpolated surface ocean pCO_2 based purely on statistical techniques. The data set will be used to examine pCO_2 variability in response to modes of climate variability in Chapter 4. This method avoids the pitfalls of relationships to other oceanic variables which are less accurate on large spatial and temporal scales. The method also provides uncertainty estimates for the interpolated pCO_2. Finally, the interpolation technique is validated against a model output data set to assess its accuracy. The method developed in this thesis will be useful not only for this study, but will also have other applications in the wider community.

Chapter 4 analyses the interannual variability of pCO_2 concentrations in relation to large-scale modes of climate variability (e.g. the El Niño Southern Oscillation and North Atlantic Oscillation). The relationship is examined using the correlation of pCO_2 levels to the climate indices as a measure of the strength of the link, its magnitude, and spatial extent.

Finally, the Summary and Conclusion chapter reviews the findings of the research, reflects on its efficacy, and provides suggestions for improving and extending the work of the thesis in future.

1.6.1 Data Used

This thesis utilises data from both the Lamont-Doherty Earth Observation (LDEO) database [*Takahashi and Sutherland*, 2009] and Surface Ocean CO_2 Atlas (SOCAT) database [*Pfeil et al.*, submitted]. Only the LDEO database was available at the outset of this study, so that was used for Chapter 2 and Chapter 3. The SOCAT database was released in September 2011. The SOCAT project implements a more thorough and transparent quality control process than the LDEO database, so that database is therefore deemed a more reliable data product. The SOCAT database is used for the analysis in Chapter 4.

Chapter 2

Autocorrelation Characteristics of Surface Ocean pCO_2

Autocorrelation characteristics of Surface Ocean pCO_2

Steve D. Jones[1], Corinne Le Quéré[2], Christian Rödenbeck[3]

[1] School of Environmental Sciences, University of East Anglia, Norwich, UK.

[2] Tyndall Centre, University of East Anglia, Norwich, UK.

[3] Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, Jena, Germany.

Published in Global Biogeochemical Cycles, doi:10.1029/2010GB004017.

All the work and analysis presented in this chapter was undertaken by Steve Jones, primarily to support the work performed in the remainder of this thesis. The co-authors on the paper provided guidance and suggestions regarding analyses and results that would be of most use and interest to the wider scientific community.

2.1 Abstract

Understanding the variability and coherence of surface ocean pCO_2 on a global scale can provide insights into its physical and biogeochemical drivers and inform future sampling strategies and data assimilation methods. We present temporal and spatial autocorrelation analyses of surface ocean pCO_2 on a 5°x5° grid using the LDEO database. The seasonal cycle is robust with an interannual autocorrelation of $\sim$0.4 across multiple years. The global median spatial autocorrelation (e-folding) length is 400$\pm$250 km, with large variability across different regions. Autocorrelation lengths of up to 3,000 km are found along major currents and basin gyres while autocorrelation lengths as low as 50 km are found in coastal regions and other areas of physical turbulence. Zonal (east-west) autocorrelation lengths are typically longer than their meridional counterparts, reflecting the zonal nature of many major ocean features. Uncertainties in spatial autocorrelation in different ocean basins are between 42% and 73% of the calculated decorrelation length. The spatial autocorrelation length in air-sea fluxes is much shorter than for pCO_2 (200$\pm$150 km) due to the high variability of the gas transfer velocity.

2.2 Introduction

The ocean is estimated to absorb approximately 25% of the total anthropogenic emissions of carbon dioxide (CO_2) released into the atmosphere every year [*Mikaloff Fletcher et al.*, 2006; *Le Quéré et al.*, 2009]. The partial pressure of CO_2 at the ocean surface (pCO_2) is a fundamental determinant of the rate at which CO_2 is absorbed by the ocean [*Fangohr and Woolf*, 2007]. Thus, understanding the spatial and temporal variability of surface ocean pCO_2 is critical to understanding the interaction between the atmospheric and oceanic carbon cycles.

There are several methods of assessing the variability of a physical variable such as pCO_2 on different scales, such as direct comparison of adjacent measurements [e.g. *Bates et al.*, 1996; *Jiang et al.*, 2008] and frequency domain analysis [*Lenton et al.*, 2006]. While there have been many recent studies of pCO_2 variability, the paucity of available measurements has limited their extent. The majority of spatial studies have focused on relatively small regions [*Jiang et al.*, 2008; *Ishii et al.*, 2009; *Krasakopoulou et al.*, 2009;

Santana-Casiano et al., 2009; *Zhang et al.*, 2010] or individual/repeated cruise tracks [*Fransson et al.*, 2009; *Padin et al.*, 2010]. Temporal studies are also restricted to specific regions, and are additionally limited in scope, focussing on either sub-daily [*Bates et al.*, 1998; *Dai et al.*, 2009; *Leinweber et al.*, 2009] or seasonal timescales [*Sarma*, 2003; *Shim et al.*, 2007; *Olsen et al.*, 2008; *Litt et al.*, 2010]. A small number of previous studies have attempted to produce a global perspective on pCO_2 variability. The lack of long-term measurement projects has prevented interannual analysis in most cases, although there are exceptions where fixed stations have been deployed [*Bates et al.*, 1996; *Gruber et al.*, 2002; *Cosca et al.*, 2003; *Wong et al.*, 2010]. Some attempts at global assessments of pCO_2 variability have been undertaken despite these limitations. *Li et al.* [2005] produced global maps of spatial autocorrelation lengths from surface ocean pCO_2 based on previous data sets of pCO_2 measurements, but their analysis was made on a coarse 10°x10° grid and their results were restricted to variability on scales of $<\sim$1,000 km only. These limitations both reduced the ability to discern long-scale autocorrelations and restricted detection of finer detail. *Sweeney et al.* [2002] examined the decorrelation lengths for a selection of cruise tracks with a view to estimating desired spatial sampling rates for future observation projects, but a full global analysis was not attempted.

This paper presents a global assessment of the spatial and temporal variability of surface ocean pCO_2, based on the measurements from the Lamont-Doherty Earth Observatory database [*Takahashi and Sutherland*, 2009] (hereafter referred to as the LDEO database). This work expands on previous analysis by using a more extensive data set, by looking at much larger spatial scales limited only by the length of individual cruise tracks, and by examining the directional features of the autocorrelation characteristics. The factors controlling the spatial variability of pCO_2 are identified by decomposing the pCO_2 signal into its temperature and residual components [*Takahashi et al.*, 2002]. The study is extended to cover the spatial variability of air-sea CO_2 fluxes. The differences in spatial variability between pCO_2 and CO_2 fluxes are identified and discussed. Finally the influence of other external drivers (winds, ocean circulation, biology) on CO_2 variability is also examined directly or through the analysis of proxy variables (sea surface height, surface chlorophyll). This study thus provides a global view of the spatial and temporal coherence of surface ocean CO_2 data, their underlying controls, and their correspondence

with the signatures of known physical and biological drivers.

The global view of oceanic pCO_2 variability presented here can inform strategies for determining sampling rates in both space and time [*Sweeney et al.*, 2002; *Lenton et al.*, 2009]. It will also prove useful in a number of modelling projects: sea-air CO_2 fluxes can be calculated using this data, which in turn can be used as prior estimates of ocean variability for inverse modelling techniques based on data assimilation [*Rödenbeck et al.*, 2003]. Furthermore, knowledge of the autocorrelation characteristics of pCO_2 can inform advanced methods of interpolating the sparse measurements available, such as those used for other physical and geochemical variables [*Levitus*, 1982]. This will provide an important improvement over the necessarily less comprehensive interpolations performed to date [*Lefèvre and Taylor*, 2002; *Takahashi et al.*, 2002; *Schuster et al.*, 2009].

2.3 Data preparation

The LDEO database consists of ∼4.1 million individual surface ocean pCO_2 measurements spanning the period 1968-2008. Outliers were detected and removed from the data to reduce the influence of erroneous entries caused by transcription errors or faulty instrumentation. The remaining measurements were converted into two separate formats for temporal and spatial autocorrelation.

The autocorrelation calculations performed in this analysis are based on two surface ocean pCO_2 data products: time series for each 5°x5° grid cell and ship track data. This section describes the treatment of the data necessary to construct the data products and their rationale. The method will be described in detail in Section 2.4 (page 27).

2.3.1 Data for temporal autocorrelation

To compute the temporal autocorrelation, the ocean was divided into 5°x5° grid cells and time series constructed for each. This grid size represents a compromise between a high-resolution analysis and the limitations of the available data. Daily mean values were calculated for each cell to produce a time series spanning the complete time period of the data set. For leap years, a 'day' was calculated as $1\frac{1}{365}$ calendar days, to produce a constant year length of 365 days throughout. While calculating the daily mean value for each grid cell, any measurements falling outside three standard deviations of the mean

were flagged as outliers in an iterative process, repeated until no outliers were detected. 17,952 measurements (0.004%) were flagged as outliers in this manner. Further outliers were removed by examining the complete daily time series for each grid cell as follows. A linear trend for the time series was calculated and temporarily removed. Any day whose mean pCO_2 level fell outside three standard deviations of the mean was flagged in an iterative process, again repeated until no outliers were detected. A total of 268 days' measurements (0.007%) were flagged across all grid cells.

The measurements flagged as outliers were removed from both the binned and original data sets, which were then used as the basis for the temporal and spatial autocorrelation analysis respectively.

2.3.2 Data for spatial autocorrelation

Calculating the spatial variability of the pCO_2 measurements requires a set of data with sufficient spatial coverage over the ocean. Using the gridded data set created for the temporal autocorrelation above was not suitable: in any given day or month there was insufficient coverage to calculate the variability, and combining multiple maps from the gridded data set to produce sufficient spatial coverage would artificially increase the variability of the data as pCO_2 levels changed over time. Using gridded data also restricts the spatial resolution of the final autocorrelation calculation, and prevents the detection of small-scale variability.

The LDEO database is constructed from measurements taken along individual cruise tracks. Each cruise represents a suitable data set for assessing the spatial variability of pCO_2 in the region through which it passes, with most cruises made up of several hundred measurements logged to sub-kilometer precision. Each cruise's measurements are also taken closely together in time, thus minimising the effect of temporal variations in pCO_2. Calculating the spatial autocorrelation for each cruise's data, then projecting the results onto a gridded map, allows a global assessment of the spatial variability of pCO_2 levels to be created.

Unfortunately, the LDEO database does not identify individual cruises: it lists only the institutes or scientists who collected the data. Extraction of specific cruise information was therefore performed by analysing the characteristics of the data as follows. The

measurements provided from a single source were grouped together and sorted by date and time. Where two consecutive measurements were taken within 10 days, both were assumed to be from the same cruise period; greater time periods between measurements were treated as boundaries between separate cruises. The 10-day period was chosen to provide a balance between maintaining coherent cruises, and accounting for reduced correlations due to large time differences between measurements. The measurements from each of these periods were split into cruises by assessing their geographical proximity. For any pair of measurements to be considered as part of the same cruise, they could not be separated by more than the distance a ship is likely to travel in the time between the two measurements. The threshold was set at a rate of 1,500 km per day, equating to an average speed of 33 knots. While this is faster than most ships can travel, it provides some flexibility to account for errors in the recorded measurement positions and/or times. Even this threshold was not sufficient to capture accurately all cruises: in some cases, all the measurements for a cruise are recorded on a single date or at short fixed intervals, presumably where accurate time records were not available. These cruises would be split erroneously using the above threshold, and so were identified and processed manually. Finally, any cruise containing fewer than five measurements was discarded, as this was a strong indication of errors in the original data such as mis-recorded ship positions. This yielded a total of 1,535 individual cruises from the LDEO database.

2.4 Method

2.4.1 Temporal autocorrelation

A temporal autcorrelation function (ACF) for each 5°x5° grid cell was calculated at monthly resolution. Since no grid cells contained a complete time series, the autocorrelations could be computed only where the original and lagged time series contained pairs of values at the same time steps. In some cases the number of paired time steps was very small, so a measure of the statistical significance of the result was required to ensure that the results were robust. The statistical significance of each ACF value was calculated as a function of the number of time steps used in the calculation using the formula:

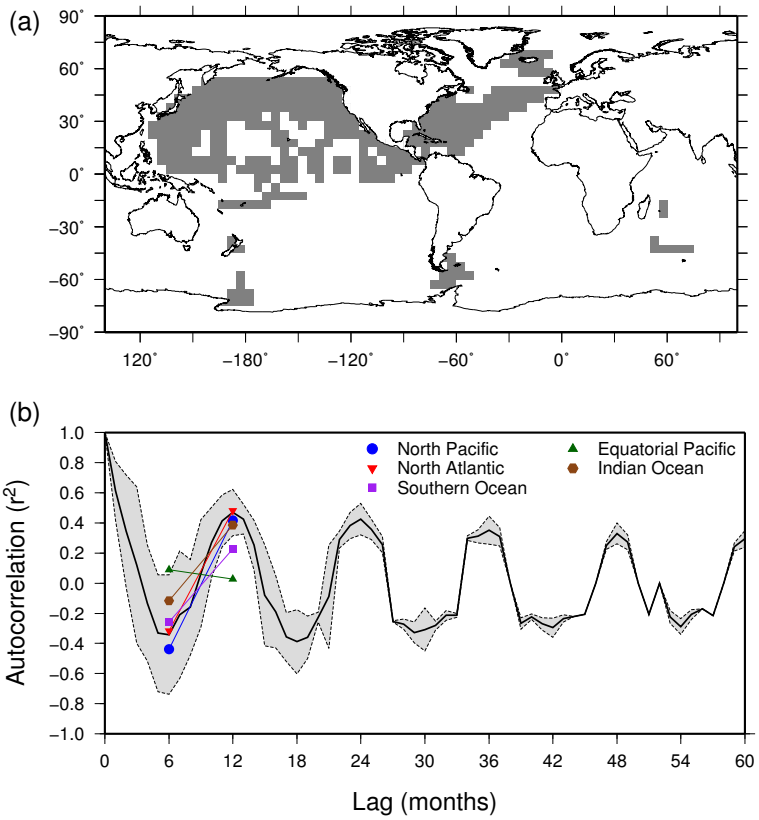

Figure 2.1: (a) Map of the grid cells that pass the statistical significance test on the monthly temporal ACF. (b) The mean monthly temporal ACF calculated from all grid cells. The gray shaded area indicates one standard deviation either side of the mean. The symbols show the progression of the ACF between 6 and 12 months in different regions, to indicate the relative influence of the seasonal cycle.

$$T = \frac{Q\left(\frac{1}{2}(1+0.95)\right)}{\sqrt{n}} \tag{2.1}$$

where T is the threshold of statistical significance, Q is the quantile of the cumulative distribution function of the normal distribution at 95% [*Wichura*, 1988] and n is the number of values used to construct the ACF. This function gives a threshold between 0 and 1. Individual ACF values between T and 0 (either positive or negative) are not statistically significant; any such values in the calculated ACFs were discarded. A value of $T \leq 0.5$ was used for this study, since the number of monthly values available for a given grid cell is relatively low. A total of 348 grid cells had ACFs containing statistically significant values at the 95% level (Figure 2.1a). The majority of these were in the North Pacific and North Atlantic, where most pCO_2 measurements are available [*Takahashi et al.*, 2009].

Temporal ACFs at daily resolution were also calculated for each grid cell, but too few

cells produced statistically significant ACFs to allow a robust analysis of the results.

2.4.2 Spatial autocorrelation

Spatial autocorrelation functions were calculated for each cruise in the LDEO database using the Moran's I technique [*Moran*, 1950], comparing the similarity of pairs of measurements within the cruise. Autocorrelation values for the cruise were calculated in distance groups of 50 km. For the 0-50 km bin, pairs of measurements separated by 50 km or less were assessed to give an autocorrelation value for the cruise at a distance lag of 50 km. Next, pairs of measurements separated by 50-100 km were examined and so on, to build a complete ACF covering the full distance of the cruise. This approach limited the smallest detectable autocorrelation length to 50 km, which meant that some detail was lost around coastlines where autocorrelations are likely to be very short. However, this was necessary to reduce the amount of computation required for the analysis to a feasible level.

Any cruise from the LDEO database that covered a distance of 50 km or less was discarded from the analysis. Similarly, any cruise with a correlation length of within 100 km of the overall cruise distance was also discarded as it is likely that the correlation length was limited by the length of the cruise. A total of 1,454 cruises remained for the spatial autocorrelation analysis. The Moran's I technique includes an assessment of the statistical significance of its results. Any value that fell below the threshold of 95% significance was discarded. The decorrelation length of the measurements from each cruise was determined by the e-folding length of the ACF.

Many of the cruises in the data set pass through different water masses, meaning that the ACF for each cruise represents the combined autocorrelation characteristics of all the water masses encountered and any variability between them is hidden. The autocorrelation analysis for each cruise was extended to reveal this variability. For each grid cell through which the cruise passed, an ACF was calculated for the measurements taken within a reference distance of the center of the cell. The reference distance was set at five times the e-folding length of the original ACF calculated for the entire cruise.

Global maps of spatial autocorrelation lengths were produced using the e-folding lengths of the ACFs calculated for the individual grid cells. Where more than one cruise

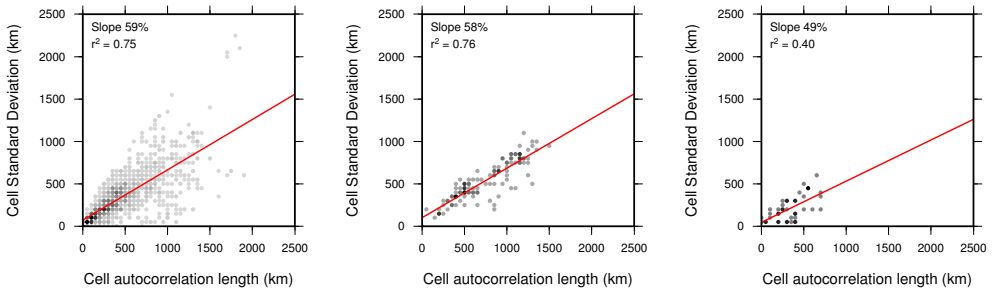

Figure 2.2: Examples of scatter plots used to estimate uncertainties of spatial autocorrelation lengths for (left) all cruises in the global ocean, (middle) zonal cruises in the eastern North Pacific, and (right) meridional cruises in the eastern North Pacific. The mean autocorrelation length is plotted against the standard deviation of contributing cruises for each grid cell, and a linear fit is made to estimate the relationship between the two. The steepness of the slope is converted to a percentage, which is used for the uncertainty.

had an ACF for a given grid cell, the mean e-folding length of all those ACFs cruises was calculated to determine the spatial autocorrelation length for that particular cell. This produced a map of decorrelation lengths for each 5°x5° grid cell. An accompanying map showing the number of cruises contributing to each cell's value was constructed to provide a measure of the confidence level for each cell. Additional maps were produced to show directional autocorrelations. A zonal map was computed using the 571 cruises travelling within 30° of the east-west direction, and a meridional map from the 521 cruises travelling within 30° of the north-south direction.

Assessing the uncertainty of the spatial autocorrelation lengths was difficult because very few cruises contribute to each grid cell over much of the ocean (Figure 2.4b, page 35). An estimate of the uncertainty for each grid cell was calculated as follows. The standard deviation of the autocorrelation lengths that were used in each grid cell was plotted against the mean autocorrelation length calculated for that cell, and a linear fit applied to the scatter plot. The slope of the fitted line was converted to an uncertainty expressed as a percentage of the grid cells' mean autocorrelation length. The uncertainty estimates were calculated for the global ocean as well as smaller ocean regions. Examples of the scatter plots and fitted slopes are shown in Figure 2.2. The linear fit used to estimate the uncertainty was robust, as illustrated by the r^2 values of the linear fits (Table 2.1, page 34).

We examined the pCO_2 autocorrelation lengths in greater detail by extracting the temperature-driven component of the pCO_2 measurements, calculated as pCO_2 at a constant temperature, and a residual component representing the effect of all other processes.

Following the method of *Takahashi et al.* [1993], pCO_2 has been observed to vary with temperature at the rate:

$$\frac{\delta ln\, pCO_2}{\delta T} = 0.0423\,^{\circ}C^{-1} \tag{2.2}$$

This allows each pCO_2 measurement to be decomposed into temperature and residual components:

$$pCO_2 = pCO_2^T + pCO_2^{residual} \tag{2.3}$$

where:

$$pCO_2^{residual} = pCO_2 \ * \ exp\left[0.0423\left(\overline{T} - T\right)\right] \tag{2.4}$$

where pCO_2 and T are the in situ pCO_2 and SST measurements respectively, and $\overline{T}$ the global mean sea surface temperature (20.29°C) cell calculated from Level-3 Standard measurements from the Aqua-MODIS satellite provided by NASA/GFSC/DAAC (http://oceancolor.gsfc.nasa.gov). Spatial autocorrelation maps of each component were produced for direct comparison.

2.4.3 Autocorrelation of drivers

Spatial autocorrelation analysis was also performed on other ocean properties to determine possible drivers for the autocorrelation of pCO_2 values. We used Chlorophyll data from the SeaWiFS satellite, sea surface temperature (SST) data from the MODIS satellite, and sea surface height (SSH) data from AVISO. The latter was used as a proxy for ocean currents, since spatial gradients in SSH are a strong indicator of current strength and direction [*Imawaki et al.*, 2001; *Van Sebille et al.*, 2010]. These are gridded data sets covering multiple years. To eliminate the influence of seasonal cycles and trends, a single grid was produced for each data set containing the temporally averaged data from the whole data set.

The nature of gridded data sets means that they cannot be used to detect very short decorrelation lengths unless they are of very high resolution, at which point the computation requirements of the Moran's I technique become unamanageable. However, using a

coarse grid allows an approximation of the spatial ACF for each grid cell to be obtained while maintaining realistic computation times. We used 1°x1° grids for each of the data sets, and the decorrelation limit was set to 0.1 instead of the e-folding length to compensate for the larger distances between data points. Even so, the minimum detectable decorrelation length was 200 km with spatial autocorrelation lag steps of 100 km instead of the 50 km obtained for the pCO_2 autocorrelation.

2.4.4 Spatial flux autocorrelation

Spatial autocorrelation analyses were also performed on air-sea CO_2 fluxes. Instantaneous CO_2 flux values were calculated for each of the individual measurements using the standard formulation:

$$F = k\, s\, \Delta pCO_2 \tag{2.5}$$

where k is the gas transfer velocity, s the solubility, and ΔpCO_2 the difference between the atmospheric and oceanic pCO_2. The gas transfer velocity k was calculated using the wind formulation by *Wanninkhof* [1992] with bomb ^{14}C corrections by *Sweeney et al.* [2007]. 6-hourly wind data were taken from the ERA-Interim Reanalysis [*Simmons et al.*, 2007] for measurements from 1989 onwards, and from the ERA-40 Reanalysis [*Uppala et al.*, 2005] for measurements prior to that date. The solubility s was calculated according to the method presented by *Weiss* [1974], using the *in situ* temperature and salinity value from the LDEO database. The Hadley Centre's EN3 data set [*Ingleby and Huddleston*, 2007] were used where salinity data were missing from the LDEO database. The atmospheric pCO_2 levels used to calculate ΔpCO_2 were taken from the corresponding latitude in the GLOBALVIEW atmospheric CO_2 database [*GLOBALVIEW-CO2*, 2008] for measurements from 1979 onwards, and from the Mauna Loa record [*Keeling et al.*, 1976] for measurements prior to 1979. Barometric pressure values were taken from the *in situ* measurements recorded in the LDEO database.

The spatial autocorrelation of air-sea flux values was calculated in exactly the same manner as for the pCO_2 values, using the same set of 1,454 cruises. Autocorrelation maps were also produced for each of the flux components k, s and ΔpCO_2 to see which had the greatest influence in determining the flux decorrelation scales.

2.5 Results and discussion

2.5.1 Temporal autocorrelation

The monthly temporal ACF shows almost no sub-seasonal variability, with a dominant seasonal cycle (Figure 2.1b, page 28). The e-folding length of this ACF falls between the first and second months. The 12-month autocorrelation is $\sim$0.46. The interannual correlation decays only very slowly ($\sim$0.33 after four years), indicating that the seasonal cycle is consistent and robust. The ACF from the original data is indistinguishable from the ACF computed from the observations with the long-term trend removed. This means that the slow decay of the temporal ACF is not due to the trend in pCO_2 levels, but is caused by other sources of inter-annual variability.

The prominence of the seasonal cycle is not consistent across all regions. Examining the six- and twelve-month lags in the ACF for five major ocean regions (Figure 2.1b, page 28) shows that the seasonal cycle is strong in the north Pacific and north Atlantic, and slightly less influential in the Indian and Southern Oceans (although there is much less data available in these regions). In the equatorial Pacific, a seasonal cycle is not evident at all. This is consistent with previous analyses of the seasonal cycle of pCO_2 levels [*Takahashi et al.*, 2009].

2.5.2 Spatial autocorrelation

The decorrelation lengths calculated for each grid cell range between 50 km and 3,150 km (Figure 2.4a, page 35), with a median of 400 km and 25%/75% quantiles of 200 km and 650 km respectively. This reflects the large variability of the world's oceans. The zonal and meridional mean decorrelation lengths (Figure 2.4b, page 35) are 450 (250-850) km and 350 (200-550) km respectively. Zonal decorrelation lengths are frequently longer than their meridional counterparts (Figure 2.3, page 34) because many ocean currents run east-west, resulting in a zonal transport of water with similar characteristics in most regions.

The uncertainties for the autocorrelation lengths were calculated in seven ocean regions as well as globally (Table 2.1, page 34). The global mean uncertainty for the map

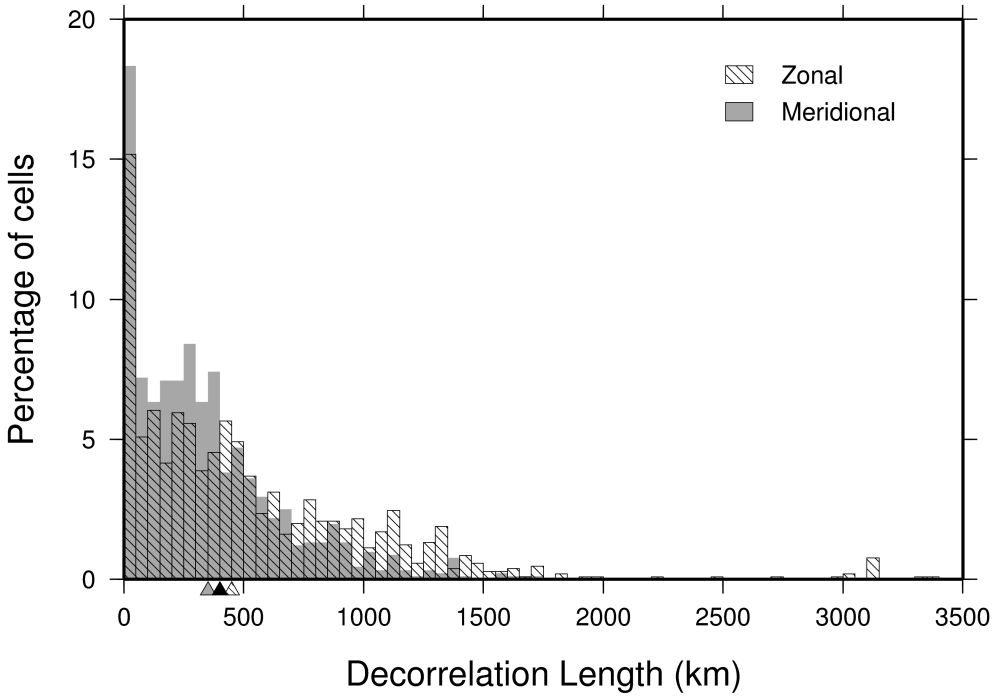

Figure 2.3: Histogram showing the frequency of zonal (diagonal stripe) and meridional (gray) decorrelation lengths as a percentage of the total number of cells for which spatial ACFs could be calculated. The arrows on the x-axis indicate the median decorrelation lengths for (gray) meridional direction, (striped) zonal direction and (black) all directions combined.

Basin	**All directions**		**Zonal**		**Meridional**	
Western North Pacific	**47%**	(0.64)	**79%**	(0.83)	**37%**	(0.52)
Eastern North Pacific	**64%**	(0.81)	**58%**	(0.76)	**49%**	(0.40)
Equatorial Pacific	**44%**	(0.60)	**34%**	(0.48)	**38%**	(0.34)
South Pacific	**65%**	(0.65)	**49%**	(0.52)	**40%**	(0.35)
North Atlantic	**73%**	(0.67)	**71%**	(0.61)	**72%**	(0.58)
South Atlantic	**42%**	(0.34)	**19%**	(0.09)	**20%**	(0.13)
Indian Ocean	**57%**	(0.39)	**35%**	(0.24)	**23%**	(0.19)
Global	**59%**	(0.75)	**46%**	(0.68)	**37%**	(0.55)

Table 2.1: Uncertainty levels for the autocorrelation lengths of pCO_2 measurements in different ocean regions. Uncertainties are calculated as the linear relationship between the autocorrelation length for each grid cell and the standard deviation of cruise autocorrelation lengths contributing to that cell. This gives the uncertainty as a percentage of the calculated autocorrelation length. Numbers in brackets show the r^2 coefficient of the linear fit to illustrate the robustness of the uncertainty estimate. The boundary between the eastern and western North Pacific is at 170°E, and the Equatorial Pacific is between 15°S and 15°N.

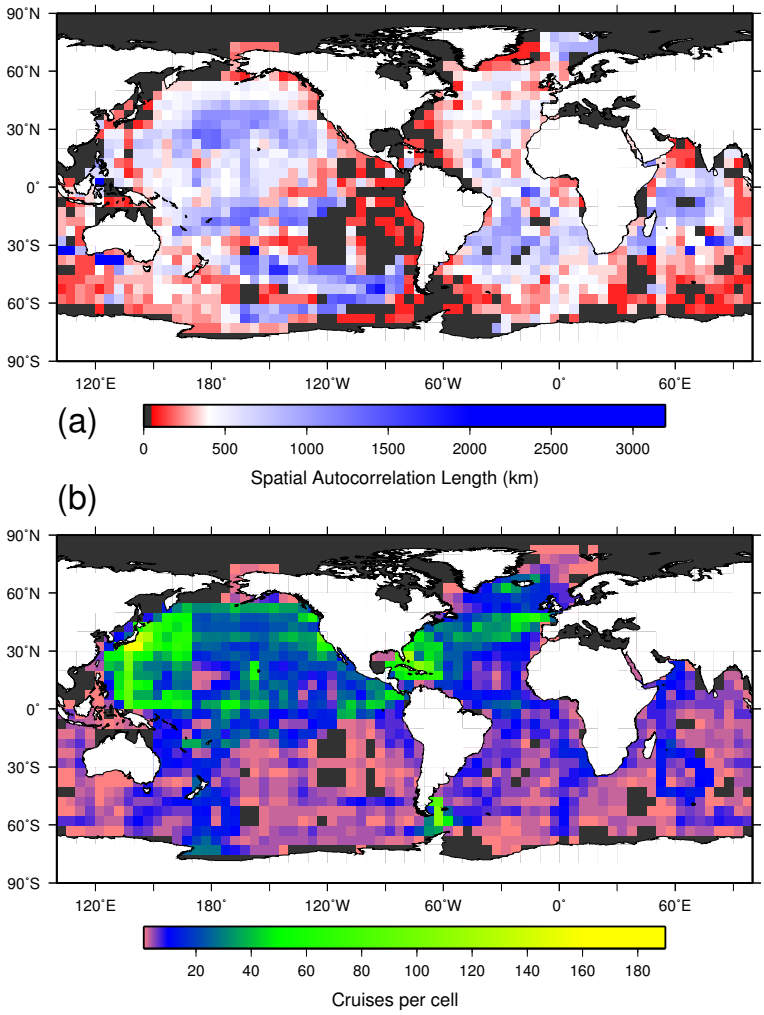

Figure 2.4: (a) The mean spatial autocorrelation length of cruises passing through each 5° grid cell. White values indicate the median autocorrelation length of 400 km, while blue and red cells show longer and shorter correlations respectively. Dark gray cells indicate regions where there is insufficient data to calculate the autocorrelation length, or the autocorrelation length is shorter than the minimum detectable distance. (b) shows the number of cruises passing through each cell.

of all cruises (Figure 2.4a) is 59% of the calculated autocorrelation length, varying between 42% and 73% in different regions. The zonal and meridional uncertainties are 46% (19%-79%) and 37% (20%-72%) respectively. Errors in the zonal and meridional autocorrelation lengths are smaller than those found in the directionless autocorrelations because they eliminate much of the variability caused by different cruises crossing or following currents. Using the same technique, the zonal and meridional errors are calculated as 46% and 37% of the autocorrelation lengths respectively. The region with greatest uncertainty is the North Atlantic, where the uncertainty is greater than 70% in all directions. This is because there are several gyres, currents and upwelling/downwelling areas

[*Schmitz*, 1996] in this relatively small region, including the Gulf Stream whose position varies over time [*Kelly*, 1991]. This means that cruises passing through this region will encounter several different water masses with different spatial variability, which may be in different locations for different cruises in the LDEO database. This accounts for the large uncertainties in spatial autocorrelation length in the North Atlantic. The varying position of the Kuroshio current and its extension [*Kawabe*, 1995] has a similar effect in the western North Pacific, which shows much higher zonal variability than the eastern North Pacific.

The map of mean autocorrelation lengths highlights many of the major ocean currents and gyres as regions where autocorrelation lengths are long (1,000 km and above), especially away from the coasts (Figure 2.4a, page 35). The North Pacific and South Atlantic gyres are clearly discernible, as are the currents of the Indian Ocean. Short autocorrelation lengths (400 km and below) are evident where waters are heterogeneous or where different water masses are in close proximity. This is most evident in the Southern Ocean, where the water characteristics are heterogeneous [*Watson and Naveira Garabato*, 2005], especially around Drake Passage and the Scotia Sea [*Heywood et al.*, 2002]. Other prominent regions of short autocorrelation lengths include the Humboldt current system off Chile, Peru and into the equatorial Pacific, where biological activity is particularly pronounced [*Morales and Lange*, 2004]; the North Atlantic around Iceland and Greenland, where the Gulf Stream is most prominent [*Dickson and Brown*, 1994]; the Kuroshio current in the western Pacific south of Japan [*Taft et al.*, 1973]; the highly variable currents of the Caribbean Sea and Gulf of Mexico [*Richardson*, 2005]; and the continental shelf of the South Atlantic Bight of the United States [*Jiang et al.*, 2008]. The North Atlantic is the only ocean basin with no obvious coherence in spatial autocorrelation lengths. This is due to the high variability of the currents in this region combined with the effects of biological activity. The distribution of autocorrelation lengths in the North Atlantic becomes much clearer when the zonal and meridional cruises are assessed separately (Figure 2.5, page 37).

The accompanying map of cruise counts for each cell shows areas of most prolific coverage (and therefore greatest confidence) in the western North Pacific off the Japanese coast, the Caribbean islands and Drake Passage in the Southern Ocean, with over 50

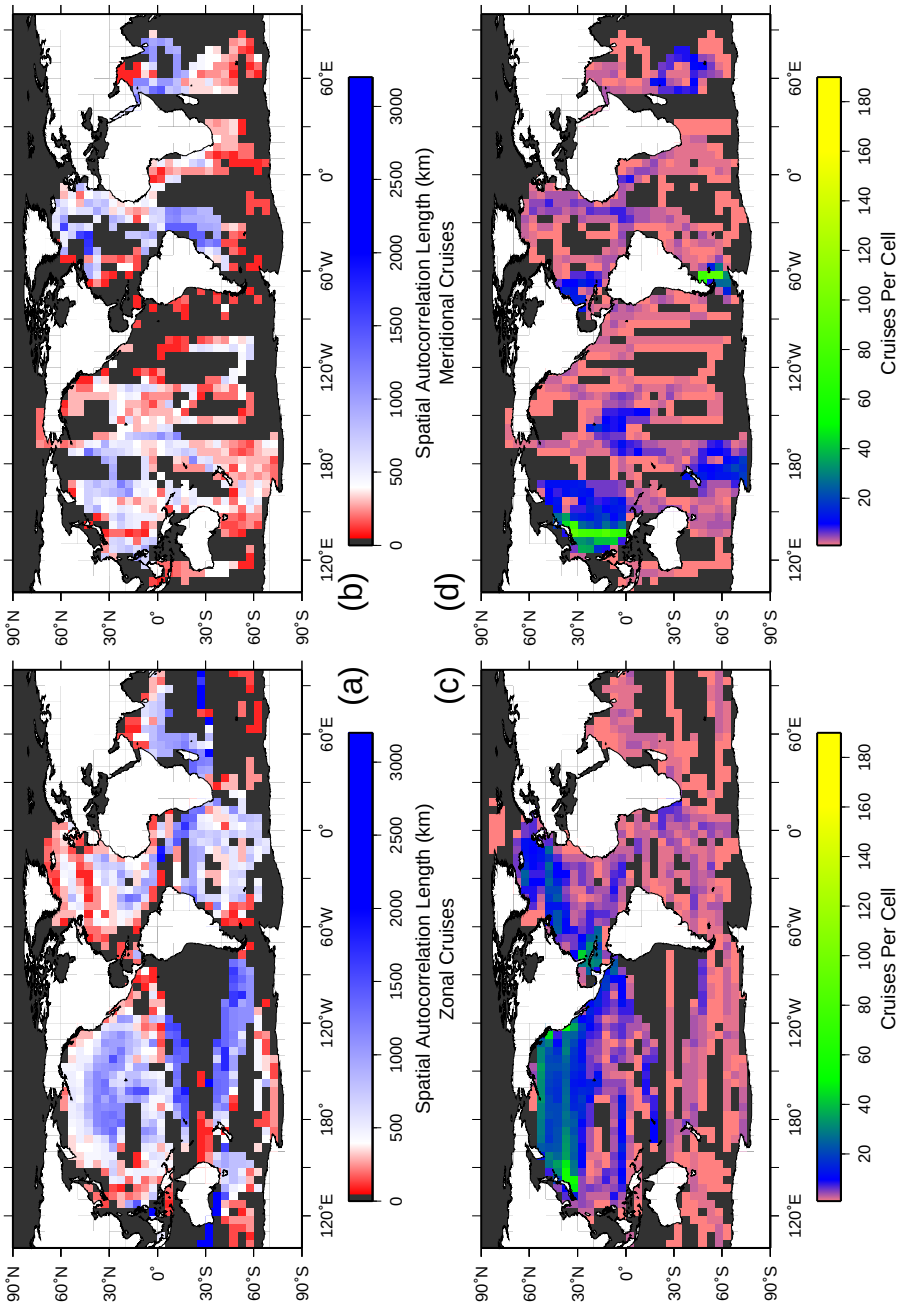

Figure 2.5: Spatial autocorrelation maps for (a) zonal and (b) meridional cruises only, with accompanying cruise counts (c) and (d) as for Figure 2.4 (Page 35).

cruises recorded in the database (Figure 2.4b, page 35). The North Atlantic and North Pacific have 10 or more cruises recorded over the majority of their areas. The remainder of the world's oceans are only minimally sampled outside repeat cruise tracks such as those between New Zealand and the Antarctic, and a repeated circular cruise track in the Indian Ocean.

Further detail of spatial autocorrelation patterns can be seen by examining the zonal and meridional cruises independently (Figure 2.5, page 37). The extended autocorrelation lengths in the North Pacific basin (1,200 $\pm$ 700 km), the South Equatorial current (1,500 $\pm$ 500 km) and the Antarctic Circumpolar current (1,300 $\pm$ 500 km) are more clearly discernible in the zonal map along the main direction of water flow, with much shorter meridional autocorrelations of 550 $\pm$ 200 km, 450 $\pm$ 150 km and 450 $\pm$ 150 km respectively. Meridional correlations dominate in the Atlantic Ocean only, particularly in the mid to high latitudes of the North Atlantic (1,400 $\pm$ 1,000 km) and the western South Atlantic (1100 $\pm$ 200 km). In the western tropical Atlantic the autocorrelations follow the bifurcation of the South Equatorial current on the coast of South America, forming the Brazil and North Brazil currents [*da Silveira et al.*, 1994]. In the eastern North Atlantic, the autocorrelations are associated with the Canary Current [*Schmitz*, 1996]. The long meridional autocorrelations in the western North Atlantic follow the Gulf Stream and the North Atlantic Current [*Flatau et al.*, 2003], showing the greatest dominance of meridional over zonal correlations. Cruises travelling east-west here will cross many currents carrying waters of different characteristics, thereby producing short autocorrelation lengths; north-south cruises, meanwhile, will not see this effect. These long autocorrelations extend as far north as Greenland and Iceland, where the North Atlantic Current loses its identity around Greenland and there is a large area of dense, sinking water at the limits of the thermohaline circulation [*Dickson and Brown*, 1994].

A full analysis of zonal and meridional autocorrelations cannot be performed for the eastern South Pacific, the region of the Southern Ocean south of South Africa, or for much of the South Atlantic due to the uni-directional nature of the cruises in this region (Figure 2.5c and d, page 37). Comparing the autocorrelation lengths of the temperature-driven and residual components (Figure 2.6a, page 39) shows that the temperature-driven component is more spatially stable in much of the ocean, with 61% (17%) of grid cells

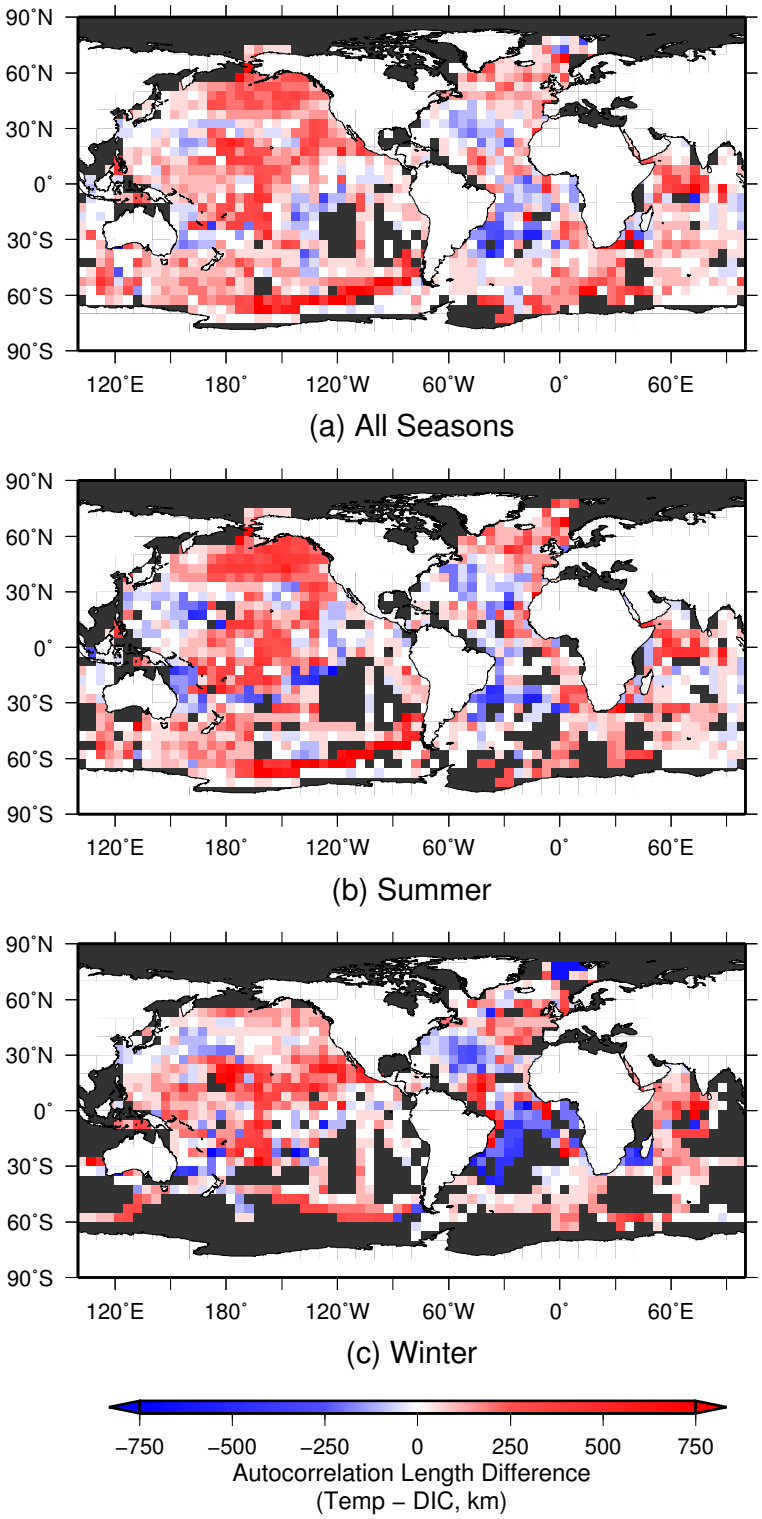

Figure 2.6: The difference between spatial autocorrelation lengths of the temperature and residual components of the pCO_2 measurements for (a) all seasons, (b) summer and (c) winter. Red (blue) regions indicate that the temperature (residual) component is more spatially stable.

reporting longer correlations for the temperature (residual) component. The residual component tends to have the longest relative autocorrelation length in the mid-latitudes of the Atlantic, with similar but weaker patterns in the Pacific mid-latitudes. This pattern compares well with analyses of the biological influence on surface pCO_2 levels [*Takahashi et al.*, 2002], indicating that this is a significant constituent of the residual component. The relative spatial stability of these two components varies with the seasonal cycle. In the summer months (June-July-August/December-January-February in the northern/southern hemisphere), the pattern of relative spatial stability (Figure 2.6b, page 39) is much the same as that for the complete year, while pattern in the winter months changes significantly (Figure 2.6c, page 39). Analysis of the seasonal differences in the two compnents (not shown) shows that this is due to a combination of the temperature component becoming less spatially stable in the winter months, and biological activity becoming more spatially stable as it decreases to a minimum in most regions.

2.5.3 Comparison with drivers

Comparing the maps of pCO_2 autocorrelation lengths with those of chlorophyll, SST and SSH (Figure 2.7, page 41) shows the extent to which the latter variables may act as drivers for the pCO_2 autocorrelations. The chlorophyll and SST maps show the same basic large-scale patterns of spatial autocorrelation, with larger autocorrelations in the central basins of the Atlantic and Pacific. Values in the eastern Indian Ocean are not well defined, since they are consistently shorter than the 200 km lower limit on detectable autocorrelation lengths for the gridded data sets and therefore show no variability. SSH autocorrelation lengths are also below the 200 km threshold across much of the global ocean, with only the Equatorial and North Pacific, tropical Atlantic and portions of the Southern Ocean exhibiting longer autocorrealtion lengths.

The pattern of the chlorophyll and SST maps is visible to some extent in the map of pCO_2 autocorrelation lengths, although it is obvious that these are not leading drivers of the autocorrelation length since the pCO_2 map shows greater spatial variability. This is confirmed with a quantitative comaprison of the maps, with pattern correlations of r^2=0.24 and r^2=0.21 for chlorophyll and SST respectively. The SSH map cannot be reliably compared to the pCO_2 map because of the limited number of regions in which the

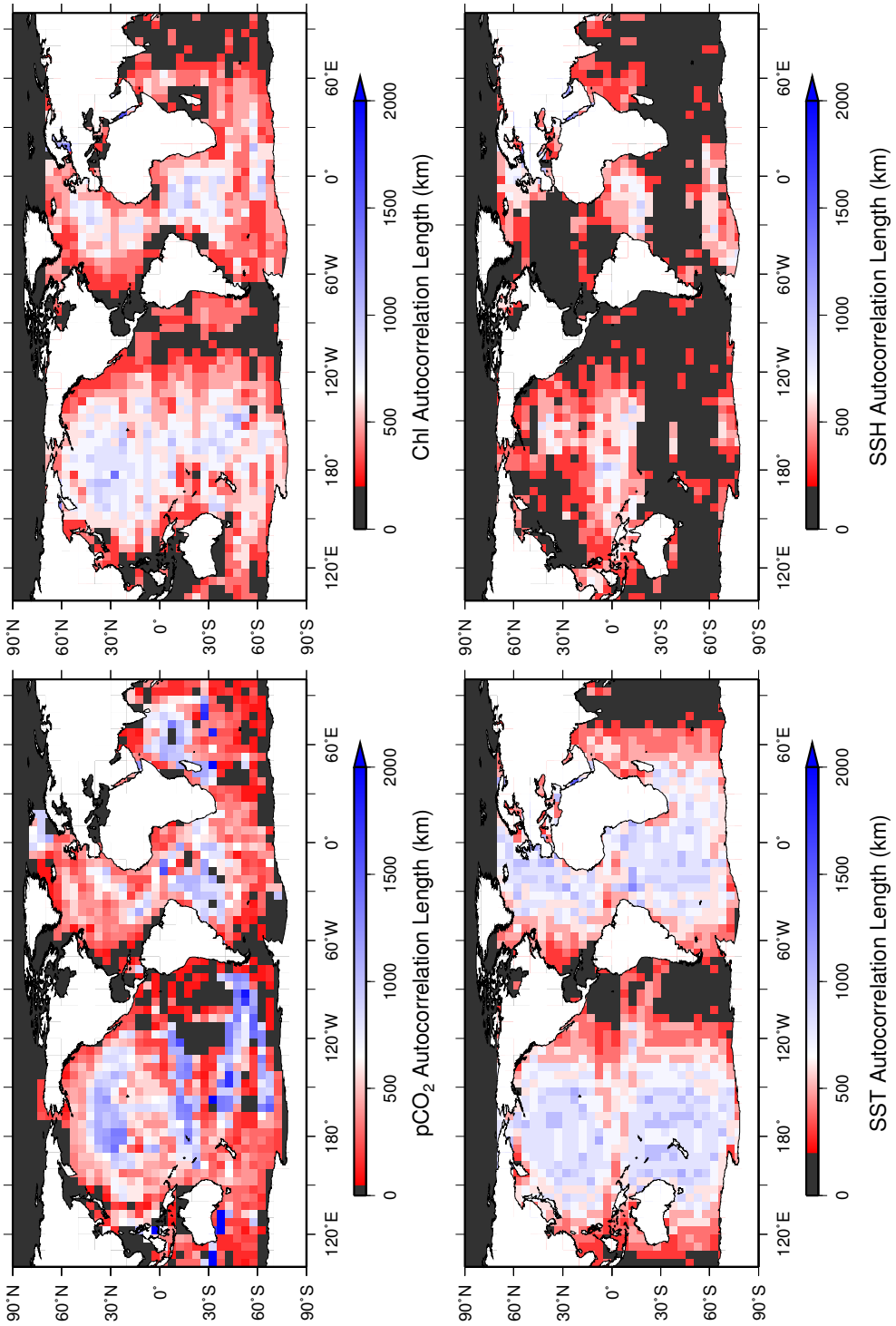

Figure 2.7: Maps of spatial autocorrelation lengths of (clockwise from top left) pCO_2 (from Figure 2.4, page 35), chlorophyll, sea surface height and sea surface temperature. Dark gray cells indicate regions where there is insufficient data to calculate the autocorrelation length, or the autocorrelation length is shorter than the minimum detectable distance.

autocorrelation length can be estimated. However, the relatively low similarity of pCO_2 autocorrelation lengths, and the fact that ocean currents and gyres are clearly visible in the zonal and meridional maps of autocorrelation (Figure 2.5, page 37), leads to the conclusion that it is the physical circulation of the oceans is likely to be the largest influence on the patterns of pCO_2 autocorrelation.

2.5.4 Flux autocorrelation

Spatial autocorrelation lengths for CO_2 fluxes are approximately half those calculated from the pCO_2 measurements (200 (150-350) km). Estimated uncertainties for the flux autocorrelation lengths are very similar to those for the pCO_2 measurements (Table 2.1, page 34). Mapping the individual components of the flux calculation (Figure 2.8, page 43) reveals the primary cause of this difference. The ocean pCO_2 and ΔpCO_2 autocorrelation lengths are essentially identical, with a mean difference that is smaller than the 50 km resolution of this analysis; atmospheric CO_2 therefore has no influence on the flux autocorrelation. Solubility autocorrelation lengths are typically longer than those of the pCO_2 measurements (600 (350-950) km), but this has the parameter with by far the smallest influence over the calculated flux value, consistent with current understanding of the carbonate system [*Takahashi et al.*, 2009]. The difference between the pCO_2 measurements and the gas transfer velocity is 150 (50-350) km, which is very close to the overall difference between pCO_2 and the total flux (150 (50-300) km). Pattern correlation tests show that the fluxes have a very similar distribution to both pCO_2 and the gas transfer velocity, with $r^2 = 0.71$ and $r^2 = 0.76$ respectively. Thus we conclude that the gas transfer velocity is most influential in causing the decreased autocorrelation length in CO_2 fluxes.

2.5.5 Validation

2.5.5.1 Bias detection

Tests for the existence of systematic biases in the data show that there are no inherent characteristics of the LDEO data set that influence the results of this study. Checks were performed to ensure that the spatial autocorrelation length for a given grid cell is not influenced by the number of cruises contributing to that value, despite observations that the regions with most cruises tend to be regions of short spatial autocorrealtion length. A

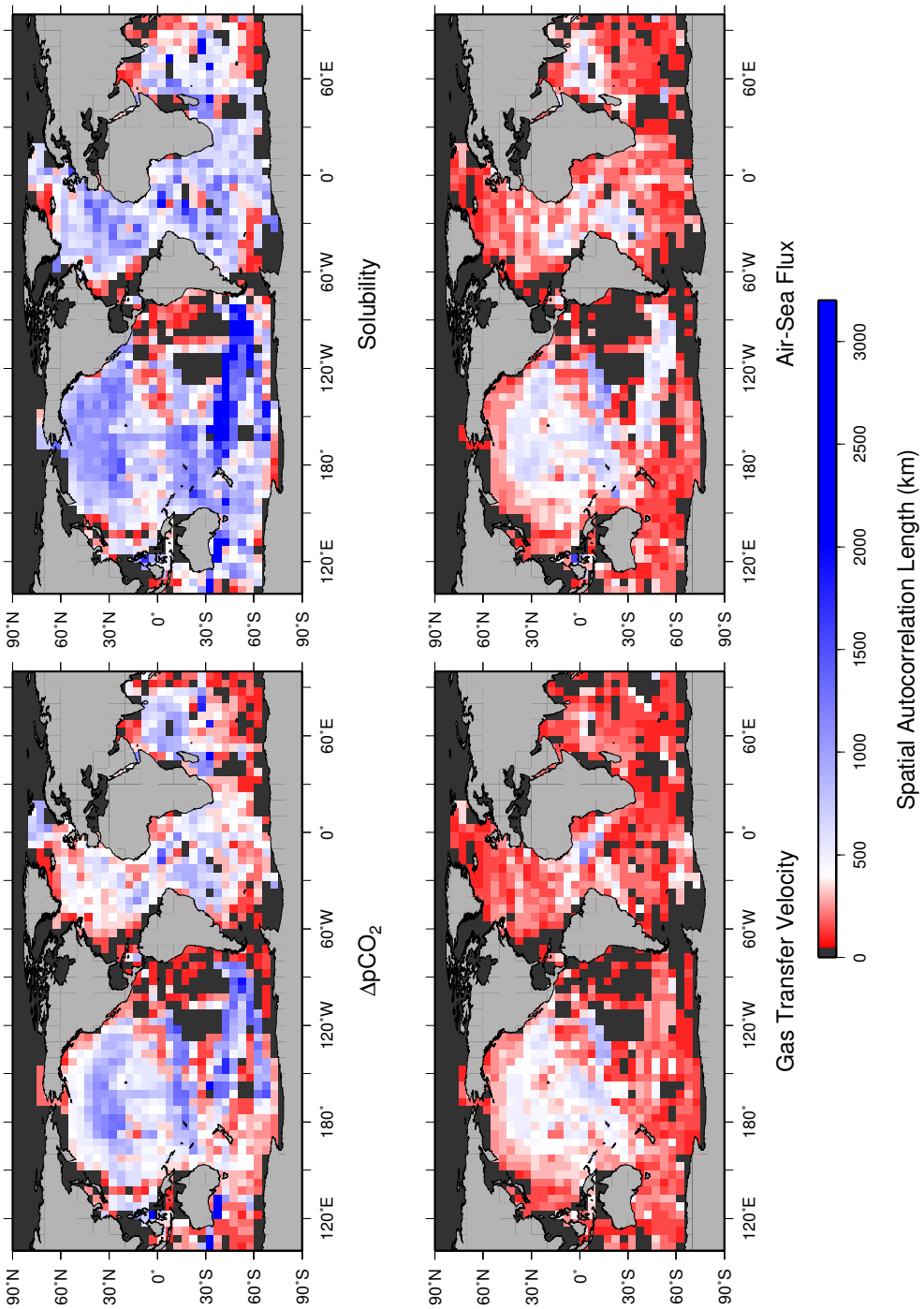

Figure 2.8: Autocorrelation lengths of components of the air-sea flux of CO_2. Top left: The difference between atmospheric and surface ocean pCO_2; Top right: solubility; Bottom left: gas transfer velocity; Bottom right: The calculated air-sea flux.

linear regression fit on the relationship between autocorrelation length and the number of cruises in each cell gives an r^2 of 0.056, confirming that there is no such relationship. Furthermore, examination of the sea surface height (SSH) (calculated using the AVISO SSH anomaly data) also shows that the regions of high cruise counts and short autocorrelation lengths are regions of high SSH variability. This indicates the high mesoscale variability caused by unstable currents where short autocorrelation lengths are expected.

2.5.5.2 Comparison with previous studies

The results of our autocorrelation analysis compare well with previous studies of pCO_2 variability, but provides near-global coverage and a level of detail that better highlights oceanographic features and allows the identification of underlying drivers. The strong seasonal cyle in the temporal ACF is in agreement with similar regional studies, both in terms of interannual variability [*Bates et al.*, 1996; *Gruber et al.*, 2002; *Wong et al.*, 2010] and the ability to fit harmonic curves to time series of pCO_2 measurements [*Schuster et al.*, 2009]. The spatial autocorrelation analysis also compares well with other studies examining both surface ocean pCO_2 and related air-sea fluxes. The gyre and current features visible in Figures 2.4 (Page 35) and 2.5 (Page 37) are similar those described by *Li et al.* [2005], but they provide a more coherent picture and details that were not captured therein because of the scale limitation and the coarse grid selected. The additional resolution and less restrictive limits used here enhance significantly the ability to detect and understand these characteristics. The short autocorrelation lengths in the Humboldt current region agree well with high spatial pCO_2 variability associated with strong CO_2 drawdowns [*Lefèvre and Taylor*, 2002]. Relatively short autocorrelation lengths also agree with high spatial variability of carbon fluxes found in the south-east Atlantic [*Santana-Casiano et al.*, 2009] and the South Atlantic Bight [*Jiang et al.*, 2008], while the "moderate" variability in the western equatorial Pacific [*Ishii et al.*, 2009] is reflected in autocorrelation lengths close to the global mean average. The autocorrelation lengths found in this study also match closely estimates of the required spatial sampling rate for pCO_2 along specific cruise tracks from previous versions of the Takahashi database [*Sweeney et al.*, 2002]. The directional autocorrelation lengths we find in the regions matching the same cruises are very close to the results from that study, which is to be expected since both studies

are based upon the analysis of individual cruises. However, our analysis shows that all available data should be examined to provide a true picture of spatial variability of pCO_2 across the oceans.

2.6 Summary and conclusion

The temporal and spatial autocorrelation analysis of the LDEO database of surface ocean pCO_2 measurements and their corresponding air-sea fluxes provides a comprehensive insight into the global variability of these critical ocean characteristics. For pCO_2 in the temporal dimension, the monthly mean ACF exhibits a robust and consistent seasonal cycle. For pCO_2 in the spatial dimension, the global median and quantile autocorrelation lengths of pCO_2 are 400 (200-650) km. For the air-sea CO_2 flux, the global median autocorrelation length decreases to 200 (150-350) km because of the spatial variability of the gas transfer velocity. In both cases zonal correlations are longer than their meridional counterparts, indicating that ocean currents play a significant role in determining these lengths. The major ocean currents and gyres have longer correlations in both pCO_2 and CO_2 fluxes than those regions with less heterogeneous characteristics, consistent with the autocorrelation lengths in sea surface height.

The results of this study will be useful to both the measurement and modelling communities. They will inform a future research into the interaction between the atmospheric and oceanic carbon cycles, and help to develop future oceanic measurement strategies. The results are particularly relevant for atmospheric CO_2 inversions, which require a-priori correlations in Bayesian inverse calculations to estimate CO_2 fluxes from atmospheric data. Our analysis suggests that inverse calculations should incorporate a-priori correlation of pCO_2 patterns and compute CO_2 fluxes using observed winds to optimize the information content of the available surface ocean data. Such a strategy would require the addition of a surface ocean box in inversions in order to merge the oceanic and atmospheric data streams most effectively.

Chapter 3

A Global Interpolation of Surface Ocean pCO_2

3.1 Abstract

We have developed an interpolation technique adapted to the specific coverage and properties of surface ocean pCO_2 observations, and used it to produce a complete daily data set of pCO_2 values from 1990-2008 on a 5°x5° global grid south of 70°N based on the database of the Lamont-Doherty Earth Observatory. The interpolation combines spatial techniques based on a 'radius of influence' to detect nearby related observations and harmonic curve-fitting to determine trends and seasonal cycles. Tests of the technique using model data show the technique to perform as well as or better than previous regional interpolations: globally, 90% of interpolated values are likely to be within 30 μatm of the actual pCO_2 value. All values are assigned an uncertainty based on the spatial and temporal range over which they are interpolated, which will guide future users in their use of the data set. Unlike previous interpolations, our approach does not rely on relationships to external variables, and can therefore be applied to other data sets of similar data density.

3.2 Introduction

The very uneven distribution of surface ocean pCO_2 measurements is a fundamental limiting factor on our ability to understand many features of the oceanic carbon cycle and to calculate global air-sea CO_2 fluxes. The majority of measurements are concentrated in the northern hemisphere (Figure 3.1, page 50), with most obtained within the last 20 years. This relative lack of spatial and temporal data coverage is in stark contrast to many other oceanic observations. For example, sea surface temperature, sea surface height and chlorophyll concentrations can all be measured via remote sensing satellites, providing almost global coverage at high resolutions. Other surface ocean variables such as dissolved oxygen concentrations are easier to measure and therefore have much more complete records than pCO_2, as sensors can be deployed on floats and gliders [*Johnson et al.*, 2009].

The limited distribution of surface ocean pCO_2 has meant that much of the past work to map these values has concentrated on the production of climatological data sets of the seasonal cycle [*Takahashi et al.*, 2002] and on long-term trends [e.g. *Takahashi et al.*, 2003, 2009], with little or no emphasis on variability at other temporal scales. Until

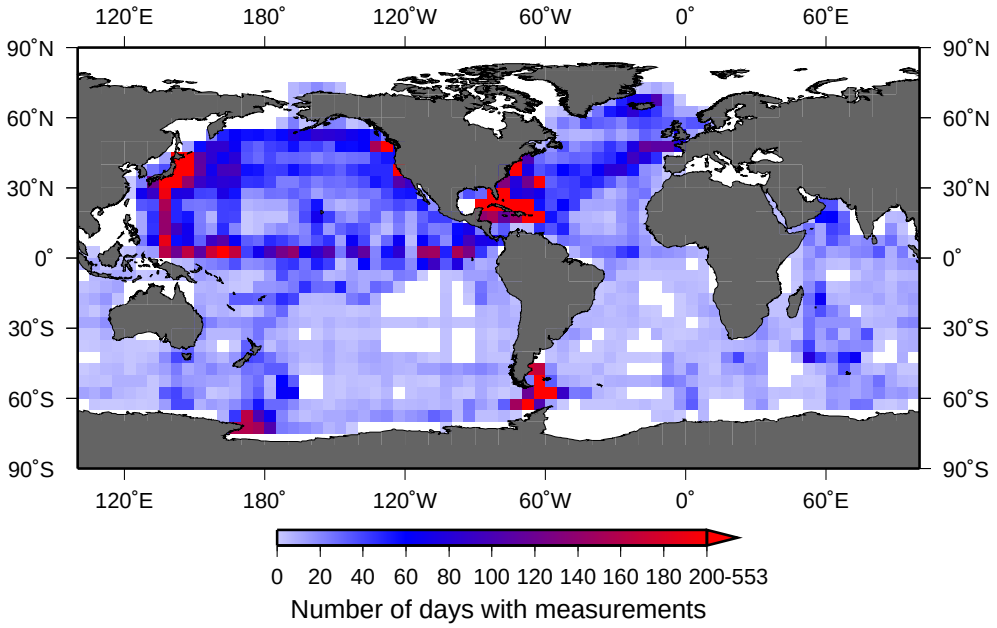

Figure 3.1: Data density of the LDEO database sampled onto a 5°x5° grid. The scale indicates the number of days between 1990-2008 for which pCO_2 measurements are available in each grid cell.

recently, only regional studies have focused on pCO_2 variability on sub-annual time scales [e.g. *Bates et al.*, 1998; *Sarma*, 2003; *Shim et al.*, 2007; *Olsen et al.*, 2008; *Litt et al.*, 2010] and on interannual variability [*Bates et al.*, 1996; *Gruber et al.*, 2002; *Cosca et al.*, 2003; *Wong et al.*, 2010]. Whilst these studies go some way towards aiding our understanding of oceanic fluxes of CO_2, a global perspective of pCO_2 characteristics remains lacking.

The recent release of a complete global database of surface ocean pCO_2 measurements by the Lamont-Doherty Earth Observatory [*Takahashi and Sutherland*, 2009] (hereafter referred to as the LDEO database) provides opportunities for a more detailed global analysis of surface ocean pCO_2 over multiple time scales. An interpolated data set of surface ocean pCO_2 measurements covering multiple years would be invaluable as a starting point for analysis of pCO_2 trends and variability and their drivers on a variety of temporal and spatial scales. It could provide insight into the response of oceanic pCO_2 to climate change and variability, provide the necessary prior estimates for atmospheric inverse methods [e.g. *Gurney et al.*, 2002], and help the validation of ocean biogeochemical model results [e.g. *Le Quéré et al.*, 2009]. Some work has already begun in assimilating the observations from the LDEO database into biogeochemical models [*Valsala and Maksyutov*, 2010].

Large-scale interpolations of other physical and biogeochemical variables have been produced in the past: the World Ocean Atlas [*Levitus*, 1982] pioneered this field with the adaptation of a spatial interpolation method developed for atmospheric pressure fields [*Cressman*, 1959; *Barnes*, 1964] using data coherence within a 'radius of influence' (the distance over which a given value is correlated with nearby values). This method was used to produce climatologies of several ocean variables. *Masarie and Tans* [1995] developed a temporal interpolation method to establish long term trends and seasonal cycles in atmospheric CO_2 levels, fitting harmonic curve equations to measurement time series [*GLOBALVIEW-CO2*, 2008]. More complex interpolation techniques have allowed interpolation in both space and time, for example by representing sea surface temperature and sea level pressure as empirical orthogonal functions [*Kaplan et al.*, 1997; *Allan and Ansell*, 2006]. Each of these has been successful because there is much greater data coverage for those variables, or because (in the case of GLOBALVIEW) the atmosphere is relatively well mixed compared to the ocean which allows for a simpler interpolation approach.

Some interpolations of surface ocean pCO_2 measurements have been produced in the past. The most widely used pCO_2 climatology [*Takahashi et al.*, 2009] was calculated using a diffusion-advection model that transports pCO_2 values to regions where no data are available. Harmonic time series have been fitted to pCO_2 measurements in individual basins to analyze longer term trends [*Schuster et al.*, 2009], but without the ability to increase spatial coverage. Others have interpolated the available measurements by analyzing the relationships between pCO_2 and related ocean properties, such as sea surface temperature, salinity, chlorophyll and mixed layer depth. These relationships have been calculated using single or multiple regression analysis [*Boutin et al.*, 1999; *Lefèvre and Taylor*, 2002; *Cosca et al.*, 2003; *Ono et al.*, 2004; *Olsen et al.*, 2004; *Park et al.*, 2006; *Jamet et al.*, 2007; *Watson et al.*, 2009; *Park et al.*, 2010], principal components analysis [*Lohrenz and Cai*, 2006] or by training neural networks [*Telszewski et al.*, 2009]. Assessing the relationships between these variables has worked well within a limited temporal and spatial scope, but it has also been shown that the relationships are not constant through space and time [*Boutin et al.*, 1999; *Lefèvre and Taylor*, 2002; *Cosca et al.*, 2003; *Jamet*

et al., 2007; *Park et al.*, 2010], thus it is problematic to apply them to larger scale interpolations. These approaches require a minimum number of available pCO_2 measurements in multiple regions and times against which reliable relationships to other variables can be established.

This paper presents a purely statistical method to interpolate surface ocean pCO_2 in space and time for the entire global ocean south of 70°N. The method uses autocorrelations of pCO_2 measurements within a radius of influence as in the World Ocean Atlas [*Cressman*, 1959; *Barnes*, 1964; *Levitus*, 1982] combined with harmonic curve fitting as in GLOBALVIEW [*Masarie and Tans*, 1995] to interpolate the available pCO_2 measurements. This purely statistical approach avoids the need to establish relationships between pCO_2 and other variables, and can therefore be more reliably extended into regions of lesser data coverage. The method includes an assessment of the uncertainty of the interpolated values. Every value in the final data set is assigned an uncertainty as a measure of the relative likelihood that the computed value is close to the real-world value. This allows subsequent users of the interpolated data set to understand explicitly the limitations of the interpolated data.

3.3 Method

3.3.1 Data preparation

The LDEO database consists of ∼4.1 million individual surface ocean pCO_2 measurements taken between 1968 and 2008. We focused on the 1990-2008 time period, which includes 85.95% of the measurements in the LDEO database. These measurements were binned into 5°x5° grid cells of daily temporal resolution. For leap years, a 'day' was calculated as $1\frac{1}{365}$ calendar days to produce a constant year length of 365 days throughout.

The complete data set was analyzed to remove erroneous entries caused by transcription errors or faulty instrumentation. Measurements falling outside three standard deviations of the daily mean for a given grid cell were discarded as outliers in an iterative process, repeated until no further outliers were detected. 17,952 measurements (0.004%) were eliminated in this manner. Further outliers were removed by examining the complete daily time series for each grid cell as follows. A linear trend for the time series in each

grid cell was calculated and temporarily removed. Any day whose mean pCO_2 level fell outside three standard deviations of the mean was removed from the cell's time series in an iterative process, again repeated until no further outliers were detected. A total of 268 days' measurements (0.007%) were discarded across all grid cells.

The radius of influence over which values could be interpolated was dependent on the autocorrelation characteristics of the pCO_2 data in both temporal and spatial dimensions. A global mean autocorrelation function (ACF) was calculated for the temporal dimension using data from those grid cells that had sufficient measurements. Spatial autocorrelation functions were calculated for each cruise in the LDEO database using the Moran's I technique [*Moran*, 1950], and the e-folding length assigned to each cell through which the cruise passed as a first guess of the autocorrelation length for that cell. This was then refined by calculating the ACF for only those measurements within a radius of five times the first guess value and recording the resulting e-folding length. Where multiple cruises passed through a given cell, the mean e-folding length was used (Chapter 2). The ACFs for a given cell can vary with compass direction, particularly in strong ocean currents. To provide the greatest accuracy, four directional spatial ACFs were calculated for each cell: north-south, east-west, northeast-southwest, and northwest-southeast. A 'direction-less' ACF was calculated for use where insufficient data existed to construct directional ACFs, using all data regardless of directionality. Figure 3.2 (Page 54) shows maps of the e-folding lengths of the ACFs for all cells.

3.3.2 Interpolation

The interpolation of the LDEO pCO_2 data combines the temporal [*Masarie and Tans*, 1995] and spatial [*Cressman*, 1959; *Barnes*, 1964] techniques used in prior interpolations of other physical variables, with minor adjustments to accommodate the specific characteristics of the pCO_2 data set. This allows a full interpolation of the pCO_2 data where the use of a single technique would not be sufficient. No interpolation was attempted poleward of 70°N as there are too few available measurements to produce reliable interpolated values.

The interpolation technique operates in a number of distinct stages which are run iteratively on each 5°x5° grid cell individually (hereafter referred to as the target cell) to

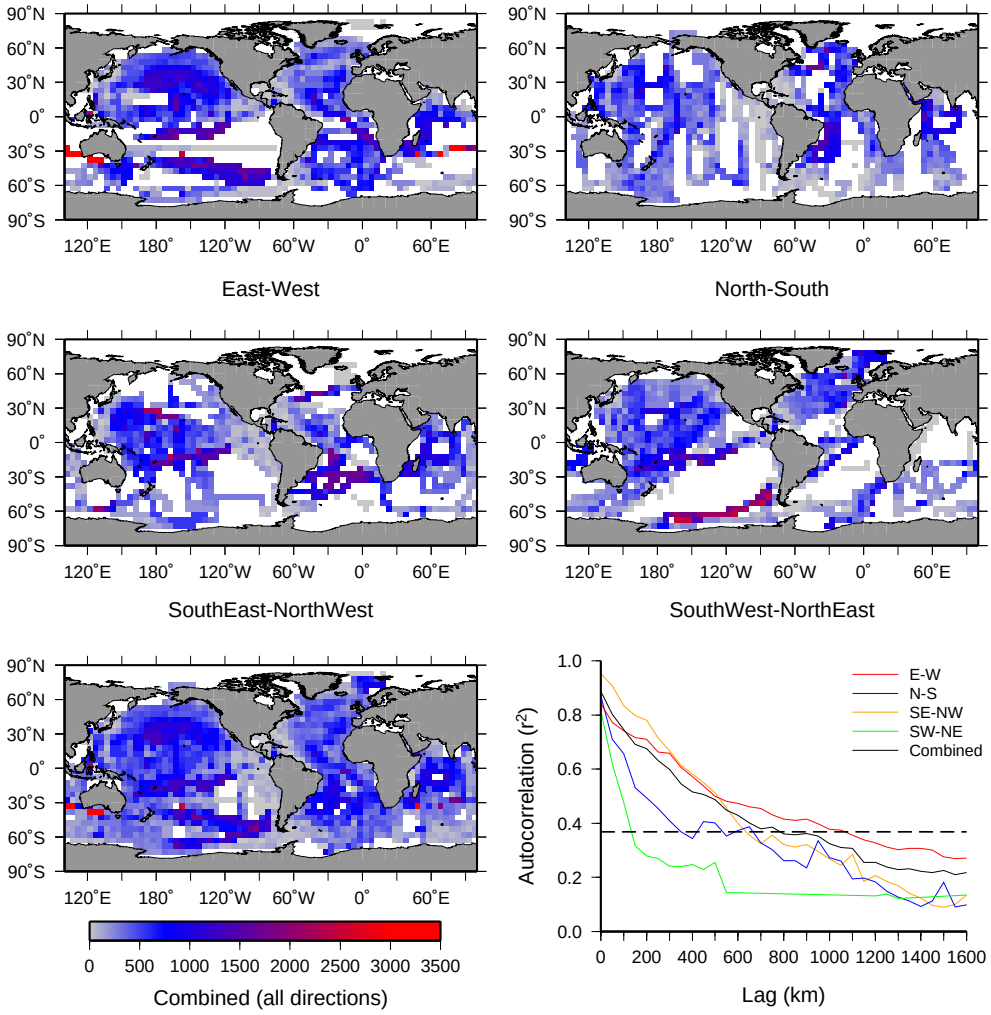

Figure 3.2: Maps of spatial autocorrelations of pCO_2 in various compass directions (top four maps) and the mean of all directions combined (bottom left). The shading of each cell indicates the e-folding length of the ACF for that cell. The graph (bottom right) shows the different directional ACFs for an example cell centered on 32.5°N 147.5°E. The e-folding threshold ($1/e$) is shown as a dashed horizontal line.

build the complete interpolated data set. Here we describe the overall progression of the technique employed, together with details of the individual steps that are taken during the interpolation. Flow charts are provided to give a visual overview of the progression of the interpolation.

3.3.2.1 Progression of the Interpolation

The complete process for creating interpolating a single grid cell is shown in Figure 3.3 (Page 55). This process is run in parallel for all grid cells simultaneously, so values could be shared between multiple grid cells in the spatial interpolation step.

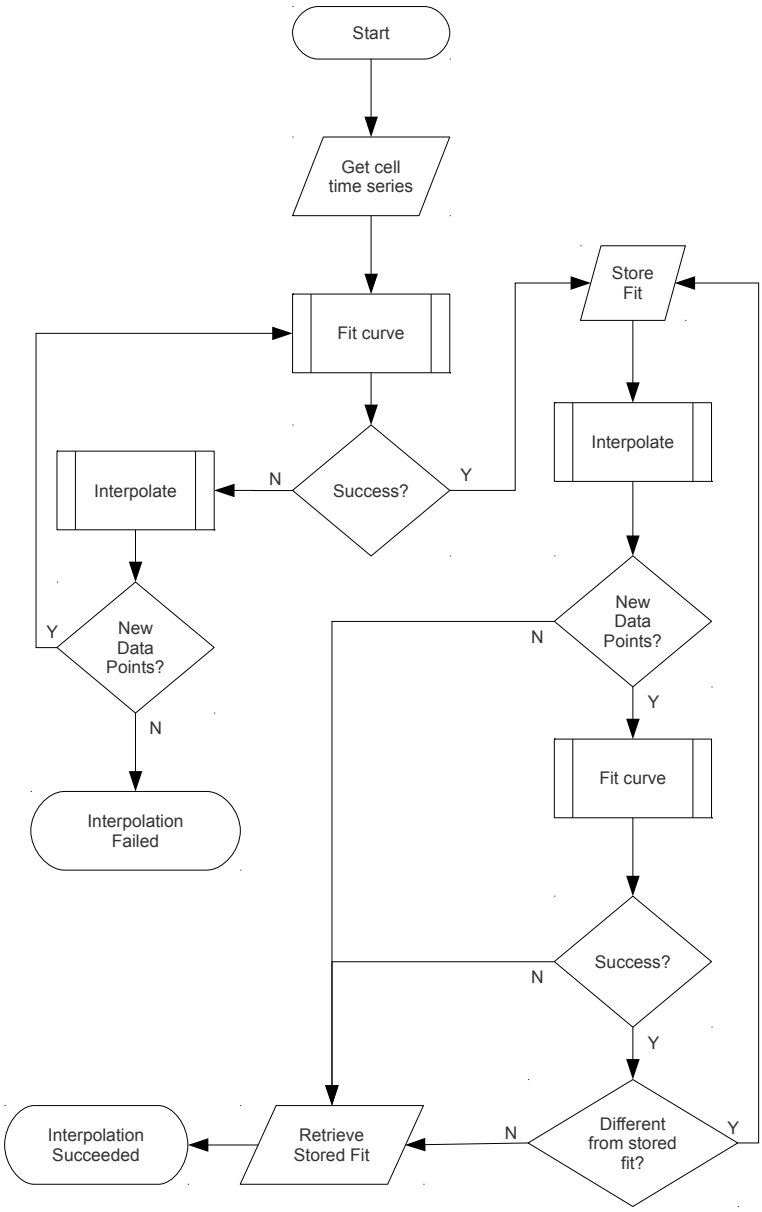

Figure 3.3: The complete process used to interpolate the pCO_2 data.

For each grid cell, a curve fit was attempted (Section 3.3.2.2, page 56). If the curve fit was not successful, the spatial and temporal interpolation were performed (Section 3.3.2.3, page 58), and the next iteration of the process begun. This process was repeated until either a successful curve fit was achieved, or no new data points could be interpolated into the cell's time series. For those regions of the ocean where the spatial autocorrelation is the longest such as the central North Pacific (Figure 3.2, page 54), up to five iterations were possible before no more values could be added, i.e. values can be interpolated from

up to five grid cells away from the target grid cell. In other regions, only one or two interpolations were possible before the spatial ACFs indicated that the neighbouring cells were unlikely to be related to those of the target cell.

Even if a curve fit was successful and passed the criteria for a valid fit, it may be that it was not truly representative of the actual pCO_2 values for that grid cell. When a curve fit was successful the results are stored temporarily, and one more iteration of the interpolation completed. The new curve fit and the stored curve fit were then compared. If the differences between them were not statistically significant at the 95% level, the original (stored) curve fit was used as the final output of the interpolation for the grid cell. If the differences were significant, this indicated that more interpolated values would produce a more realistic curve fit, so the interpolation was repeated until either the differences were no longer significant or no more values could be added via spatial interpolation.

3.3.2.2 Curve Fitting

As described above, the initial step of interpolating data within a grid cell fitted a curve to the time series of that cell. The formula of the curve is a simplified version of that used by *Masarie and Tans* [1995], which included a polynomial term to detect changes in the long-term trend. This was omitted from the present study as the available data was insufficient to detect varying interannual trends in the time period being examined. The curve used here was of the form:

$$f(t) = a_0 + a_1 t + \sum_{k=1}^{n} \left[b_{2k-1} sin\left(2\pi k t\right) + b_{2k} cos\left(2\pi k t\right) \right] \tag{3.1}$$

where t is the time in days since January 1st 1990, a_0 is the y-axis intercept, a_1 is the linear trend, and n is the maximum number of harmonics used to represent the seasonal cycle. n is initially set to 4 to allow the fitted curve to encompass deviations from a purely sinusoidal progression of the seasonal cycle caused by biological activity and temperature changes [e.g. *Lüger et al.*, 2004; *Körtzinger et al.*, 2008].

The fitted curve was assessed against a number of criteria to ensure that it produced a realistic result. The criteria ensured that the curve was based on data covering a reasonable time period with measurements representing a large proportion of the calendar year; that the fitted curve was representative of the range of pCO_2 values and exhibited a plausible

Criterion Name	Description	Justification
Total time range	The timespan covered by the earliest and latest measurements in the time series must be at least five years.	Short timespans of measurements can result in unrealistically steep linear trends.
Standard deviation	The standard deviation of the available measurements must not exceed 75 μatm.	Curve fits applied to time series with only extreme low and high measurements are frequently unrealistic.
Populated months	Measurements must be available in at least 8 of the 12 calendar months at some point in the time series.	Unless at least three of the four annual seasons are represented in the time series, the fitted curve is unlikely to represent a realistic seasonal cycle.
Curve ratio	The amplitude of the fitted curve must be between 50% and 150% of the range of values represented by the measurements. The upper and lower limits of the fitted curve must not exceed the limits of the measurements by more than 75 μatm.	A fitted curve whose amplitude is too small or too large does not represent an accurate fit to the measurements.
Seasonal peaks	Plankton blooms can produce a secondary peak in the sinusoidal seasonal cycle. Only one such additional peak should exist in the fitted curve. The size of the secondary peak must not exceed 33% of the total magnitude of the seasonal cycle.	Fits of multiple harmonics can frequently produce an over-fitted curve with multiple complex peaks in the seasonal cycle. This is unrepresentative of the known annual cycles of pCO_2 concentrations.
Linear trend limit	The fitted linear trend (a_1 in Eq. 3.1) must be in the range -2.5 $\leq$ trend $\leq$ 4.75 μatm yr^{-1}.	Linear trends outside these limits are unlikely to be realistic.

Table 3.1: Criteria used to determine whether or not a curve fitted to a time series of pCO_2 values is plausible.

seasonal cycle; and that the trend of the fitted curve was within known reasonable limits. The complete set of criteria are listed in Table 3.1 (page 57). If the fit failed to meet all criteria, the value of n was reduced by 1 and the curve fit repeated until either a good fit was achieved or no good fit was achieved after n is reduced to 1, at which point the curve fitting was deemed to have failed. A flowchart showing the progression of the curve fit is presented in Figure 3.4 (Page 58).

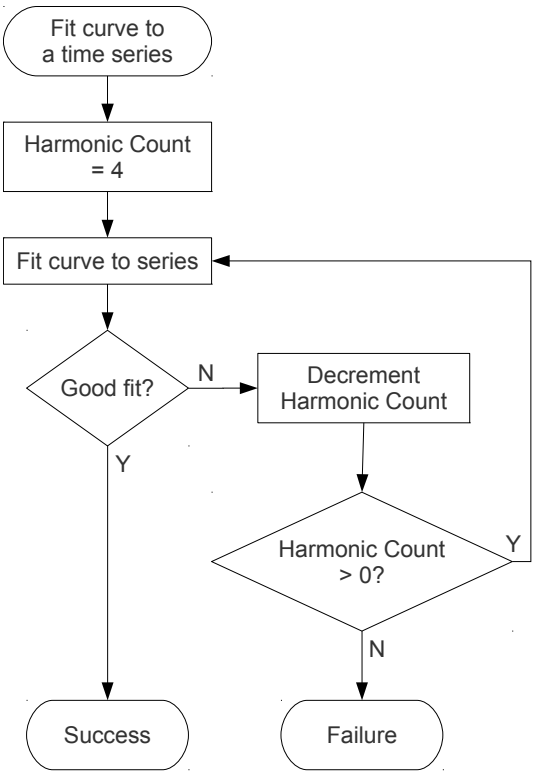

Figure 3.4: The process used to fit a curve to a single grid cell's time series.

3.3.2.3 Spatial and Temporal Interpolation

If the curve fitting for a given target cell failed, spatial and temporal interpolation methods were employed to add values to that cell's time series to increase the chances of obtaining a successful curve fit. Figure 3.5a (Page 59) shows the process of performing the spatial interpolation.

Each step of the cell's time series was processed in turn. If that step already had a value, either from the original pCO_2 measurements or from a previous iteration of the spatial interpolation, no further action was taken. If no value was present in the time series, values were obtained from the same time step in grid cells surrounding the target cell. For the first interpolation iteration, the directly neighbouring cells were used. For subsequent iterations, cells two steps away from the target cell were chosen (see Figure 3.5b). If any values existed for the time step in the surrounding cells, they were combined using a weighted mean and added to the target cell's time series. The weights used were between 0 and 1, and used to give more influence to those values that were more likely to be representative of the target cell. Where possible, the directional ACF was used to match

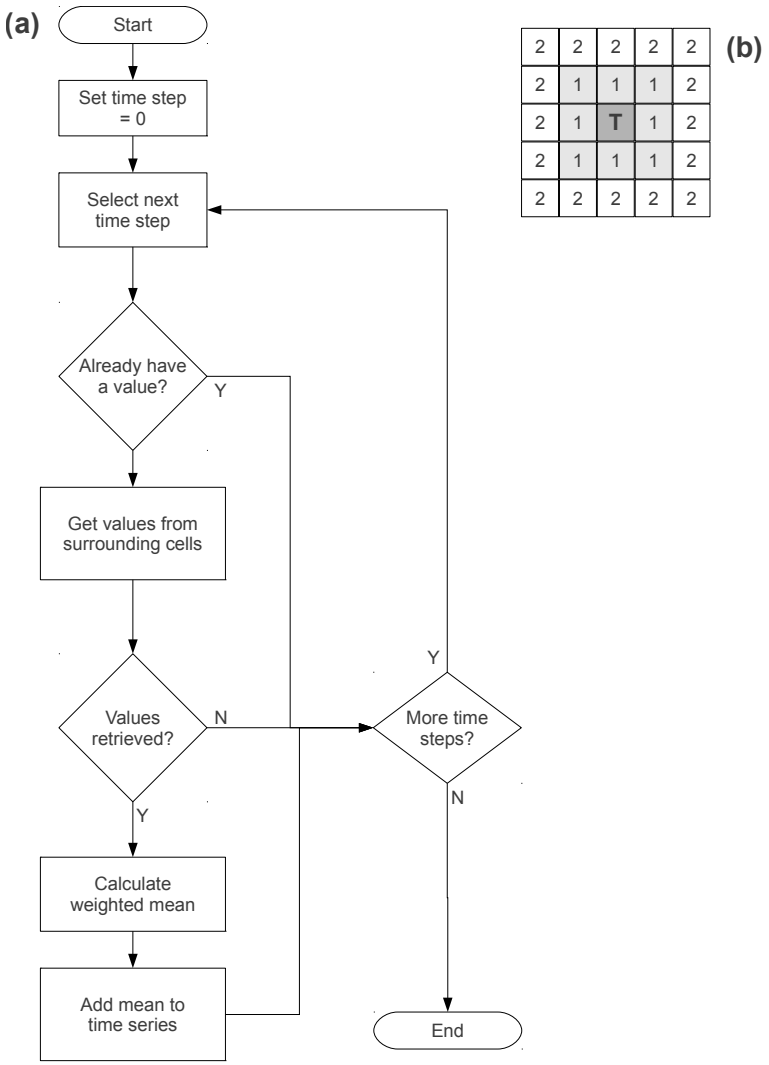

Figure 3.5: The workflow followed in performing the spatial interpolation.

the bearing between the centre of the target cell and interpolated cells; otherwise the directionless ACF was used (Figure 3.2, page 54). If the directionless ACF was also not available, the neighboring cell's directional ACF was used as the best available substitute. The value of the ACF at the distance between the target and interpolation cells was used as a weighting for the interpolated measurement unless it was below the e-folding threshold ($1/e$, or 0.37); in this case the value was not used in the interpolation. The root-mean-squared (RMS) of the weights was assigned as a weighting to the interpolated value, which was then used by subsequent iterations of the curve fitting algorithm to assign greater importance to more representative values in the time series. Since the interpolation was an iterative process, it was possible that a neighbouring cell's time series contained spatially

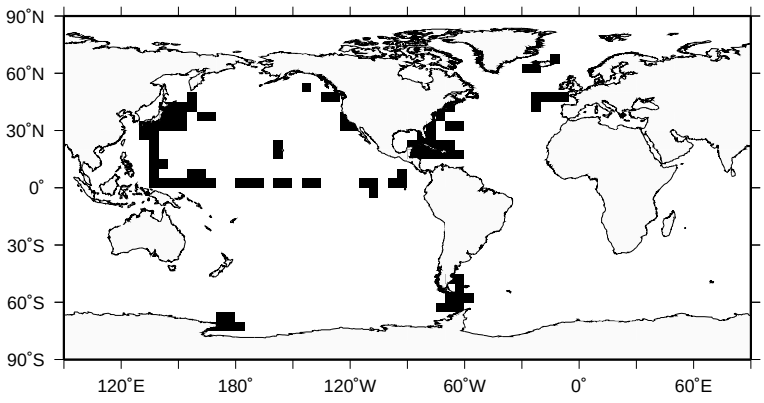

Figure 3.6: The grid cells used in the construction of the daily resolution temporal ACF.

interpolated values from a previous iteration of the process. These values were included in the spatial interpolation, but given much lower weightings due to being interpolated multiple times (the weighting assigned in both interpolation steps were multiplied). Note that no values from fitted curves were interpolated as part of the spatial interpolation.

In some grid cells there were insufficient data points in the time series for the curve fitting algorithm to achieve a fit at all, even after the spatial interpolation. To increase the likelihood of a curve fit being achieved, all values in the cell's time series (both measurements and spatially interpolated values) were copied to the 14 days' time steps before and after the value. This represented the e-folding length of a daily-resolution temporal ACF similar to the monthly temporal ACF developed in Chapter 2 (Figure 2.1b, page 28), but constructed from the few grid cells where daily-resolution measurements were sufficiently dense to produce a statistically significant ACF (Figure 3.6). Although such an ACF is unlikely to be representative of the true daily ACF for surface pCO_2, its use here simply increased the chances of the curve fitting algorithm achieving a fit to the available data. The copied values were assigned weightings from the daily temporal ACF. Once the curve fit was achieved, the temporally interpolated values were discarded to leave only the fitted curve.

3.3.2.4 Completing the Interpolation

The interpolation process described above provided fully interpolated time series for 1,104 of the 1,750 ocean grid cells south of 70°N (63%) after a single iteration. After

five iterations, a further 168 cells were populated. 478 grid cells (27%) remained unprocessed after five iterations of the interpolation algorithm, and the limits to the extent of the spatial interpolation meant that these could not be processed further.

The remaining grid cells were interpolated separately by spatially interpolation the complete time series, including fitted curves that were previously ignored, from the neighbouring successful cells. As with the original spatial interpolation (Section 3.3.2.3, page 58), the interpolated values were weighted according to the target cell's spatial ACFs and coincident values from multiple cells were combined in the target cell using a weighted average.

3.3.3 Calculation of uncertainty

Every value in the interpolated data set is assigned an uncertainty according to the temporal and spatial distance over which the value was obtained. Uncertainties for interpolated values are assigned according to variograms calculated for each grid cell. Variograms are computed for each cruise in the database by creating a scatter plot of the distance between measurement locations and the difference in pCO_2 at those locations (Figure 3.7, page 62). These values are then binned into intervals of 50 km to reduce processing time. The variogram value for each interval is calculated as the mean ΔpCO_2 value plus one standard deviation of all the values within that bin. The fitted variogram is assigned to each grid cell through which the cruise passes in the same manner as the spatial ACFs above, and combined with other variograms for that cell to produce mean variograms in each of eight compass direction. A global average variogram is used in the temporal dimension; variograms require relatively long periods of frequent measurements, and there are few regions where the available measurements are suitably distributed.

Original measurements from the LDEO database are given an uncertainty of ± 2.5 μatm as in the original data set [*Takahashi and Sutherland*, 2009]. Uncertainties for interpolated values from nearby grid cells are assigned as the value of the variogram (Figure 3.7) from that cell at the distance between the value and the target cell; temporally interpolated values have their uncertainty assigned from the global mean temporal variogram (Figure 3.8a, page 63). Uncertainties for values from the fitted curve (Equation 3.1,

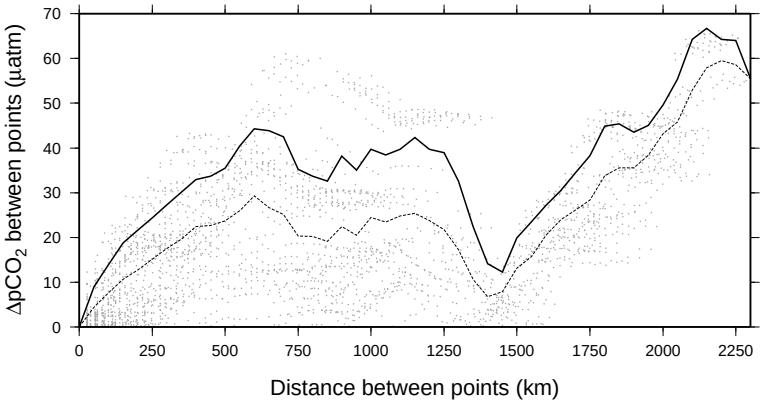

Figure 3.7: Variogram computed for a cruise in the eastern equatorial Atlantic Each dot represents the change in pCO_2 between two points of the cruise (vertical axis), and the distance between those points (horizontal axis). The value for each 50 km bin (thick line) is calculated as the mean value for all points at that distance (thin dashed line) plus one standard deviation.

page 56) are based on the uncertainties assigned to the interpolated values. Each interpolated value is assigned to the day of the year on which it occurs, to produce a set of daily uncertainties for a complete year (Figure 3.8b, page 63, blue line). Where multiple interpolated values within a grid cell fall on the same day, the mean uncertainty is used. If there are days of the year for which no interpolated values (and therefore no uncertainties) are available, the missing values are interpolated between the two known uncertainties either side of the missing values. The interpolation is an average of the two values at each end of the gap in values, weighted according to temporal distance from each point (Figure 3.8b, page 63, red line).This gives a complete progression of uncertainty values for the complete 365-day year.

3.3.4 Building the final data set

Following from the methodology described above, there will be up to three values that can be used for each day in each grid cell where a curve fit has been successfully applied: a measurement from the LDEO database, an interpolated value from a nearby cell and/or time, and a value from the fitted time series (Equation 3.1, page 56). The value chosen for the final data set is the one with the lowest uncertainty assigned during the interpolation (Figure 3.8c, page 63).

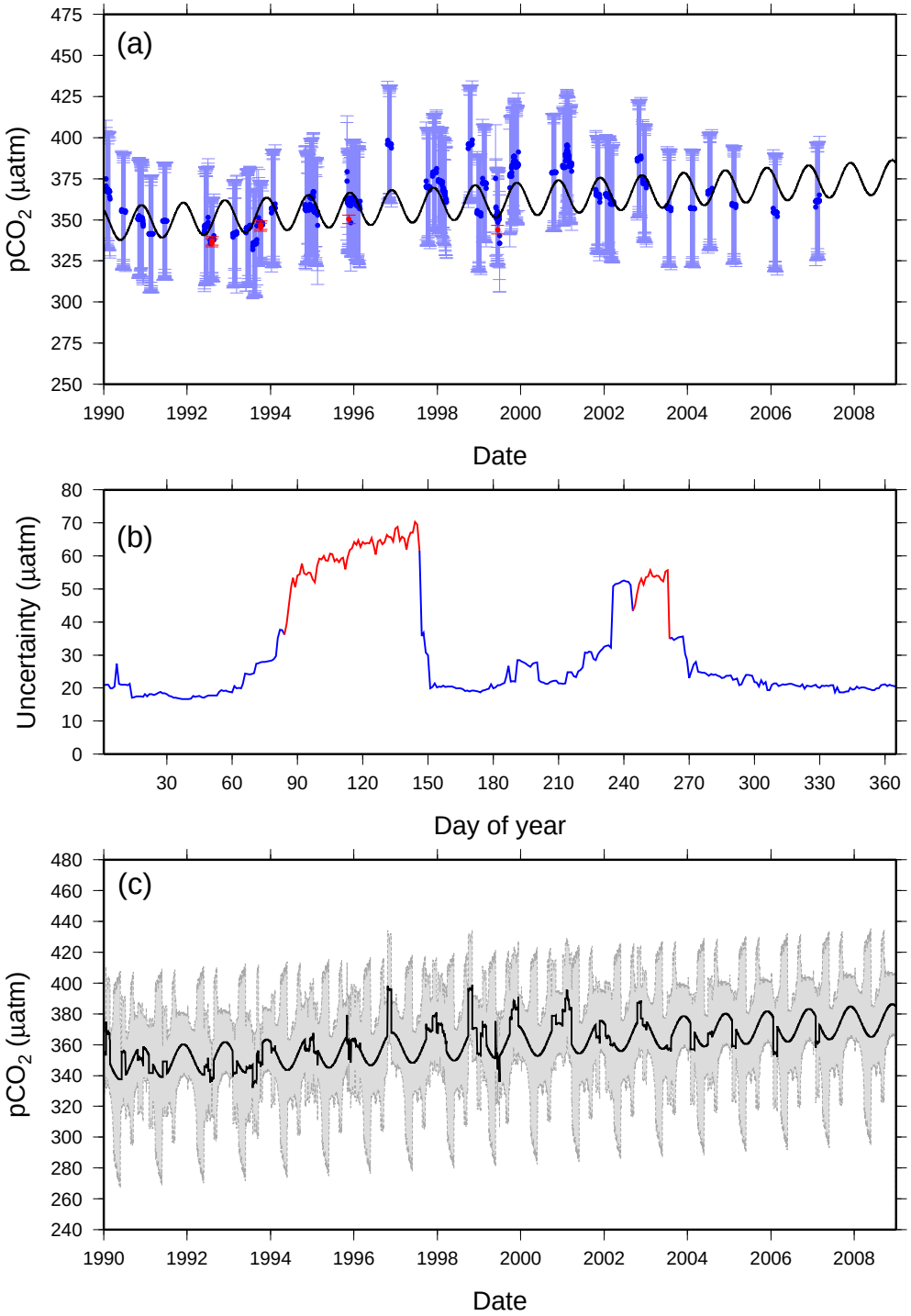

Figure 3.8: Calculation of uncertainty values for the curve fitted to the grid cell centered on 2.5°S 147.5°E. The time series (a) shows (red) the cell's measurements and (blue) interpolated values with uncertainties overlaid on the fitted curve (black). (b) shows the mean uncertainty for each day of the year (blue) and interpolated uncertainties (red) where no other values are available. (c) shows the final interpolated time series (black line) with uncertainties (gray).

3.4 Method validation

Validation of the interpolation method involved testing its ability to recreate a complete data set against model output. Daily pCO_2 values were used from a simulation of the PlankTOM5 model over 1990-2008 (updated from [*Buitenhuis et al.*, 2010]), converted to a 5°x5° grid to match the characteristics of the main interpolation. The regridded PlankTOM5 output was sub-sampled to provide the same spatial and temporal density of input values as the LDEO data set. The resulting PlankTOM5 data set was interpolated using the method described above. The autocorrelation characteristics of the model output were computed in the same manner as for the LDEO database. Spatial ACFs were calculated using measurements from the days and grid cells that matched the LDEO cruises, and assigned to those cells accordingly.

The interpolated PlankTOM5 data set can be compared with the original PlankTOM5 results to assess the performance of our interpolation method. 72% of the values have a root mean squared (RMS) error of $\leq$20 μatm and 89% have an RMS error of $\leq$30 μatm (Figure 3.9, page 65). The cells with the largest RMS errors are concentrated around regions with few or no measurements, namely the eastern equatorial Pacific, the south Atlantic and the Southern Ocean. There is a strong correspondence between the uncertainties assigned during the interpolation and the error determined with the model fit: the values with the highest errors are also those with the largest interpolation uncertainties.

Trends in the interpolated and original PlankTOM5 data show good agreement, with the interpolated trends marginally lower than the original data (1.60 μatm yr^{-1} versus 1.63 μatm yr^{-1} on average). The interpolated trends show greater spatial variability than the original data (Figure 3.10, page 65), with the largest difference in the South Atlantic where the interpolation produces negative pCO_2 trends that are absent from the original data. Negative trends are also seen in the northern Indian Ocean and individual cells elsewhere. These poorly-performing areas have low data density and correspondingly large uncertainties.

The seasonal cycle in the interpolated and original PlankTOM5 data sets are also in good agreement in nearly all regions (Figure 3.11, page 66). The RMS error of the seasonal amplitude is 28.5 μatm, with a pattern correlation of r^2=0.34. Zonal and meridional correlations are r^2=0.59 and r^2=0.51 respectively.

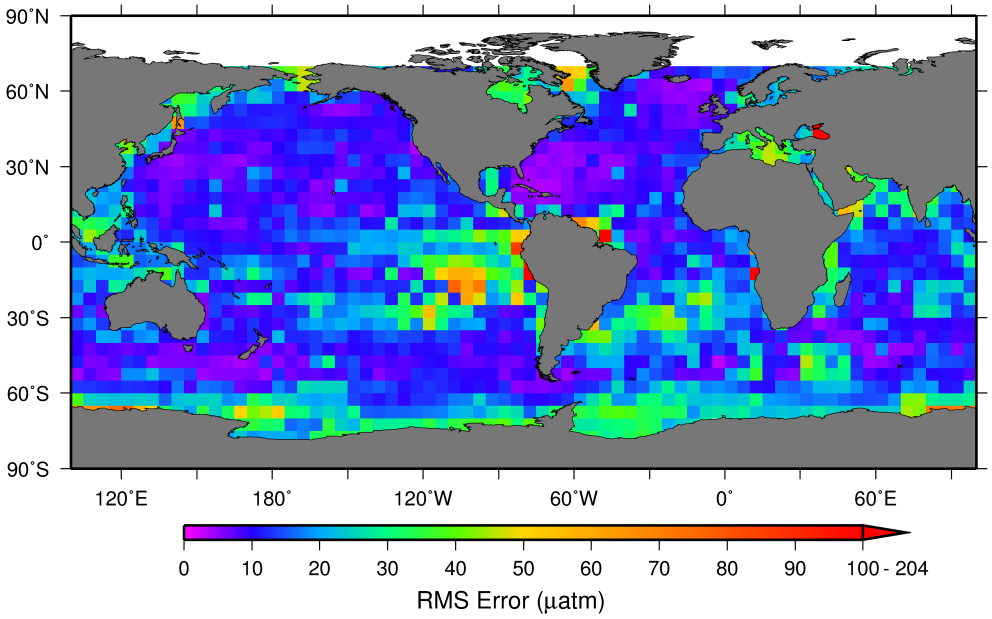

Figure 3.9: The root mean squared error (in μatm) of each grid cell of the interpolated PlankTOM5 data compared to the original model output sub-sampled using the LDEO spatial and temporal density.

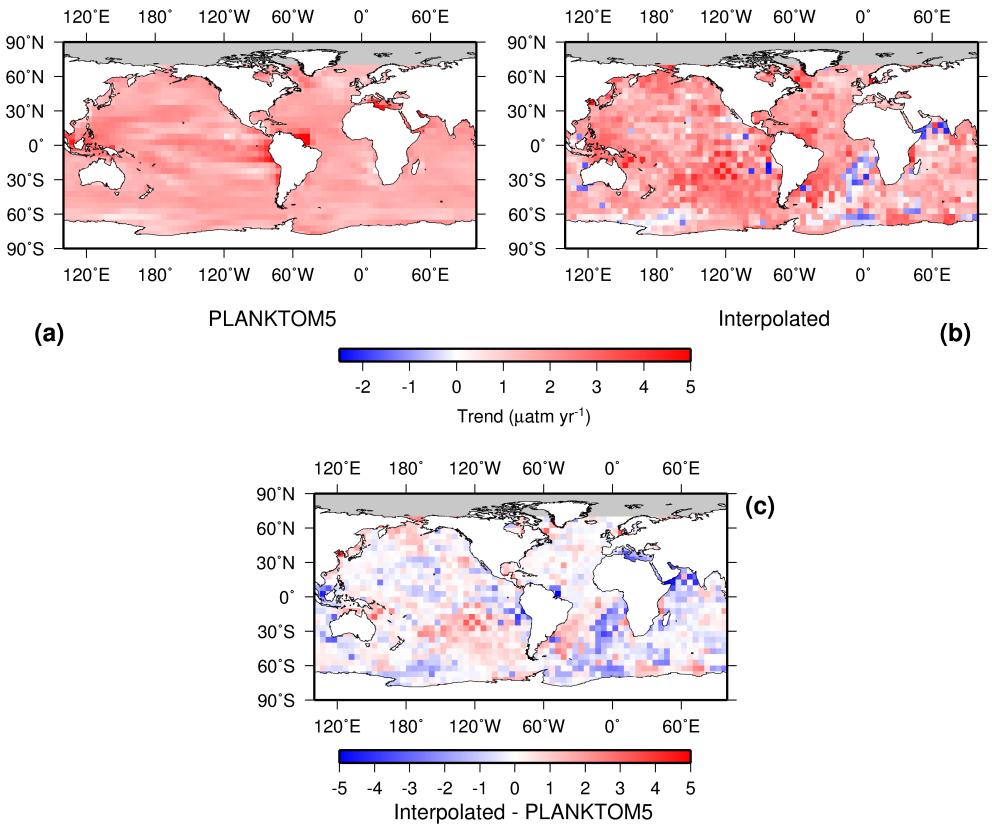

Figure 3.10: Global maps of linear trends in pCO_2 values (in μatm yr^{-1}) from 1990-2008 for (a) sub-sampled PlankTOM5 model output and (b) the interpolated data from the sub-sampled model output. (c) shows the differences between the two (in μatm yr^{-1}).

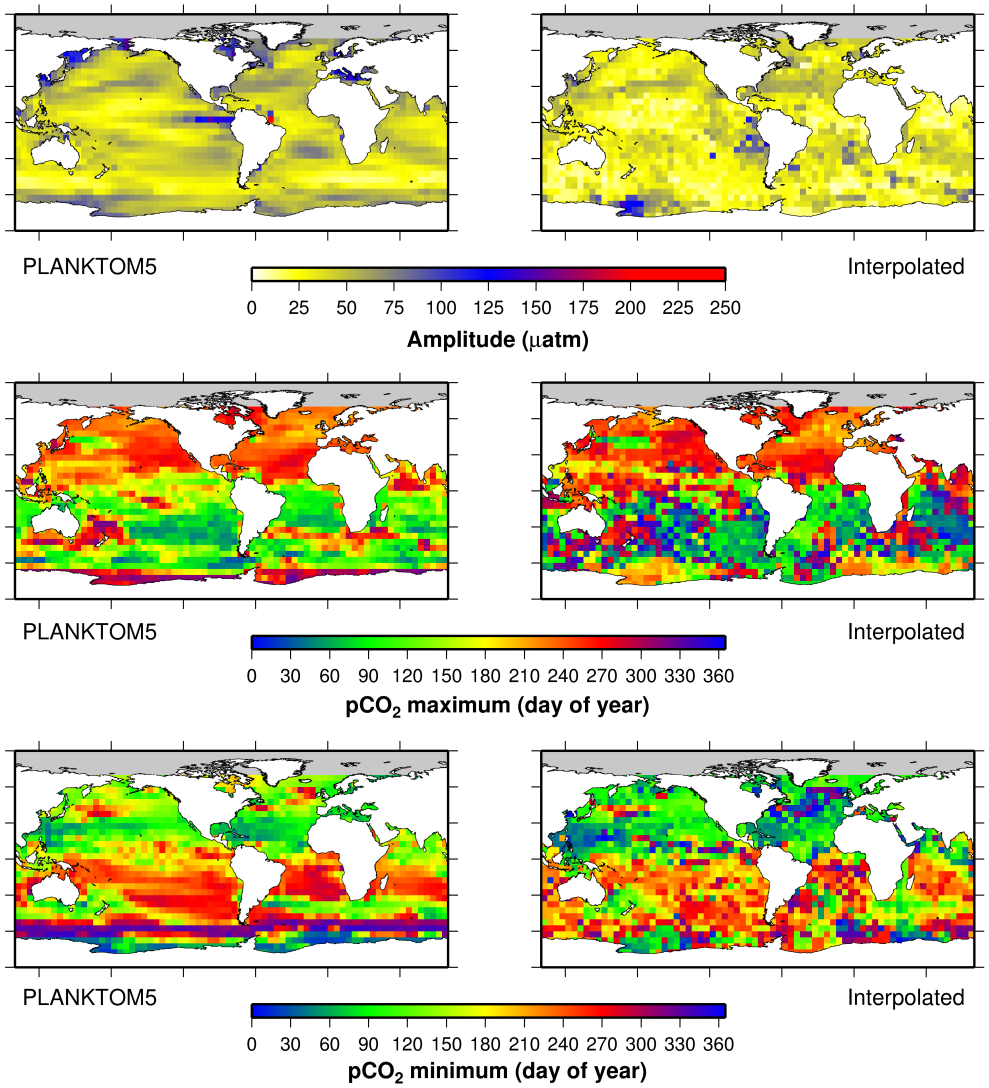

Figure 3.11: Comparison of the seasonal cycle of pCO_2 in the PlankTOM5 model (left) to the interpolation of the PlankTOM5 data. The comparison is in terms of (top) seasonal amplitude, (middle) the day of the pCO_2 maximum, and (bottom) the day of the pCO_2 minimum.

The results of these tests give a strong level of confidence that interpolating the LDEO database using this technique will give a good estimate of pCO_2 values over much of the global ocean and an appropriate assessment of the uncertainty. The performance of this interpolation also compares well with that of alternative techniques. Table 3.2 (page 67) shows comparisons of the errors calculated for interpolated values in this study with those in previous regional interpolations. In most cases the errors are very similar, and the largest differences occur when this study out-performs the others.

Study (Region)	**Method**	**Period**	**Error (μatm)**	**This Study Error (μatm)**
Wanninkhof et al. [1996][a] (Equatorial Pacific)	NO_3/SST regression	Aug. 1991	48	34
		Nov. 1991	3	4
Ono et al. [2004] (North Pacific)	SST/Chl regression	May 1997[b]	21	12
Jamet et al. [2007] (North Atlantic)	SST/Chl/MLD regressions	Winter 1994-1995	12.38-16.76[c]	11.34
		Spring 1994-1995	13.35-14.53[c]	12.53
		Summer 1994-1995	11.44-14.27[c]	10.25
		Autumn 1994-1995	8.98-17.33[c]	10.82
Watson et al. [2009] (North Atlantic)	SST/MLD regression	2002-2007[d]	1.8	1.66
	Neural network		0.77	
Telszewski et al. [2009] (North Atlantic)	Neural network	2004[e]	8.1	9.8
		2005[e]	12.6	9.7
		2006[e]	12.5	9.3

Table 3.2: Comparison of the mean RMS errors from interpolating model sub-sampled model output (see text) and error measures from other interpolation studies. [a]Only values after 1990 were compared. [b]Monthly means compared in both studies. [c]Three different regression were used, giving a range of errors. [d]Annual mean values are compared in both studies. [e]Used a 1°x1° resolution compared to this study's 5°x5° daily resolution; monthly means compared in both studies.

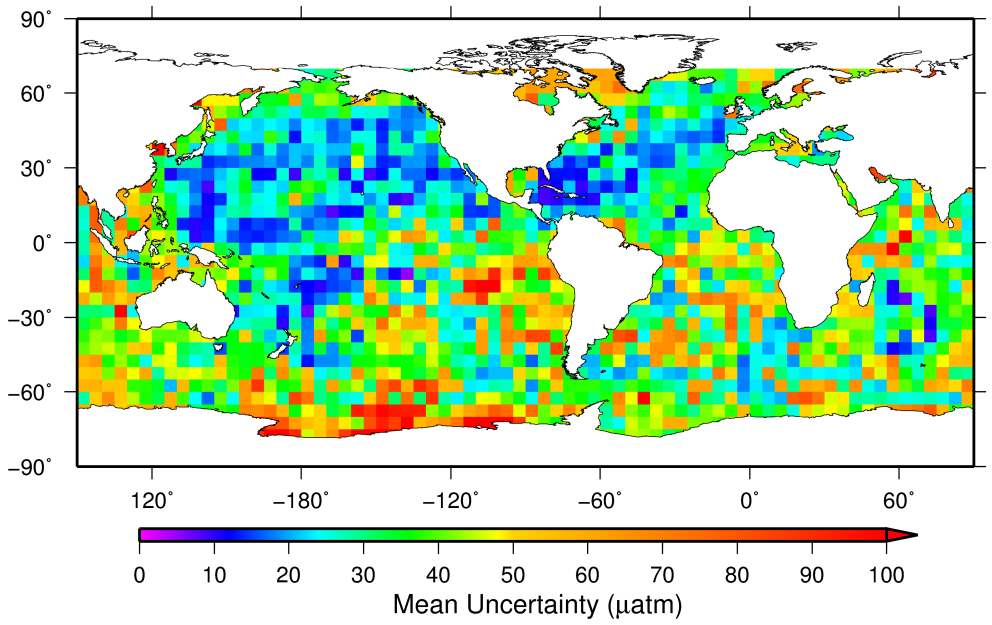

Figure 3.12: Mean uncertainty assigned to the interpolated values of each grid cell for the LDEO database. The color scale is identical to that used for the RMS errors (Figure 3.9, page 65) for easy comparison.

3.5 Results of interpolating the LDEO database

The implementation of this interpolation method on the LDEO database produces a data set consisting of daily pCO_2 values in the period 1990-2008 on a 5°x5° grid south of 70°N. The pattern of uncertainties (Figure 3.12) corresponds well with the data density of the original data set (Figure 3.1, page 50), with the smallest uncertainties in the north Pacific and north Atlantic and the largest uncertainties in the eastern equatorial and south Pacific, south Atlantic and Southern Ocean. The assigned uncertainties are typically larger than the predicted errors (Figure 3.9, page 65).

The interpolated pCO_2 data set exhibits a number of features that reflect some major aspects of the oceanic CO_2 cycle, visible in the seasonal mean pCO_2 field for the year 2000 (Figure 3.13, page 69). The mid-latitudes of the North Pacific and North Atlantic have low pCO_2 concentrations in winter and moderate concentrations in summer, due primarily to seasonal temperature changes tempered by biological activity and vertical transport of pCO_2 [*Ayers and Lozier*, 2012; *Schuster and Watson*, 2007]. Similar features are found in the mid-latitudes of the Southern Hemisphere [*Takahashi et al.*, 2002]. There is large seasonal variability of pCO_2 in the Southern Ocean, due to the combined effects of

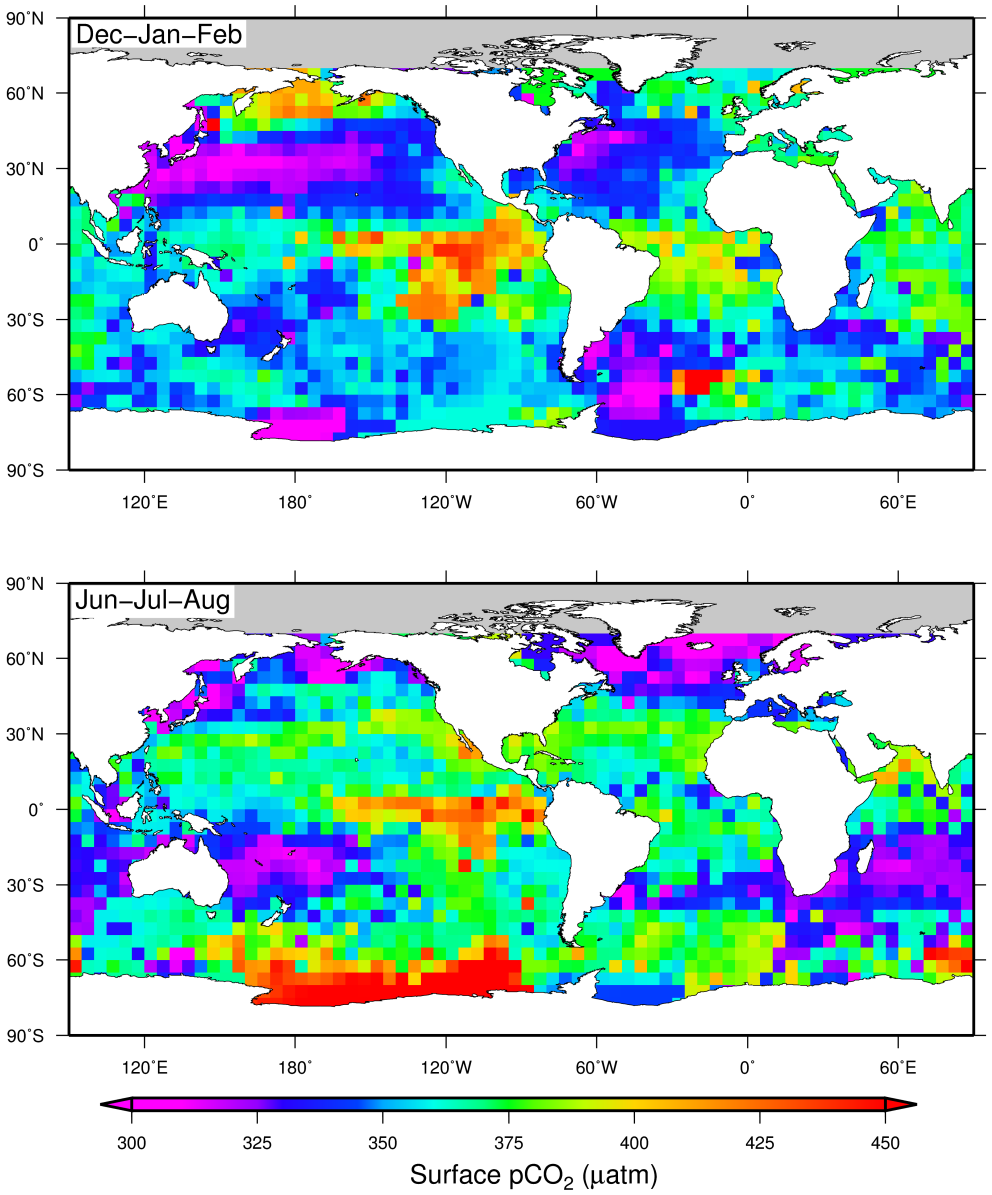

Figure 3.13: The mean pCO_2 field from the interpolated data set for two seasons: (top) December 1999 to February 2000, and (bottom) June to August 2000.

temperature changes and significant changes in biological activity due to light limitation in winter [*Metzl et al.*, 2006]. The consistently high levels of pCO_2 in the eastern Equatorial Pacific are a result of upwelling of CO_2-rich water at the eastern boundary of the Pacific along with transport of similar waters northward along the coast of South America [*Feely et al.*, 1999; *Etcheto et al.*, 1999].

The annual mean pCO_2 field and seasonal cycle from the interpolation compares well with that published by *Takahashi et al.* [2009] (Figure 3.14, page 71). Comparing the Takahashi climatology, which is normalised to the year 2000, and the annual mean

pCO_2 for 2000 from the interpolation shows very good similarity between the two (Figure 3.14a). The RMS error between the two maps is 13.69 μatm, and an overall pattern correlation of r^2=0.55. Separating the zonal and meridional mean correlations removes many of the effects of variability between individual cells, and gives a better picture of the coherence between the two data sets. This shows that the overall structure of the two data sets is very similar; the zonal mean correlation is r^2=0.64, and the meridional mean correlation is r^2=0.85. The lower zonal mean correlation is due to a divergence in the high northern latitudes, where the Takahashi climatology exhibits a sudden reduction in pCO_2 in both the North Pacific and North Atlantic. This difference is due to the interpolation's inability to work with the much sparser data in these high latitudes, meaning that values from lower latitudes are interpolated northwards.

The amplitude of the seasonal cycle in each data set (Figure 3.14b, page 71) is also very similar for much of the ocean. The RMS error between the two is 28.4 μatm, with a pattern correlation of r^2=0.43. The zonal and meridional pattern correlations of the seasonal amplitude are r^2=0.77 and r^2=0.40 respectively. Many of the spatial structures in the northern Pacific and Atlantic Oceans are very similar, as is the large seasonal amplitude in much of the Southern Ocean. The major difference between the two datasets is the larger seasonal cycles produced in the equatorial Pacific by the Takahashi climatology. This region is characterized by large interannual variability compared to the amplitude of the seasonal cycle, due in large part to ENSO activity [*Cosca et al.*, 2003; *Doney et al.*, 2009]. Reproduction of pCO_2 values in this region is therefore very difficult, and it is unsurprising that the Takahashi climatology and the interpolation give different outcomes. These differences are exacerbated by the fact that measurements from strong El Niño periods are removed prior to the calculation of the Takahashi climatology [*Takahashi et al.*, 2009], but were not removed for this study.

Comparisons of the timing of maximum and minimum pCO_2 values (Figure 3.14c and d) show strong zonal coherence between the Takahashi climatology and this study, with r^2=0.89 in both cases. Meridional differences are much larger (r^2=0.18 and r^2=0.36 respectively). The most likely cause for the low similarity is the differing method of interpolation between the two studies. The Takahashi climatology employs a transport model to determine missing values, and since the majority of oceanic transport is zonal

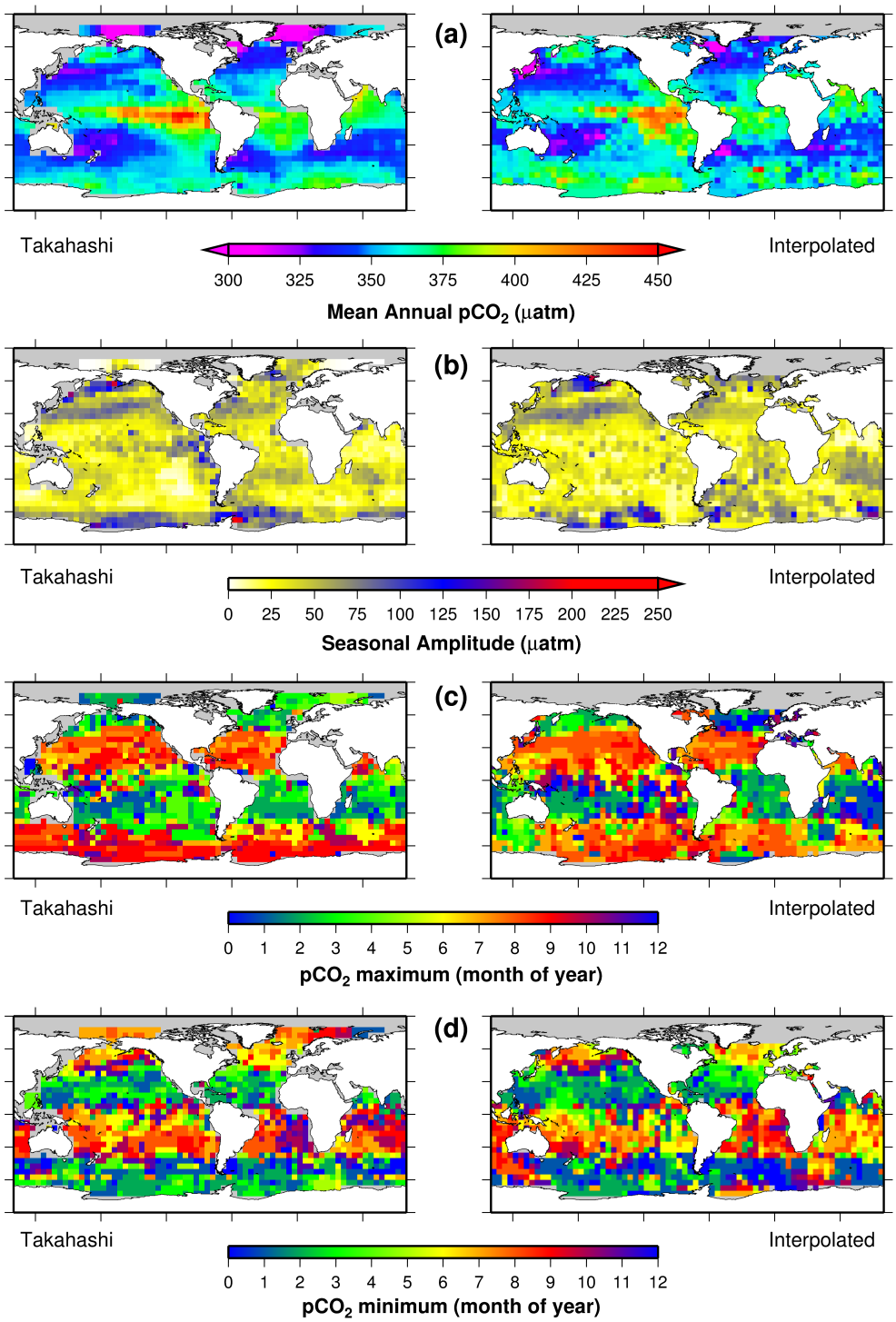

Figure 3.14: Comparison of (left) the climatology published by *Takahashi et al.* [2009] and (right) the mean pCO_2 concentration and seasonal cycle of the interpolated data set. The comparison is in terms of (a) the mean pCO_2 concentration in 2000, (b) seasonal amplitude, (c) the month of the pCO_2 maximum, and (d) the month of the pCO_2 minimum.

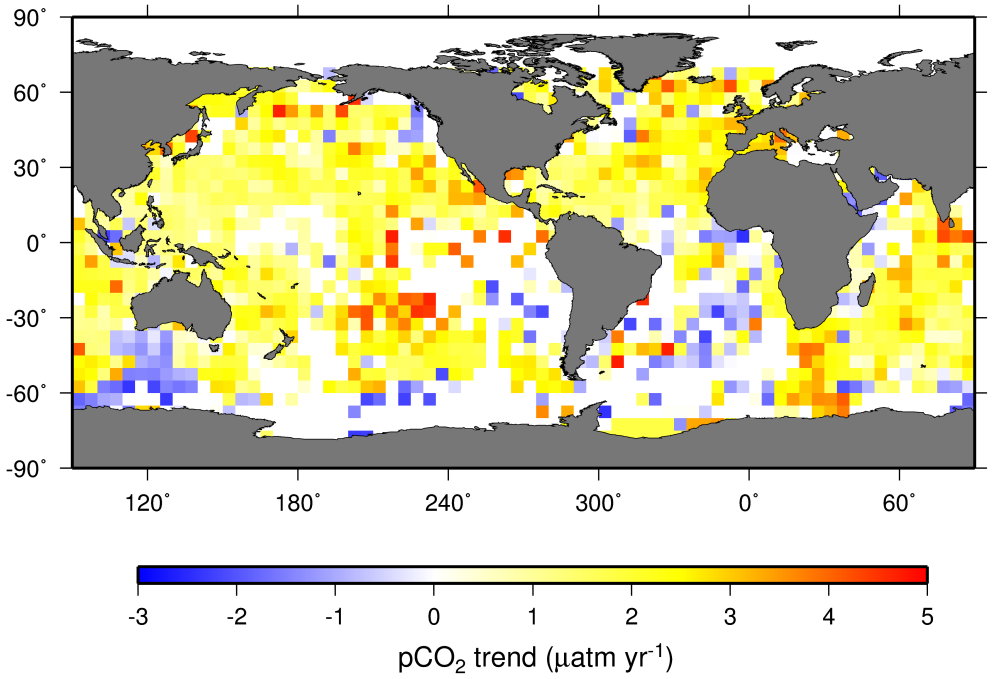

Figure 3.15: The linear trend of pCO_2 (in μatm yr^{-1}) for each grid cell of the interpolated data set.

there is much greater inter-cell similarity in this direction than found using our purely statistical technique.

The scarcity of pCO_2 measurements in much of the ocean has meant that global estimates of long-term pCO_2 trends are not available directly from the LDEO data set, so a thorough comparison of the trends in the interpolated data set (Figure 3.15) is not possible. However, it is likely that the negative trends seen in the South Atlantic are artifacts of the interpolation method, since the same trends were shown to be incorrect when validating against model output (Figure 3.10, page 65). The region of negative trends south of western Australia are also likely to be artifacts for the same reason, although attempts to assess trends from measurement data in this region have also been unsuccessful [*Inoue et al.*, 1999] so it is difficult to assess the likelihood of accuracy here. The North Atlantic and North Pacific, meanwhile, have yielded suitable trends for comparison [*Le Quéré et al.*, 2009]. Comparisons of the trends for individual grid cells in these regions (Figure 3.16) show that the two sets are comparable, although with considerable variability. This is due to the different methods of applying trend fits to the sparse data. A complete study of oceanic pCO_2 trends is beyond the scope of this study.

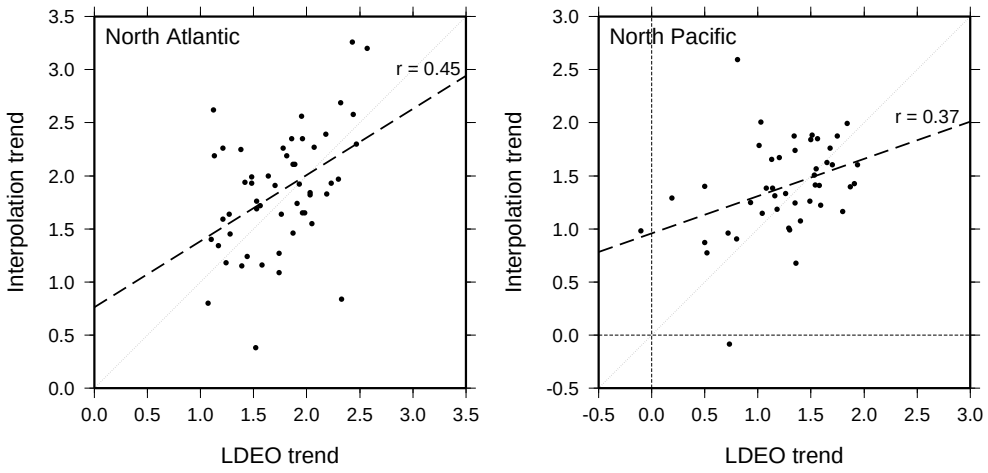

Figure 3.16: Comparison of linear trends calculated from the LDEO database [*Le Quéré et al.*, 2009] and the interpolated data set, for (left) north Atlantic and (right) north Pacific.

3.6 Calculation and verification of air-sea CO_2 fluxes

A complete interpolated data set of surface ocean pCO_2 such as that presented above allows calculation of air-sea CO_2 fluxes of the same temporal and spatial resolution. This in turn allows detailed examination of the global air-sea flux. The fluxes obtained from the interpolated data set were compared with the Takahashi climatology (adjusted to represent the year 2000) [*Takahashi et al.*, 2009] both globally and in a number of ocean regions.

Air-sea CO_2 fluxes were calculated using the standard formulation (Equation 1.5, page 11). Atmospheric CO_2 values were obtained from the GLOBALVIEW atmospheric CO_2 database [*GLOBALVIEW-CO2*, 2008]. Atmospheric CO_2 was adjusted to sea level pressure using data from the NCEP/NCAR Reanalysis [*Kalnay et al.*, 1996]. The gas transfer velocity was calculated using 6-hourly wind data from the CCMP Level 3 wind product from JPL [*Atlas et al.*, 2011], and the Reynolds SST data set [*Reynolds et al.*, 2002] and the Hadley Centre's EN3 salinity data set [*Ingleby and Huddleston*, 2007] were used to calculate solubility. Finally, fluxes were limited in high latitudes due to sea ice cover with data from the Hadley Centre's sea ice cover database [*Rayner et al.*, 2003].

Maps of the global air-sea CO_2 flux for February and August 2000 are shown in Figure 3.17 (page 75). These are accompanied by the same maps from the Takahashi climatology for comparison. Overall the two products are very similar, with all the major

features duplicated across the maps with some variation in their spatial extent. Even relatively small features, such as the February outgassing in the northern North Pacific and between Iceland and Greenland are reproduced (Figure 3.17a and b). The largest difference occurs in the Southern Ocean in August. The Takahashi climatology (Figure 3.17d) shows a strong outgassing along a narrow band at approximately 60°S, which follows the edge of the ice cover. This band is also present in the fluxes from the interpolation (Figure 3.17c), but is much less well defined. This is caused by the relative lack of pCO_2 measurements in the Southern Ocean, which means that those that are available are interpolated between neighbouring grid cells. The same effect can be seen to a lesser extent in Equatorial Pacific. There is a region of outgassing related to what is known as the 'cold tongue' of upwelled waters spreading from the eastern edge of the Pacific and extending to approximately 180° longitude. This 'tongue' is restricted to a narrow zonal band in the Takahashi climatology, but in the interpolated data it extends much further south, again due to data being interpolated into regions of few measurements. The difference in the meridional extent of the cold tongue in both February and August is due to the fact that measurements taken during El Niño events are excluded from the Takahashi climatology. These events suppress the upwelling of carbon-rich water in the eastern Equatorial Pacific, which in turn reduces the rate of outgassing [*Feely et al.*, 2006], explaining the smaller extent of the outgassing region in the fluxes calculated from the results of this study's interpolation.

The seasonal cycle of the fluxes in different regions of the global ocean has been calculated for both the interpolated data set and the Takahashi climatology (Figure 3.18, page 76). The fluxes in the North Pacific and Equatorial/South Atlantic (Figures 3.18a and e) match well. For the former this is not surprising, since this region contains the most measurements and therefore constrains the interpolation. The latter is less expected due to the small number of measurements available in this region (Figure 3.1, page 50). The Southern Ocean (Figure 3.18d) exhibits a similar shape in both cases, but the interpolation has a lower sink efficiency and acts as a source of atmospheric CO_2 during more of the year. This is due to the extended region of outgassing along the ice boundary discusses earlier. There is a large difference in the seasonal cycles in the South Pacific (Figure 3.18c). This is again due to the sparse measurements, and means that outgassing in the Equatorial Pacific is interpolated southwards into the South Pacific region, and

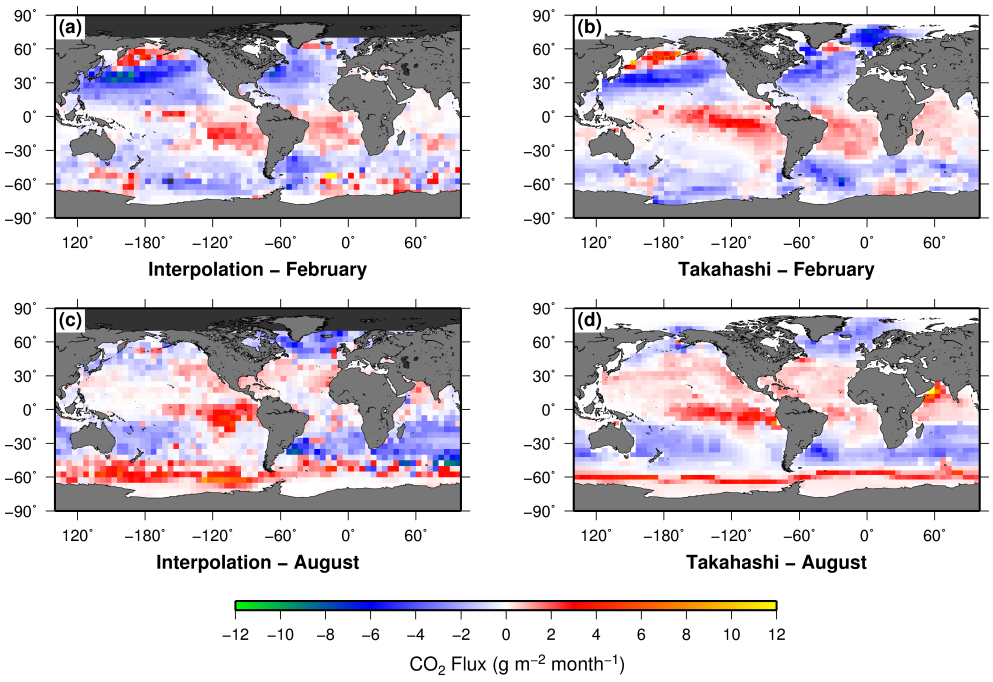

Figure 3.17: Maps of air-sea CO_2 flux for (top) February and (bottom) August for the year 2000. Maps on the left are calculated from this study's interpolated data set, while those on the write are from the Takahashi climatology [*Takahashi et al.*, 2009]. Negative (blue) and positive (red) represent fluxes into and out of the ocean respectively.

neurtralises the increased sink observed during the winter months. The Equatorial Pacific itself (Figure 3.18b) shows little seasonal variability as would be expected in this region. In both the interpolated data and the Takahashi climatology the region is a year-round source of atmospheric CO_2. The rate of outgassing is much lower in the interpolated data because it includes the effects of ENSO events, which were removed from the Takahashi climatology.

The Indian Ocean (Figure 3.18g) is similar in both data sets except for a much larger sink in the interpolated data during the winter months. The converse is true in the North Atlantic (Figure 3.18f), with a much larger sink seen in the Takahashi climatology during winter. In both cases this is due to differences in the way that pCO_2 measurements are interpolated. Takahashi used a transport model to 'move' data between grid cells and does so for all values regardless of data density. This study copies pCO_2 values between grid cells and only does so when required. In the Indian Ocean, there are measurements in a few season of strong oceanic sinks in wintertime. In copying these values across cells, this study's interpolation maintains the strength of the sink across a larger area of the Indian

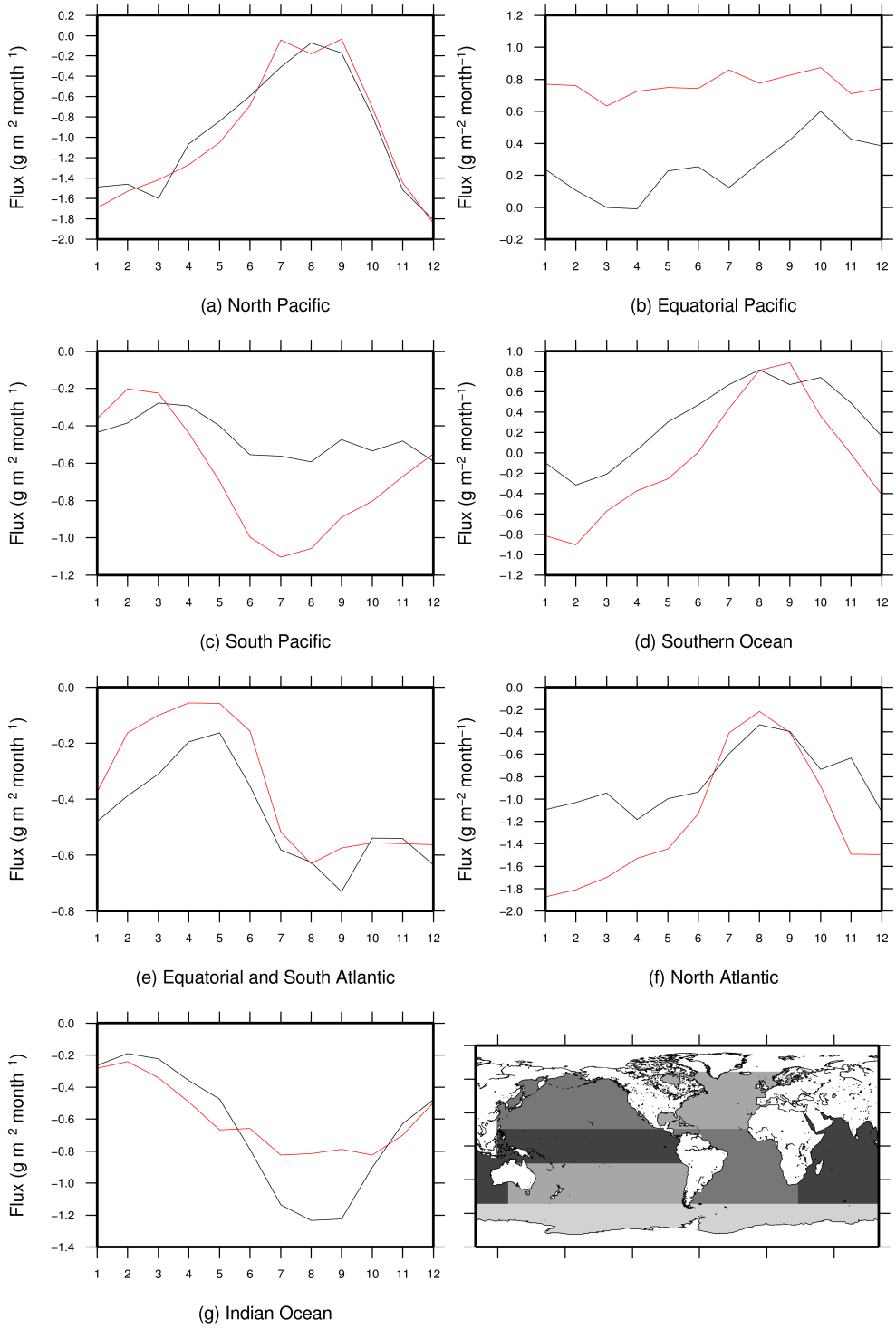

Figure 3.18: Seasonal cycles of air-sea CO_2 fluxes in seven ocean regions for the interpolated data set (black lines) and the Takahashi climatology (red lines). The map in the bottom right shows the various regions: North Pacific (15°N to 70°N); Equatorial Pacific (15°N to 15°S); South Pacific (15°S to 50°S); Southern Ocean (South of 50°S); Equatorial and South Atlantic (15°N to 50°S); North Atlantic (15°N to 70°N); Indian Ocean (North of 50°S).

Ocean, while the physical transport approach used by Takahashi allows the strong sink to dissipate as the low-CO_2 water dissipates through the ocean. In the North Atlantic there are a large number of available measurements (Figure 3.1, page 50), so this study does not spatially interpolate the data to such a large extent. Takahashi approach, meanwhile, continues to transport water between grid cells. The different results of these approaches can be seen in the comparison of the February flux maps (Figure 3.17a and b, page 75). There is a region of strong CO_2 sink off the east coast of North America, which spreads across much of the North Atlantic in the Takahashi climatology. In this study there are plenty of measurements available adjacent to this region, so spatial interpolation is not required and the region of strong sink is not extended so far across the Atlantic.

3.7 Conclusion

This chapter presents a new interpolation method adapted to the available measurement coverage of surface ocean pCO_2 values over the 1990-2008 period south of 70°N, demonstrating comparable or improved accuracy compared to other approaches. The interpolated data set includes uncertainties based on the spatial and temporal distance over which values have been interpolated, which will be useful to guide data selection in future studies. Such a data set can be used to assess pCO_2 variability over multiple temporal and spatial scales, and help to establish the most beneficial placement and frequency of future pCO_2 observation programs to reduce the uncertainties in our knowledge of this key oceanic variable. Our interpolated data set can also provide prior estimates required in atmospheric inversion models, and data to evaluate forward model simulations.

This new interpolated data set provides coherent coverage for the global ocean with greater temporal and spatial extent than existing data products. The technique relies neither on knowledge of other oceanic variables nor on the physical characteristics of the ocean. There is therefore no need to establish these relationships between variables in situations where there is insufficient data to do so, nor to take into account how they might change over space and time and recalculate them accordingly. It is therefore possible to interpolate data in regions where these previous approaches are inapplicable. The independent statistical nature of the technique means that it can be readily applied to other environmental global data sets with comparatively little effort.

Chapter 4

The Ocean pCO_2 Response to Modes of Climate Variability

4.1 Abstract

We quantify the magnitude and extent of the response of the partial pressure of surface ocean CO_2 (pCO_2) to interannual climate variability. Our results are based on observations taken during the period 1990-2007, and informed by model output to assess their representativeness. We find a strong relationship between pCO_2 and the El Niño Southern Oscillation in the Equatorial Pacific (15°S-15°N; r = -0.47), where pCO_2 levels changed by -6.6 $\pm$ 1.0 μatm per index unit (μatm iu^{-1}), which corresponds to a change in air-sea flux of 0.118 $\pm$ 0.018 Pg C yr^{-1} iu^{-1} over 4.029 $\times$ 10^7 km^2. pCO_2 levels were low when the SOI was in its negative (El Niño) phase. The Pacific Decadal Oscillation (PDO) has the strongest effect in the North Pacific (North of 15°N; r = 0.25), although the pCO_2 response of -0.8 $\pm$ 0.4 μatm iu^{-1} is weak. We do not find a statistically significant correlation between the North Atlantic Oscillation and pCO_2 in the North Atlantic North of 15°N, although the pattern of correlations across the basin supports previous findings. The PDO shows a statistically significant correlation with pCO_2 in the North Atlantic, albeit with a weak pCO_2 response of -1.4 $\pm$ 1.0 μatm iu^{-1} and a corresponding change in air-sea flux of 0.057 $\pm$ 0.024 Pg C yr^{-1} iu^{-1} over 3.112 $\times$ 10^7 km^2.

4.2 Introduction

The interannual variability of the global climate is complex and difficult to predict. However, various large-scale oscillations repeat on multi-year time scales, and many of the related and consequent effects have been observed and quantified. These 'modes of variability' are quantified using climate indices. They are commonly defined as teleconnections in atmospheric pressure systems [*Barnston and Livezey*, 1987; *Ropelewski and Jones*, 1987; *Hurrell*, 1995; *Limpasuvan and Hartmann*, 1999; *Thompson and Wallace*, 2000] or as regional changes in sea surface temperature [*Trenberth*, 1997; *Mantua et al.*, 1997; *Zhang et al.*, 1997].

These modes of climate variability affect the carbon cycle and the partial pressure of carbon dioxide (pCO_2) at the ocean surface, either directly through the temperature-dependence of pCO_2, or indirectly through the impact of ocean circulation and biological fluxes on dissolved inorganic carbon (DIC). Relationships between pCO_2 and climate

indices have been observed in the equatorial Pacific [e.g. *Inoue et al.*, 1996; *Etcheto et al.*, 1999; *Feely et al.*, 2006; *Sheu et al.*, 2010] and North Atlantic [*Corbière et al.*, 2007; *Schuster et al.*, 2009], but the magnitude of the response of pCO_2 to changes in the indices and the spatial extent of those responses has not been directly quantified before.

The recent release of a global database of surface ocean pCO_2 measurements (the Surface Ocean CO_2 Atlas (SOCAT) [*Pfeil et al.*, submitted] presents an invaluable opportunity to quantify the impact of climate variability on surface ocean pCO_2, and to determine the spatial extent of influences. Here we compare the interannual variability of surface ocean pCO_2 to changes in several climate indices between 1990 and 2007, and determine the direction, magnitude and spatial extent of their relationship. This work will contribute to improving our understanding of the relationship between climate variability and biogeochemical processes in the ocean, and help quantify interannual variations in air-sea CO_2 fluxes.

4.3 Methods

4.3.1 Calculation of anomalies

The seasonal cycle and long-term trend are the largest components of the temporal variability of surface ocean pCO_2 (Chapter 2). These were removed from all the data sets used here as explained below, leaving anomalies that represented the interannual variability alone.

4.3.2 SOCAT measurements

The SOCAT measurements on their own provide little opportunity to determine how the interannual variability (IAV) of pCO_2 is related to large scale modes of climate variability, because the seasonal cycle is large in many regions and difficult to isolate with the available measurements. To minimise this problem, we first interpolated the SOCAT data to produce spatially and temporally complete set of pCO_2 data. The SOCAT database contains over 6 million measurements covering the period 1968-2007. However, only 10% of the available measurements were taken before 1990, and these are too sparsely distributed to produce a reliable interpolation. We therefore used only the measurements from the

period 1990-2007. The measurements were binned on a 5°x5° grid with daily resolution. Where multiple measurements exist for a particular day, the mean value was calculated. In leap years, each 'day' was calculated as $1\frac{1}{365}$ days to provide a consistent year length throughout. A curve was fitted to each grid cell's time series of the form:

$$f(t) = a_0 + a_1 t + \sum_{k=1}^{n} [b_{2k-1} sin\,(2\pi kt) + b_{2k} cos\,(2\pi kt)] \tag{4.1}$$

where f is the fitted curve, t is the time in years since 1990 and n is a number of harmonics representing the seasonal cycle. Up to four harmonics were fitted where the data density allowed. If a curve could not be fitted to the time series, values from neighbouring grid cells were added according to a 'radius of influence' [*Levitus*, 1982] based on the spatial autocorrelation characteristics of surface pCO_2 (Chapter 2). These interpolated values were weighted according to the decorrelation length of pCO_2 values in that region (Chapter 3). This process was repeated until a realistic curve fit was achieved.

The curve fitting process produced a linear trend and mean seasonal cycle for each grid cell, which were removed from the cells' original time series to leave anomalies from the expected seasonal cycle. These were converted to monthly mean anomalies to match the temporal resolution of the climate indices.

4.3.3 Models

The interannual variability of the SOCAT observations alone is not sufficient to provide a robust result. Since the measurements in each grid cell are temporally sparse, it is not possible to tell whether or not the results obtained are representative of the values that would be obtained from a complete data set. We therefore used model data to constrain the output of the results obtained from the SOCAT data.

We use the pCO_2 output data of eight model simulations from five different models contributing to the RECCAP project [*Canadell and Ciais*, in preparation]: the Bergen model MICOM-HAMOCC [*Assmann et al.*, 2010], the CSIRO model [*Lenton and Matear*, 2007], the WHOI model BEC [*Thomas et al.*, 2008], the ETH model (two simulations) [*Graven et al.*, submitted], the UEA model NEMO-PlankTOM5 (three simulations) [*Le Quéré et al.*, 2010]. The two ETH model runs used different gas transfer formulations: ETHk19 used a coefficient of 0.31 cm hr^{-1} s^2/m^2 for the gas transfer formulation leading

to a global mean velocity of 19 cm hr^{-1} for CO_2; ETHk15 used a coefficient of 0.24 cm hr^{-1} s^2/m^2 leading to a global mean velocity of 15 cm hr^{-1} for CO_2. The UEA model runs were each forced with a different wind reanalysis product: the NCEP/NCAR reanalysis [*Kalnay et al.*, 1996], ERA-Interim from ECMWF [*Dee et al.*, 2011] and the CCMP Level 3 wind product from JPL [*Atlas et al.*, 2011].

The pCO_2 data from each model was regridded to match the 5°x5° grid used for the SOCAT data. The regridded data from the eight simulations were averaged to produce an ensemble mean pCO_2 field for comparison with the SOCAT data. The long-term linear trend was removed from each grid cell of the ensemble mean, along with the mean seasonal cycle calculated from the detrended time series by computing the average value for each calendar month. The remaining anomalies represented the interannual variability equivalent to those extracted from the SOCAT data.

The anomalies produced above were more comprehensive than those from the SOCAT data since there were no temporal gaps in the model output. A second set of anomalies were produced by selecting values from the model output at those times and grid cells where anomalies exist in the SOCAT data. Comparing the models' complete and sampled anomalies provided information regarding the realism of the observed SOCAT relationships with climate indices.

4.3.4 Comparison with climate indices

The calculated pCO_2 anomalies were partitioned into three ocean regions (Figure 4.1, page 85): The Equatorial Pacific between 15°S and 15°N; the North Pacific between 15°N and 60°N; and the North Atlantic between 15°N and 60°N. The anomalies in these regions are correlated with a set of indices representing major modes of climate variability (Figure 4.2, page 86). The North Atlantic Oscillation (NAO) [*Barnston and Livezey*, 1987; *Hurrell*, 1995] is an index of the difference in sea level pressure (SLP) between Iceland and the Azores, which influences surface temperatures, winds and precipitation rates. ENSO indices are a measure of SST anomalies in the Equatorial Pacific, signalling the suppression of upwelling in the east. In this study we use the El Niño Region 3.4 SST index (hereafter referred to as Niño34) [*Trenberth*, 1997] which measures ENSO events at their largest spatial influence across the Equatorial Pacific in the region covering 5°S to

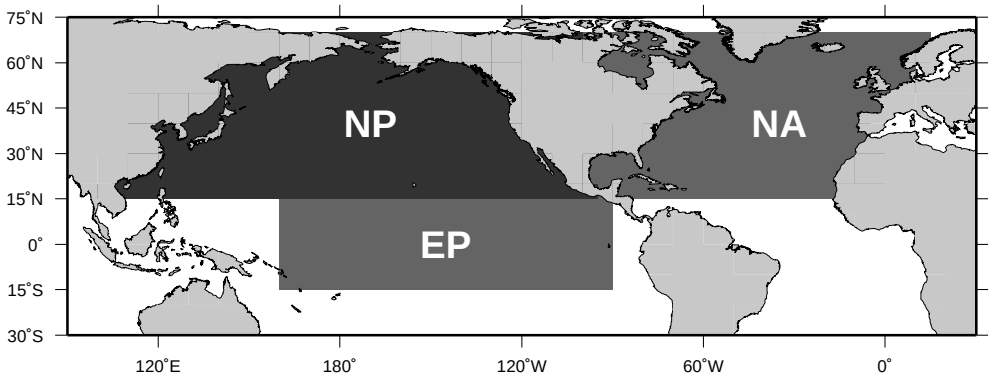

Figure 4.1: The three ocean regions used in this study.

5°N and 120°W to 170°W. The Southern Oscillation Index (SOI) measures a dipole of sea level atmospheric pressure between Tahiti and Darwin, Australia [*Ropelewski and Jones*, 1987], and is frequently used to measure the atmospheric effects of ENSO activity. The Niño34 and SOI indices measure different aspects of the same climate phenomenon, but whether the indices have equivalent relationships to aspects of the biogeochemical cycle is currently unknown. Finally, the Pacific Decadal Oscillation (PDO) is a measure of shifts in climate regimes across the North Pacific whose effects can be detected in a number of climatic variables including SST, SLP, air temperature and precipitation [*Mantua et al.*, 1997]. The index is based on the spatial pattern of SSTs in the North Pacific. The PDO has far less high frequency (monthly) variability than the NAO because it is based on SST rather than SLP. Each of the indices was detrended and deseasonalised to ensure a consistent comparison with the SOCAT and mean model anomalies. Correlations were calculated for the pCO_2 measurements outside the three regions shown in Figure 4.1, but the results were not robust because there were insufficient measurements.

The NAO exhibits very high month-to-month variability which hides the interannual signal. Since this index is primarily a winter-summer shift in relative atmospheric pressure between two locations [*Hurrell*, 1995], three alternative versions of the index were produced and correlated with the pCO_2 anomalies (Figure 4.2, page 86): A six-month mean (January-June, July-December), winter and summer means (December-February and June-August), and winter only (December-February).

As the indices are indicative of large-scale features of the global climate, there are similarities between them. Cross-correlation of the indices (Table 4.1, page 87) shows that the

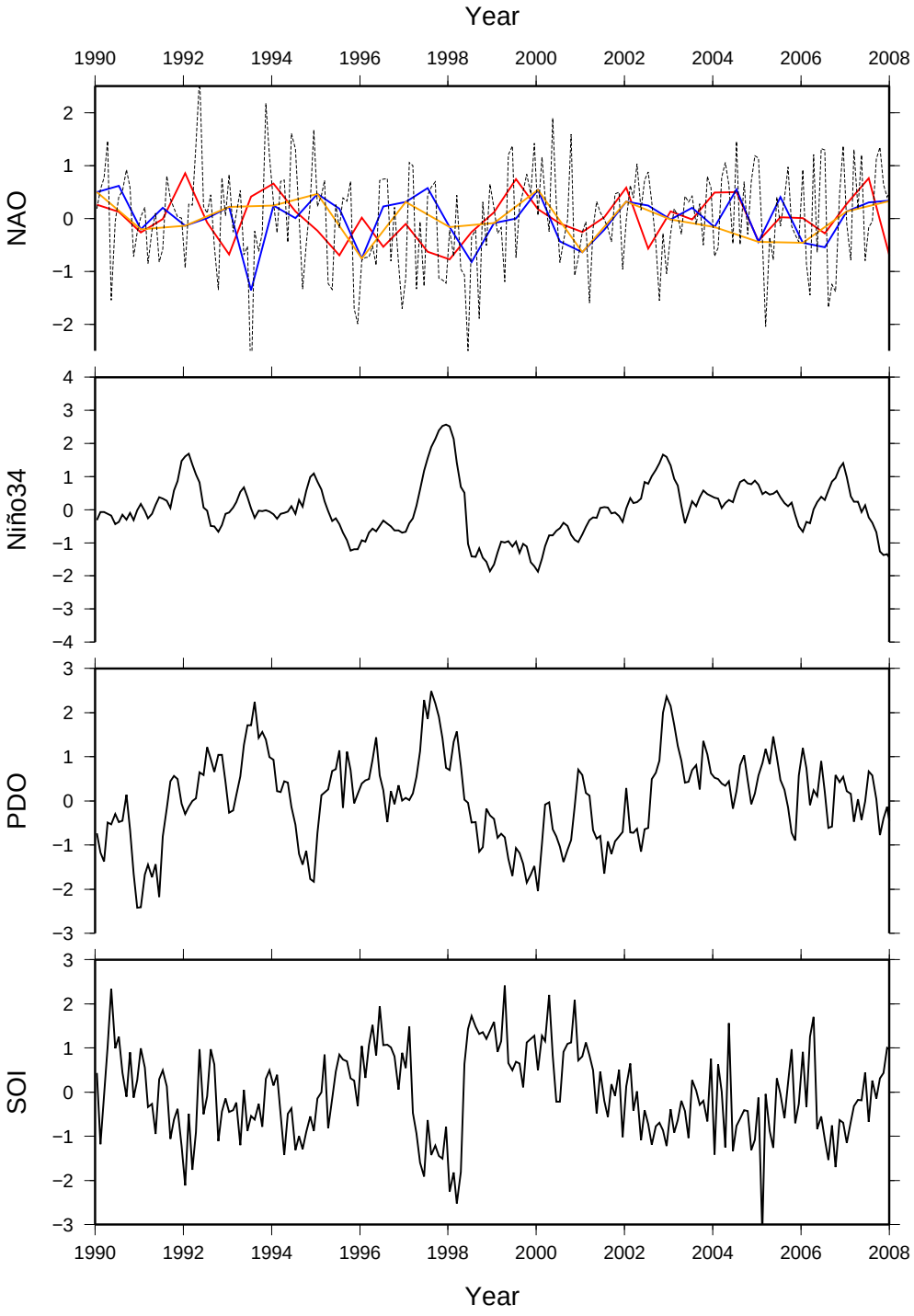

Figure 4.2: Detrended and de-seasonalised climate indices used in this study. All indices are measured at one-month intervals. The top panel shows the variations of the NAO index used in the study: The original monthly index (dashed line), the six-month mean (red), the December-January-February/June-July-August (DJF-JJA) means (blue) and the DJF mean (orange).

Niño34 and PDO indices are quite well correlated, which is caused by ENSO's influence as a driver of the PDO [*Newman et al.*, 2003]. The SOI is strongly anti-correlated with Niño34, as ENSO is a coupled atmosphere/ocean system with effects in both domains.of the changes in SST related to ENSO propagate into the atmospheric system [*Rasmusson and Wallace*, 1983]. There is also a small negative relationship between the NAO and PDO.

4.3.5 Basin-wide correlations

Basin-wide correlations were calculated by combining all the pCO_2 anomalies for each region into a single time series and correlating it with the climate indices. The correlation coefficient and slope of the fitted line from the correlation calculation show the robustness and strength of the relationships in each case. This simple approach provides a perspective on the relationships over large spatial scales, incorporation as many of the available pCO_2 measurements as possible.

4.3.6 Per-cell correlations

Assessment of the relationships between pCO_2 and climate indices on sub-basin scales allowed a finer understanding of the spatial extent of those relationships. Each set of pCO_2 anomalies (SOCAT observations, complete model mean and sampled model mean) was correlated with each of the indices for all 5°x5° grid cells to assess the spatial extent of the impact of climate modes on pCO_2. The correlation coefficient (r) of each grid cell indicates the strength of the pCO_2-index relationship. The sign and magnitude of the pCO_2 response to the changes in index values calculated as the slope of the linear fit (Figure 4.8, page 97) provides information on the underlying processes.

For the SOCAT data (Figure 4.3, page 88) and the sampled model data (Figures 4.4 and 4.5 c and d, pages 89 and 90), grid cells with 5 or fewer measurements were discarded

	NAO	**Niño34**	**PDO**
Niño34	−0.17		
PDO	−0.24	0.42	
SOI	−0.09	−0.69	−0.39

Table 4.1: Cross-correlations (r) of the indices of climate variability used in this study.

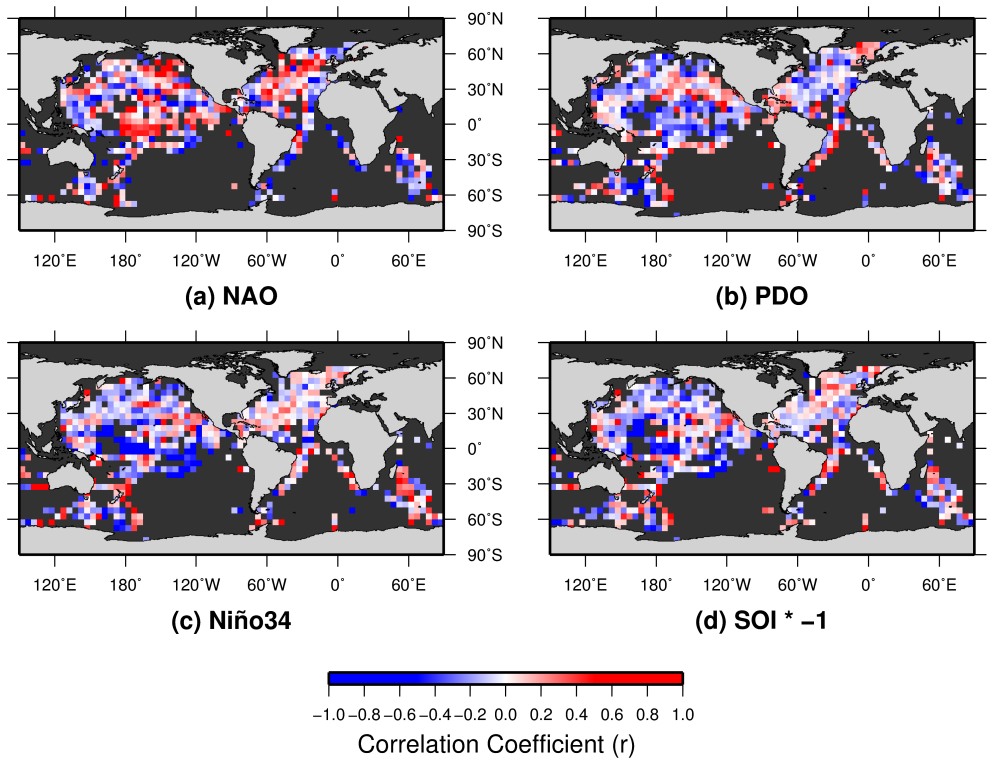

Figure 4.3: Correlations of observed surface pCO_2 anomalies with the climate indices. The pCO_2 anomalies have been computed by removing the mean seasonal signal and the trend for each variable (see text). Monthly anomalies are used for the PDO, Niño34 and SOI indices. 6-monthly anomalies are used for the NAO. The pCO_2 data are from the Surface Ocean CO_2 Atlas [*Pfeil et al.*, submitted]. See text for the source of the climate indices. All data are analysed over the 1990-2007 time period. The SOI correlations have been inverted for simpler comparison.

because the sample was too small. In the maps generated from the complete model output (Figures 4.4 and 4.5 a and b), correlations calculated for the eight individual model simulations were used to establish the model agreement. The models were said to agree if at least three quarters of the simulations showed the same sign of correlation coefficient in a given grid cell. For all indices the area of inter model agreement covers at least 79% of the ocean, giving confidence that the model mean is representative of the characteristics of the relationship between pCO_2 and the climate indices.

The correlations obtained from the sampled model mean (Figures 4.4 and 4.5 c and d, pages 89 and 90) were examined in greater detail to determine where they reproduced the complete model output (Figures 4.4 and 4.5 a and b), and identify where sufficient data is available for our analysis. This was done first by randomly sampling the complete model mean at each grid box, to create 100,000 samples with the same number of data as in the sampled model mean. Correlations were then computed for all samples, and a probability

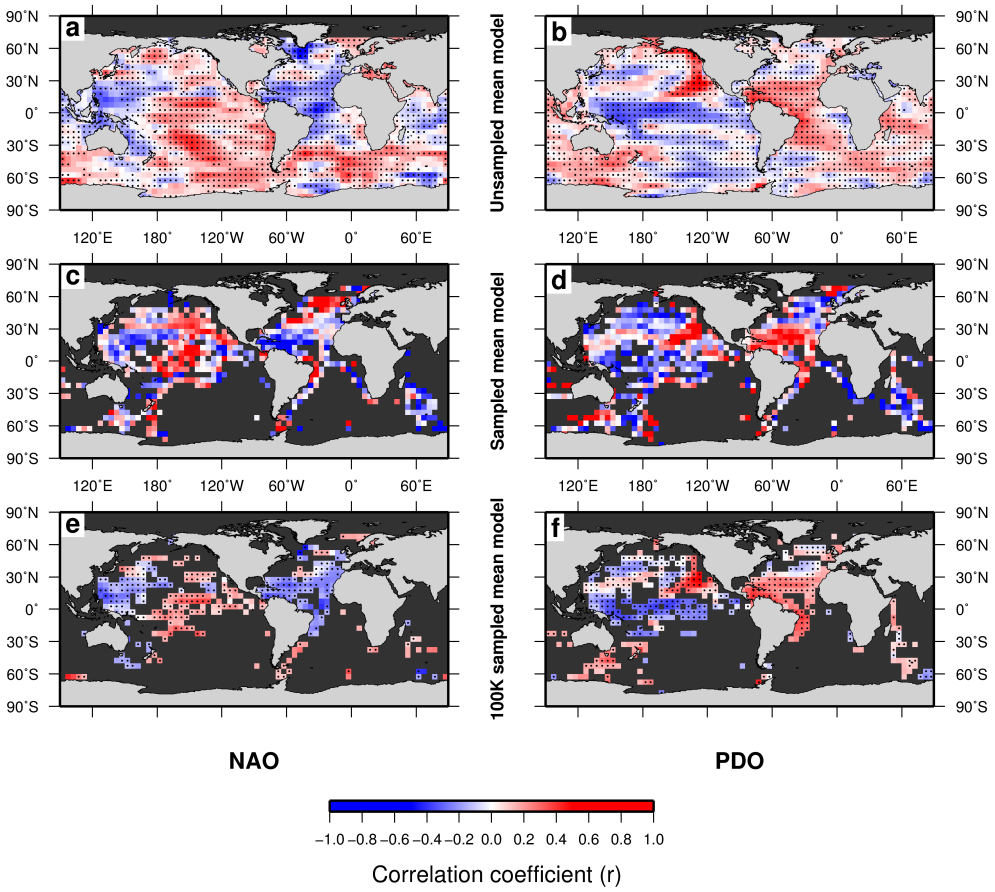

Figure 4.4: Correlations of modelled pCO_2 anomalies with the NAO (left) and PDO (right) climate indices. The modelled pCO_2 anomalies are from an ensemble average of eight model simulations. Top: Correlations calculated using the complete ensemble model mean output (every month between 1990 and 2007). Dots indicate cells where 75% or more of the eight models agree on the sign of the correlation. Middle: Correlations calculated from the model output sampled at the same locations in space and time as the SOCAT observations. Bottom: Statistical estimate of the most likely correlation when using the same number of data points as observed for each box but randomly distributed in time. The dots show cells whose correlation sign is consistent with both the sampled and unsampled models, and therefore represent the locations where the uneven sampling does not bias the correlation results. The dotted mask is applied to the correlations computed with the observations in Figure 4.3 (page 88).

distribution of correlations was produced for all grid cells (see examples in Figure 4.6, page 90). The peak value in the histogram is the most likely correlation coefficient for that grid cell. The grid cell was deemed representative of the true correlation if 63.2% ($1-\frac{1}{e}$) or more of the values in the histogram were of the same sign (i.e. clearly different from zero), or if the peak value $\pm$ 0.1 encompassed at least 36.79% ($\frac{1}{e}$) of all sample values (i.e. the correlation is consistent regardless of sampling). Between 60% and 68% of the grid cells were found to be representative of the true correlation. Second, the correlation coefficients for each cell calculated from the 100,000 samples (Figures 4.4 and 4.5 e and

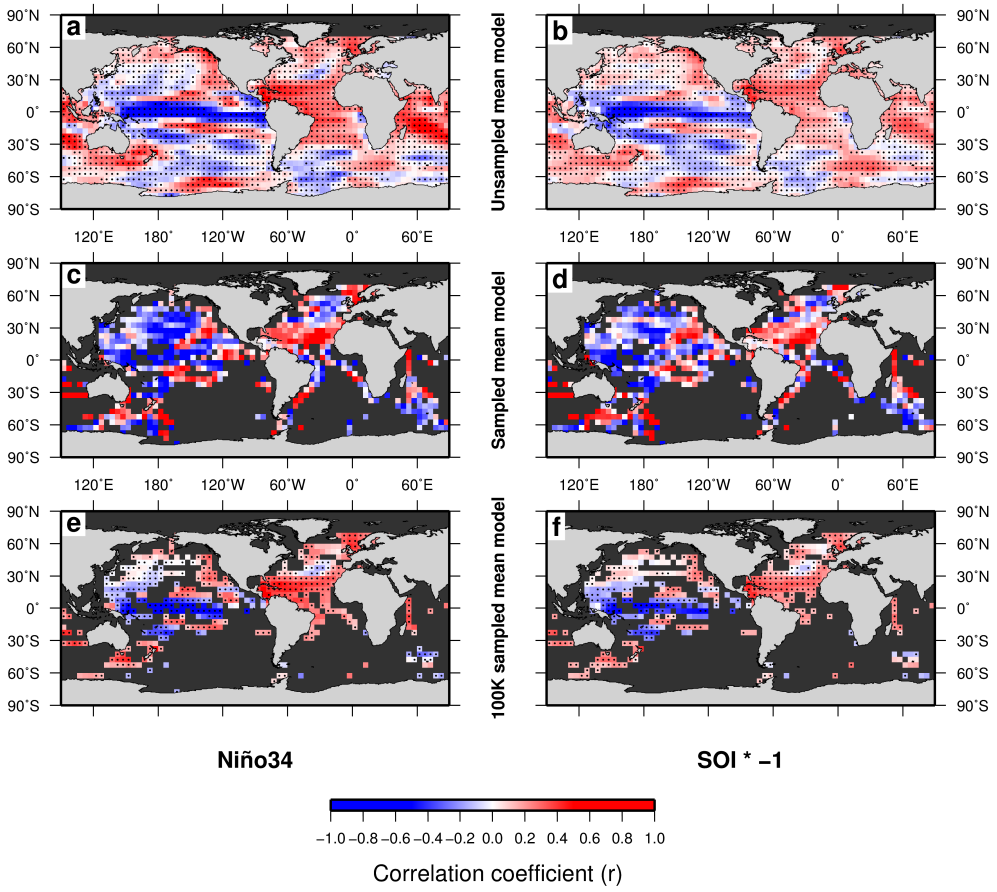

Figure 4.5: As Figure 4.4 (page 89), but for (left) Niño34, and (right) the SOI. The SOI correlations have been inverted for simpler comparison.

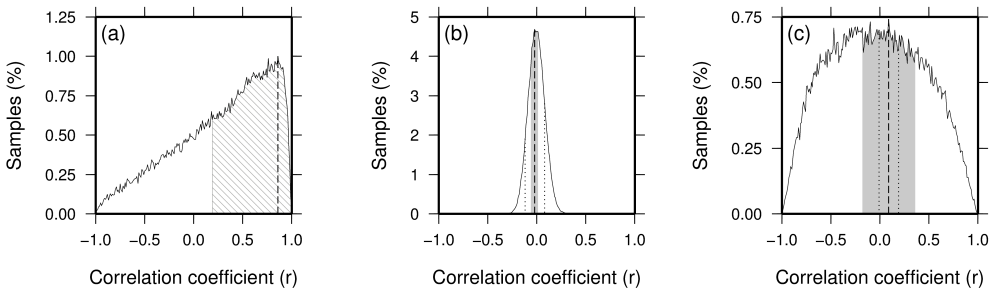

Figure 4.6: Example of histograms of the 100,000 correlation coefficients computed at each grid cell from the sampled model output, used to determine whether or not the calculated correlations against the Niño34 index were robust. The vertical dashed line indicates the most common correlation coefficient. In (a) more than 63.21% ($1-\frac{1}{e}$) of the samples were positive (striped) and the cell is considered robust. In (b) the range of correlation coefficients encompassing 36.79% ($\frac{1}{e}$) of values (shaded) is within 0.1 of the most common value (dotted lines) and the cell is considered robust. In (c) neither condition is met and the cell is discarded. These examples are taken for three grid cells at locations (147.5°W, 17.5°N), (157.5°E, 37.5°N), and (75.5°E, 57.5°S) respectively.

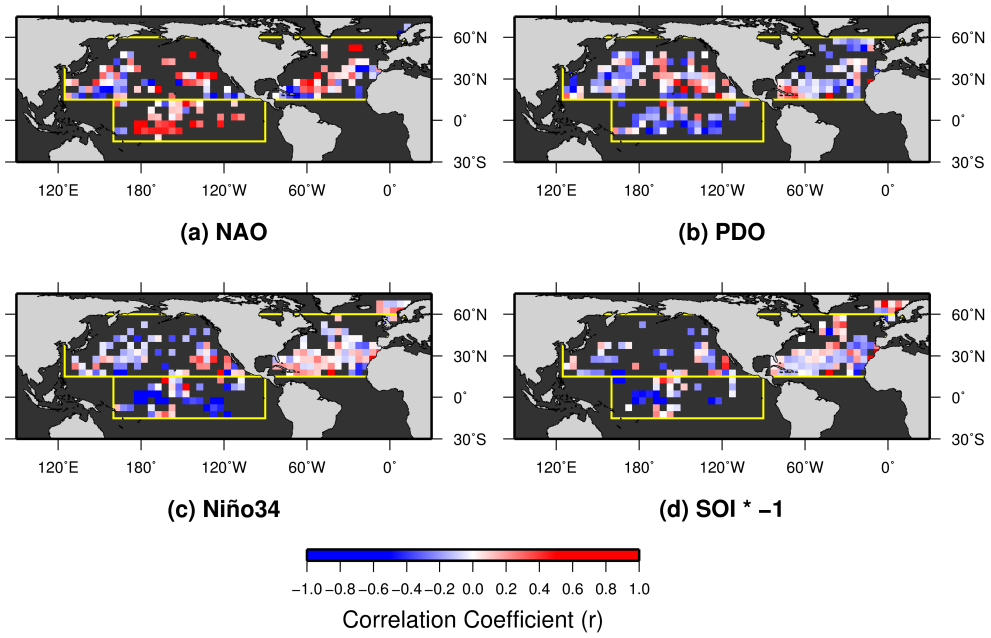

Figure 4.7: Correlations of observed surface pCO_2 anomalies with four climate indices masked to keep only the cells deemed representative of the likely true value of the correlation (see caption of Figure 4.4 (page 89) and the text). The SOI correlations have been inverted for simpler comparison.

f) were compared to the correlations calculated from the complete model output (Figures 4.4 and 4.5 a and b, pages 89 and 90) and to the observed correlations (Figures 4.4 and 4.5 c and d). The grid cell was kept if the signs of the correlations matched across all three datasets. The correlation signs were considered to match if either all three cells were of the same sign or if they were within $\pm$ 0.1 of each other. Approximately 70% of the representative grid cells were in agreement across the four climate indices (Figures 4.4 and 4.5 e and f).

Cells that did not meet these criteria were discarded. The representative cells are marked with black dots in Figures 4.4 and 4.5 e and f (pages 89 and 90). Grid cells were mostly rejected because they had too few observations or because the total time span covered by those observations was too short. However, neither factor was conclusively dominant; some representative cells had few measurements covering a long time period, while others has a large number of measurements within a relatively short time period. The SOCAT correlations were masked so that only the representative cells established from the model data were included (Figure 4.7, page 91).

4.3.7 Calculation of impact on fluxes

The magnitude of the pCO_2 response to changes in climate indices was used to estimate the corresponding change in air-sea CO_2 fluxes. The standard formulation for calculating air-sea fluxes:

$$Flux\ rate = k\ s\ \Delta pCO_2 \tag{4.2}$$

gives the air-sea flux in $mol\ m^{-2}\ yr^{-1}$ where k is the gas transfer velocity, s is the solubility of CO_2 in sea water, and ΔpCO_2 is the difference in partial pressure of CO_2 across the air-sea interface. The total flux of carbon in a given region (in $Pg\ C\ yr^{-1}$) can therefore be calculated by multiplying the flux rate by the area of that region and converting moles into weight of carbon:

$$Flux = k\ s\ \Delta pCO_2\ a\ 12 \times 10^{-15} \tag{4.3}$$

where a is the area of the region in m^2. The change in flux caused by changes in climate indices was calculated by replacing the ΔpCO_2 term in equation 4.2 with the magnitude of the pCO_2 response to changes in indices, i.e. the slope of the linear fit used in calculating the correlation between pCO_2 and the index values:

$$\Delta Flux = k\ s\ m\ a\ 12 \times 10^{-15} \tag{4.4}$$

where m is the magnitude of the response.

Estimates of the gas transfer velocity were calculated using SSTs from the NEMO/ PlankTOM5 model [*Buitenhuis et al.*, 2010] and winds from the NCEP/NCAR reanalysis [*Kalnay et al.*, 1996] with the wind parameterisation of *Wanninkhof* [1992].

4.4 Results and Discussion

4.4.1 Correlations

The strongest correlations between climate variability and pCO_2 anomalies are found in the Equatorial Pacific with the Niño34 (r = -0.47), PDO (r = -0.25) and SOI (r = 0.31) indices (Table 4.2, page 93). A strong negative correlation with Niño34 is expected as

	North Pacific	Equatorial Pacific	North Atlantic
NAO	0.04	0.04	0.09
Niño34	−0.12	**− 0.47**	−0.07
PDO	**−0.25**	**−0.25**	**−0.17**
SOI	0.11	**0.31**	0.04

Table 4.2: Region-wide average correlation coefficients between observed pCO_2 anomalies and climate indices. Bold entries indicate correlations that are statistically significant at the 95% level. The North Pacific and North Atlantic cover each ocean basin between 15°N and 60°N. The Equatorial Pacific encompasses the region 15°S to 15°N.

the index is a measure of changes in SST related to the suppression of upwelling of cold, carbon-rich waters in the eastern Equatorial Pacific during El Niño events, and enhancement of the upwelling during La Niña events [*Feely et al.*, 1987; *Inoue and Sugimura*, 1992]. An increase in SST alone would lead to an increase in pCO_2 levels [*Takahashi et al.*, 1993]. However, the SST effect is overwhelmed by the decrease in upwelled carbon throughout the Equatorial Pacific. The SOI is positively correlated with pCO_2 in the Equatorial Pacific (r = 0.31), which is a direct result of the inverse relationship between the SOI and Niño34 (Table 4.1, page 87). The PDO-pCO_2 relationship is similar to but weaker than that of Niño34 due to its basis in North Pacific. The negative (positive) relationship between Niño34 (SOI) and pCO_2 is consistent with previous studies [*Inoue et al.*, 1996; *Etcheto et al.*, 1999; *Feely et al.*, 1999, 2002, 2006; *Sheu et al.*, 2010]. The strongest relationships between pCO_2 and the Niño34, SOI and PDO indices are restricted to a narrow band between 10°S and 10°N, extending from 175°E to 120°W (Figure 4.7b and c, page 91). This corresponds very closely with the region where the Niño34 index is defined (5°S – 5°N, 170°W – 120°W) [*Trenberth*, 1997]. Correlations outside this narrow region are much weaker, indicating that the suppression of the upwelled waters in the east during El Niño events is responsible for almost all the pCO_2 variability in this region. The positive correlations with the SOI (Figure 4.7d) are not as spatially coherent as with Niño34 and the PDO, but are dominated by a small region of very high correlations between 175°E and 160°W.

There is a small negative correlation between Niño34 and pCO_2 in the North Pacific. Although this is not statistically significant, it does agree with previous studies of this region [*Inoue et al.*, 1987, 1995; *Brix et al.*, 2004]. The PDO-pCO_2 correlations are statistically significant in this region (r = -0.25). The distribution of correlation coefficients

(Figure 4.7b) suggests that there is a west-east dipole between negative and positive correlations in the North Pacific (Figure 4.7c). There is some evidence that such a west-east dipole exists in several drivers of pCO_2 [*Weare et al.*, 1976; *Chierici et al.*, 2006], including SST. An empirical study by *Takahashi et al.* [1993] showed that pCO_2 levels are correlated with changes in SST, changing by approximately 4.23% for each 1°C change in SST. The changes in pCO_2 observed in the North Pacific are consistent with changes in the PDO index and related SST changes [*Weare et al.*, 1976], but the magnitude of the change (-0.8 $\pm$ 0.4 μatm per index unut (μatm iu^{-1})) is much smaller than would be expected if only SST were affecting pCO_2 levels. The pCO_2 changes in response to the SST changes measured by the PDO index are therefore mitigated by PDO-related changes to other biogeochemical processes such as vertical mixing and biological activity [*McKinley et al.*, 2006; *Valsala et al.*, 2011].

The influence of the NAO on pCO_2 in the North Atlantic is debated. Previous observational studies have concluded that a link exists in at least some areas [*Gruber et al.*, 2002; *Corbière et al.*, 2007; *Schuster et al.*, 2009; *McKinley et al.*, 2011], while others have found no such relationship [*Schuster and Watson*, 2007; *Watson et al.*, 2009; *Padin et al.*, 2011]. Model-based studies generally find a link between the NAO and pCO_2 levels caused by increased mixed layer depth and decreasing SST during positive phases of the NAO [*Thomas et al.*, 2008; *Ullman et al.*, 2009; *Tjiputra et al.*, 2012], with largest variability in the North and indications of a dipole between the sub-tropics and sub-polar regions [*Le Quéré et al.*, 2000]. Using the monthly NAO index we found a small but not significant correlation between pCO_2 and the NAO (r = 0.09). Using the six-month (January-June, July-December), seasonal (December-February, June-August) and winter-only (December-February) NAO indices (Figure 4.2, page 86) we found correlations of r = 0.21, r = 0.01, and r = 0.21 respectively. None of the correlations were statistically significant (Table 4.4, page 96). The distribution of correlations using the six-month mean (Figure 4.7a, page 91) shows positive correlations in the central and northern North Atlantic, and negative correlations in the extreme east and west. This agrees with the dipole found by *Le Quéré et al.* [2000] and *Thomas et al.* [2008], but contradicts the spatial variation found by *McKinley et al.* [2011]. None of the correlations of individual grid

cells in this region are statistically significant. These results reinforces the conflicting outcomes found in previous studies, suggesting that there may be a link between the NAO and pCO_2 but a definitive assessment is very difficult either because it is not strong, or because variability at other time scales masks the signal.

pCO_2 in the North Atlantic is most clearly correlated with the PDO, with a small but statistically significant correlation (r = -0.17). The NAO and PDO are negatively correlated (r = -0.24), so any relationship in the North Atlantic with the NAO is likely to be inverted for the PDO. The more robust relationship with the PDO is probably due to the fact that this index has less high-frequency variability and better highlights the interannual variability. While teleconnections between Pacific climate variability and North Atlantic SSTs have been observed via a postulated 'atmospheric bridge' [*Lau and Nath*, 1996], these are usually found to lag by several months [*Klein et al.*, 1999]. Instantaneous relationships such as those found here are occasionally observed, but the underlying mechanism is not established [Wanninkhof, pers. comm.]. The spatial variation of the correlations between the PDO and pCO_2 in the North Atlantic shows some evidence of the dipole observed by *Le Quéré et al.* [2000] (Figure 4.7a and b, page 91), indicating that the PDO correlations may coincide well with NAO correlations in the North Atlantic. This suggests that it may be possible to infer NAO correlations from inverted PDO correlations, although the results from this study are not sufficiently robust to confirm that such an inference is possible.

4.4.2 Magnitude of the response

The strongest response of pCO_2 to changes in the various climate indices corresponds almost exactly with those regions where the correlation is strongest (Table 4.3, page 96, page 96). In the Equatorial Pacific, pCO_2 levels change by -6.6 $\pm$ 1.0 μatm iu^{-1}, though a wide range of responses exists across the Equatorial Pacific region (Figure 4.8c, page 97). Where the correlation is strongest responses can be as high as -18 μatm iu^{-1} (Figure 4.9a, page 98). This magnitude of change in pCO_2 is sufficiently large to neutralise the gradient between atmospheric and oceanic pCO_2 and therefore reduces the Equatorial Pacific's role as a source of atmospheric CO_2 during large El Niño events [*Feely et al.*, 2006]. The pCO_2 response to the SOI is much smaller (3.6 $\pm$ 0.9 μatm iu^{-1}). This shows that the

	North Pacific	Equatorial Pacific	North Atlantic
NAO	0.2 ± 0.4	0.6 ± 1.1	0.7 ± 0.7
Niño34	-0.7 ± 0.4	$\mathbf{-6.6 \pm 1.0}$	-0.7 ± 0.7
PDO	$\mathbf{-0.8 \pm 0.4}$	$\mathbf{-3.1 \pm 1.0}$	$\mathbf{-1.4 \pm 0.6}$
SOI	0.6 ± 0.4	$\mathbf{3.6 \pm 0.9}$	0.3 ± 0.6

Table 4.3: The magnitude of the response of pCO_2 to changes in climate indices (in μatm per index unit) in three ocean regions. Uncertainties are calculated as the RMS of the residuals from the least-squares linear fit used to calculate the correlation. Bold entries indicate the statistical significance of the correlations as seen in Table 4.2 (page 93).

NAO Version	Correlation	P-value
Original (monthly)	0.09	0.29
6-month means	0.21	0.23
DJF-JJA	0.01	0.96
DJF only	0.21	0.46

Table 4.4: Correlation coefficients, with p-values, for pCO_2 anomalies in the North Atlantic compared with four versions of the North Atlantic Oscillation. Lower p-values have greater statistical significance (p = 0.05 represents significance at the 95% level). The 6-month mean is used in the remainder of this study because it has the lowest p-value.

ocean-based index gives a much clearer indication of ENSO effects on pCO_2 than the SOI even though both indices are measures of the same climatic phenomenon.

The response to changes in Niño34 and the PDO in the North Pacific is much smaller than in the Equatorial Pacific, with values of -0.7 ± 0.4 and -0.8 ± 0.4 μatm iu^{-1} respectively. For Niño34, this is because ENSO has a much smaller impact outside the Equatorial Pacific. The PDO-pCO_2 correlation coefficient is identical in both regions (r = -0.25). Thus the difference cannot be attributed to a difference in the relationship, but highlights the real differences in climate impact between the two regions. The response to the PDO in the North Atlantic is also relatively small, with a basin-wide response of -1.4 ± 0.6 μatm iu^{-1}. The response to the NAO is 0.7 ± 0.7 μatm iu^{-1}, although there are small regions with higher responses than are seen for the PDO (Figure 4.9c and d, page 98). Again, this suggests that the mechanisms behind the indices do not include factors that have a strong influence on pCO_2. This is expected since both the PDO and NAO measure changes in aspects of the climate that have only an indirect effect on oceanic processes.

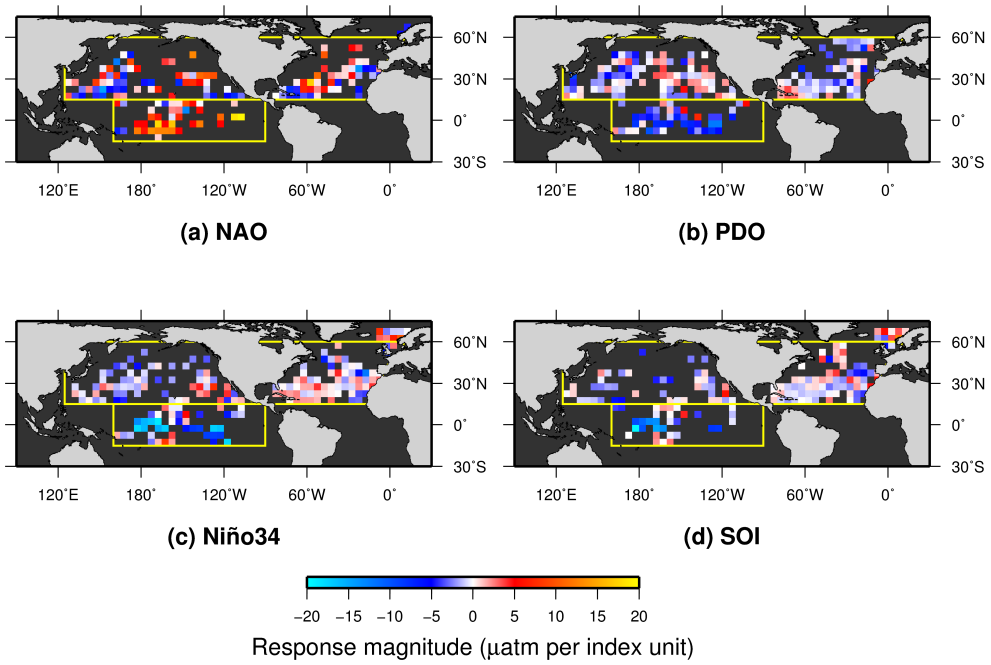

Figure 4.8: Magnitude of the response of pCO_2 levels per unit change in index values. Grid cells are masked as in Figure 4.7 (page 91). The SOI responses have been inverted for simpler comparison.

4.4.3 Changes in flux

The estimated changes in flux are closely related to the magnitude of the pCO_2 response to changes in the climate indices (Table 4.5, page 98). The largest response, the Equatorial Pacific response to Niño34 yields a change in flux of 0.118 ± 0.018 Pg C yr^{-1} iu^{-1} (Petagrams of carbon per year per index unit) over an area of 4.029×10^7 km^2. This equates to a change in flux of 0.21 ± 0.03 Pg C yr^{-1} for the 1992-1994 ENSO event, and 0.40 ± 0.06 Pg C yr^{-1} for the 1997-1998 ENSO event. This rate of change is smaller than early estimates of the effect of ENSO on pCO_2 fluxes (*Feely et al.* [1999] estimated a change in fluxes of 0.4 Pg C yr^{-1} for the 1992-1994 ENSO event), but larger than more recent estimates of 0.17 Pg C yr^{-1} for the 1997-1998 ENSO event [*Park et al.*, 2010]. The change in flux of the Equatorial Pacific related to the SOI and PDO indices is proportionally smaller than the Niño34-related change in line with the smaller magnitude of the pCO_2 response.

None of the indices induces a large change in air-sea fluxes across the North Pacific. The largest change in flux occurs in response to the PDO, with a change of 0.028 ± 0.014 Pg C yr^{-1} iu^{-1} over an area of 4.78×10^7 km^2. The largest change in the North Atlantic

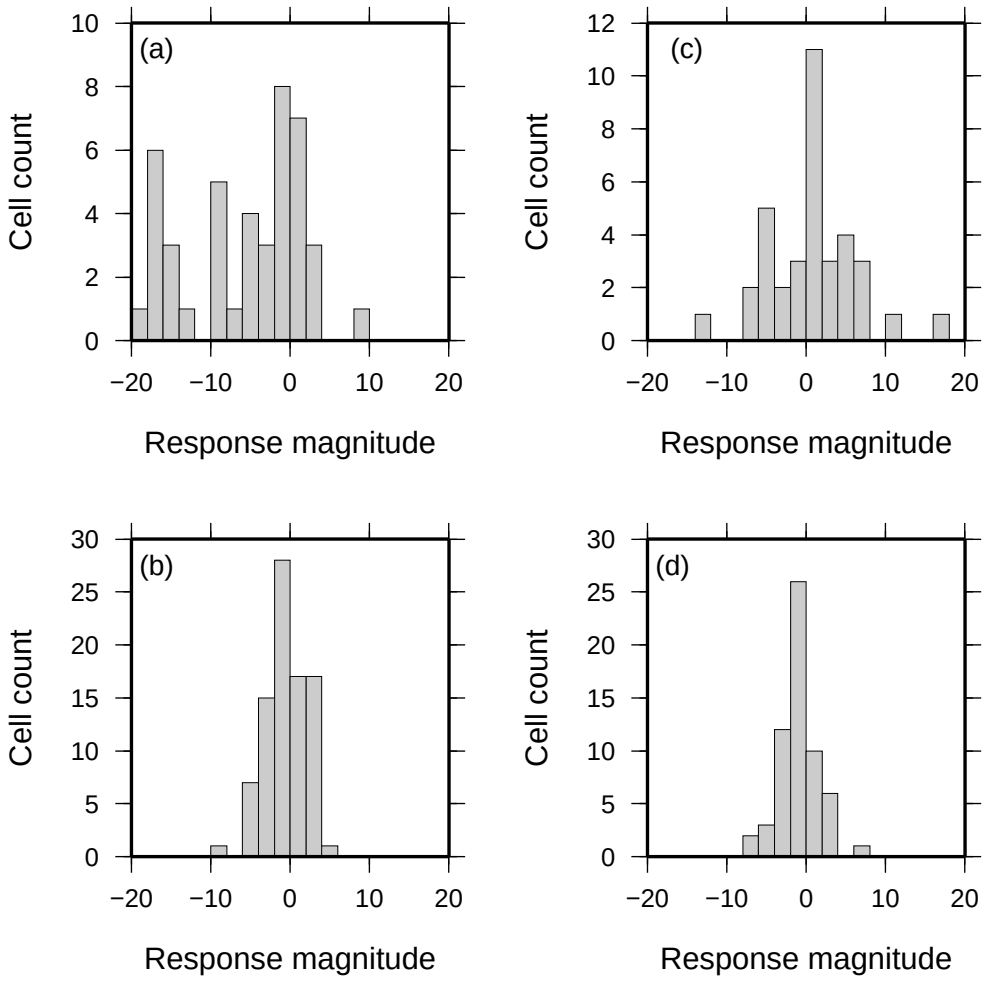

Figure 4.9: Distributions of response magnitudes for (a) Niño34 in the Equatorial Pacific; (b) PDO in the North Pacific; (c) NAO in the North Atlantic; (d) PDO in the North Atlantic.

	North Pacific 4.78×10^7 km^2	**Equatorial Pacific** 4.029×10^7 km^2	**North Atlantic** 3.112×10^7 km^2
NAO	-0.007 $\pm$ 0.014	-0.011 $\pm$ 0.020	-0.029 $\pm$ 0.029
Niño34	0.024 $\pm$ 0.014	**0.118 $\pm$ 0.018**	0.029 $\pm$ 0.029
PDO	**0.028 $\pm$ 0.014**	**0.056 $\pm$ 0.018**	**0.057 $\pm$ 0.024**
SOI	-0.021 $\pm$ 0.001	**-0.065 $\pm$ 0.016**	-0.012 $\pm$ 0.024

Table 4.5: Changes in air-sea fluxes in response to a change of +1 in each climate index across each ocean region in Pg C yr^{-1}. Positive numbers indicate increasing oceanic uptake of CO_2. The uncertainties are based only on the uncertainty in the pCO_2 response to changes in the climate indices; uncertainties in the flux calculation are not considered.

is related to the PDO, where the change in flux is almost identical to that found in the Equatorial Pacific (0.057 $\pm$ 0.024 over 3.112 $\times$ 10^7 km^2 and 0.056 $\pm$ 0.018 Pg C yr^{-1} iu^{-1} over 4.029 $\times$ 10^7 km^2 respectively). This strengthens the evidence for a link between the PDO and the North Atlantic.

4.5 Summary and conclusion

We have analysed the magnitude of the response of surface ocean pCO_2 to climate variability, using model output to identify locations where the data sampling is sufficient. The extension of the analysis beyond simply establishing that relationships exist allows more accurate assessments of that effect climate indices and their underlying mechanisms have on air-sea CO_2 fluxes and the ocean's biogeochemical cycle.

The strongest effect on pCO_2 levels is caused by the influence of ENSO in the Equatorial Pacific. pCO_2 concentrations change by an average of -6.6 $\pm$ 1.0 μatm iu^{-1}, with extremes of up to -18 μatm iu^{-1} in some areas. From this, a change of one index unit in Niño34 leads to a change of 0.118 $\pm$ 0.018 Pg C yr^{-1} over 4.029 $\times$ 10^7 km^2. The PDO and SOI, both closely related to ENSO, also exhibit statistically significant correlations in the Equatorial Pacific. The PDO also has statistically significant correlations with pCO_2 in both the North Pacific and North Atlantic. Although the North Atlantic correlation is small (r = -0.17), it is stronger than that with the NAO (r = 0.07) probably due to the high short-term variability of the latter index. The magnitude of the response to the PDO is also larger than for the NAO, as is the air-sea flux response (0.057 $\pm$ 0.024 and -0.029 $\pm$ 0.029 Pg C yr^{-1} iu^{-1} respectively over 3.112 $\times$ 10^7 km^2). The relationship between the PDO and pCO_2 levels in the North Atlantic is an unexpected result, and deserves further investigation that is beyond the scope of this study.

The relationship between pCO_2 and the NAO in the North Atlantic is not statistically significant, and therefore does not confirm the findings of some previous studies. However, the spatial distribution of correlations across the North Atlantic does agree with that of some modelling studies. The large variability of the NAO means that extracting any potential signal of an NAO-pCO_2 relationship is very difficult. Our results show that the average pCO_2 response to changes in the NAO is 0.7 $\pm$ 0.7 μatm iu^{-1}. This means that even if a link to the NAO is found its influence on pCO_2 levels is very small.

Chapter 5

Summary and Conclusions

5.1 Summary of findings

The central aim of this thesis is to improve our quantitative understanding of surface ocean pCO_2 and air-sea CO_2 flux variability on both temporal and spatial scales. This has been achieved with a study of the temporal and spatial autocorrelation characteristics of the available pCO_2 measurements, and a detailed assessment of the relationship between surface ocean pCO_2 and various large scale modes of climate variability. A new methodology has been developed to interpolate and extrapolate pCO_2 in space and time to assist this effort.

5.1.1 Autocorrelation analysis of pCO_2

The pCO_2 measurements from the LDEO database [*Takahashi and Sutherland*, 2009] were examined using autocorrelation techniques. A temporal autocorrelation function was defined at monthly resolution which was used to distinguish between different magnitudes of the seasonal cyle in pCO_2 at different latitudes. The autocorrelation is robust over multiple years (r = 0.46 after 12 months and r = 0.33 after 48 months) showing that the interannual variability has relatively little effect on the seasonal cycle of pCO_2. Spatial autocorrelation analysis provides very good indications of the pCO_2 variability in most areas of the ocean. Many ocean currents and sub-tropical gyres are visible in the autocorrelation patterns of pCO_2 with autocorrelation lengths of up to 1,400 km. Regions of heterogeneous water masses and variable currents have much shorter autocorrelation lengths (400 km or less), especially in coastal regions. Analysis of the underlying causes of the spatial variability has shown that ocean circulation is the primary driver of spatial variability in pCO_2. Division of the pCO_2 measurements into temperature and non-temperature components using the formulation of *Takahashi et al.* [1993] shows that the driving forces of spatial variability act more directly on the temperature-driven components of pCO_2 than on biological and other influences. Similar analysis of the air-sea CO_2 flux and its component parameters reveals that the spatial scales of variability are approximately half those of pCO_2, primarily due to the high spatial of variability wind speeds.

The assessment of spatial variability produced in this thesis will facilitate the design of future observation networks for pCO_2 [*Sweeney et al.*, 2002; *Lenton et al.*, 2009], by

providing information which will help determine the optimal distance between measurements necessary to capture the full spatial characteristics of the pCO_2 field. This information is particularly timely due to the imminent ability to deploy pCO_2 measurement devices on autonomous floats and gliders [*Johnson et al.*, 2009], providing an opportunity to distribute pCO_2 measurements far more widely than the current ship-board measurement network. The monthly resolution of the temporal autocorrelation function (ACF) is not so useful for informing measurement network design because pCO_2 is known to vary significantly on even sub-diel timescales [*Bates et al.*, 1998; *Dai et al.*, 2009]. Unfortunately there was insufficient data to create a daily resolution ACF which would have been of greater use to inform the design of measurement programmes. Scientists should therefore aim to collect measurements at a minimum of 6- or 12-hour intervals initially to establish the daily and weekly variability of pCO_2. Understanding how this short term variability affects air-sea flux calculations on these time scales will help to determine the requirements of temporal measuring frequency in different regions. Model studies that produce pCO_2 fields at sub-monthly or even sub-daily time resolution may also be useful for this purpose if their results can be validated against observations and the same frequency. Measures of spatial variability will also be useful as a data input to various modelling projects, particularly inverse models, where air-sea CO_2 flux variability can be used as a prior estimate of ocean variability and help reduce the degrees of freedom in the inversions [e.g. *Rödenbeck et al.*, 2003].

5.1.2 Comparison with modes of climate variability

The contribution of climate variability to interannual variability in pCO_2 was assessed using modes of climate variability. The links between pCO_2 and climate indices have been observed previously [e.g. *Inoue et al.*, 1996; *Schuster et al.*, 2009], but the magnitude and spatial extent of the climate-driven variability in pCO_2 is poorly known. Furthermore, conflicting results have been published for the North Atlantic [*Corbière et al.*, 2007; *Schuster et al.*, 2009; *Watson et al.*, 2009; *Padin et al.*, 2011]. The work undertaken in this thesis provides new information on the strength of the pCO_2-climate relationships on interannual timescales based on the most recently available pCO_2 database

[*Pfeil et al.*, submitted]. High sea surface temperatures in the Equatorial Pacific associated with El Niño events lead to reduced pCO_2 concentrations as upwelling of carbon-rich waters is suppressed. pCO_2 levels change by 6.6 $\pm$1.0 μatm per index unit (μatm iu^{-1}), the strongest response found in the study. This equates to changes in the air-sea flux of 0.118 ± 0.018 Pg C yr^{-1} iu^{-1} over an area of 4.029×10^7 km^2. The Pacific Decadal Oscillation (PDO) is the only index that has statistically significant correlations in all three basins examined. The relationship between the PDO and pCO_2 in the North Atlantic is unexpected and worthy of more detailed study in the future. No significant correlation was found with the North Atlantic Oscillation in the North Atlantic, reflecting the high variability of the index and biogeochemical processes in that region. The spatial pattern of correlations in the North Atlantic matches some findings of previous studies in this region, suggesting that a consistent relationship does exist but requires further analysis to quantify it fully.

The direction and magnitude of the pCO_2 response to climate indices provides useful information to help assess the likely impacts on air-sea CO_2 fluxes of changes in interannual variability as a result of climate change. Alterations in air-sea fluxes in response to changes in underlying climate will have a direct effect on atmospheric CO_2 and dampen or accelerate climate change. Anthropogenic activity is likely to alter the behaviour of modes of climate variability [*Trenberth et al.*, 2007], so the resulting change in air-sea fluxes will act as a feedback to human-induced climate change. Whether the feedbacks in response to changes in interannual climate variability will be positive or negative is still uncertain [*Latif and Keenlyside*, 2009].

5.1.3 Interpolation of pCO_2 data

The interpolation method developed to assess the effects of modes of climate variability on pCO_2 is a valuable product in its own right, providing a globally and temporally complete estimate of pCO_2 values for a period of 19 years accompanied by an assessment of uncertainties. Validation of the results has shown that the method is comparably accurate to or more accurate than other published approaches. Tests using model output show that the technique provides useful information even in regions of the ocean with few available measurements. The estimation of uncertainties for each data point is a key part of the data

set. Thus future users can establish their own thresholds when using the interpolated data to fit their requirements.

The globally and temporally complete data set of pCO_2 measurements presented in this thesis will be of use to the scientific community studying the carbon cycle. The data set on its own can be used for any number of analyses of surface pCO_2 trends and variability and the underlying drivers of those changes. Chapter 4 of the thesis presents just one example of such a study. Other examples of potential studies based on the interpolated pCO_2 data are described in Section 5.2.6 (page 111).

5.2 Further work

5.2.1 Improvements to the autocorrelation analysis

There are limited opportunities to extend the temporal autocorrelation analysis performed in Chapter 2 due to the limited amount of available data at a suitable temporal resolution. It may be possible to create separate zonal autocorrelation functions at monthly resolution for different latitudes in the northern hemisphere, although the North Atlantic and North Pacific ACFs would have to be compared to ensure that different ocean basins have similar autocorrelation properties before the zonally averaged ACFs could be considered realistic. A zonal analysis cannot be performed in the southern hemisphere due to a lack of measurements.

The spatial autocorrelation would benefit from in increase in resolution, both in terms of grid size and the resolution of the spatial ACFs (using lag steps of less than 50 km). An increased grid resolution will improve the ability to locate specific dynamic features, providing more detailed information for those designing observation networks or studying mesoscale or smaller features of the oceanic biogeochemical cycle. Increasing the resolution would require some compromise in data coverage because there will be more grid cells without any measurements. A more refined lag resolution would extract more detail from areas of very high variability, most notably around the coasts. Both the increased resolution and smaller lag would dramatically increase the computation requirements of the autocorrelation analysis. The results presented here required approximately 50,000 hours of processing time, so an increase in resolution may not be feasible in the short term.

5.2.2 Developing pCO_2 sampling strategies

One of the major potential uses for the spatial autocorrelation analysis is to help develop sampling strategies for different parts of the world's ocean to ensure sufficient detail of pCO_2 variability is captured to perform the types of analysis desired by the scientific community.

Several sampling strategies have been produced for different ocean regions. Some have focused on temporal sampling requirements [*Garçon et al.*, 1992; *Mémery et al.*, 2002], which cannot be improved by the spatial autocorrelations from this study. Spatial sampling strategies have been developed using a variety of techniques. *Takahashi et al.* [2002] subsampled pCO_2 climatologies to ensure detection of large-scale features. *Sweeney et al.* [2002] calculated the uncertainty in ΔpCO_2 required in a number of regions to ensure that subsequent flux calculations yielded an uncertainty of ± 0.1 Pg C yr^{-1}, and the sampling distances required to achieve that uncertainty. *Lenton et al.* [2006, 2009] utilised Fourier analysis in both space and time to determine the sampling frequency required to capture the majority of the spatial and temporal structure of pCO_2 variability in high latitude seas.

Using the results of this thesis, the maps of decorrelation length presented in Figures 2.4 and 2.5 (pages 35 and 37) can be used as a first estimate of the maximum distance between adjacent samples that maintains some relationship between them, and therefore allows for reasonable interpolation between those samples. However, it is likely that the sampling frequency suggested by these maps is overly conservative. For example, *Lenton et al.* [2006] state that zonal sampling should only be required at 30° ($\sim$2,500 km) intervals in the Southern Ocean, while the zonal autocorrelation map (Figure 2.5a, page 37) shows regions of the of much shorter autocorrelation length.

One possible cause of this discrepancy is that small-scale spatial structures in pCO_2 will lead to short decorrelation lengths. However, if these smaller structures are consistent across a large region, a coarser sampling frequency may still allow a reasonably reliable reconstruction of the pCO_2 field. Spatial Fourier analysis such as that performed by *Lenton et al.* [2006, 2009] has the ability to detect these smaller scale features as variations of a specific wavelength. If that wavelength is a dominant source of spatial variability in pCO_2, the resulting spatial ACF will reflect that wavelength over any other feature, in

turn restricting the decorrelation length. Removing that portion of the ACF ascribed to the dominant Fourier wavelength will allow analysis of the remaining variability. Thus it will be possible to calculate a sampling frequency that accounts for both the 'regular' variaibility and the more noisy variability that cannot easily be detected using Fourier analysis alone.

5.2.3 Extending the analysis of interannual variability

The analysis of relationships between modes of climate variability and pCO_2 was limited to the North Pacific, Equatorial Pacific and North Atlantic where the most pCO_2 measurements are available. It may be possible to extend the analysis of the relationship between pCO_2 and climate indices to some areas of the Southern Hemisphere, particularly in the Southern Ocean south of Australia and the south-western Indian Ocean (Figure 4.3, page 88). Similarly, there are some additional modes of climate variability whose influence could be examined, such as the Southern Annular Mode [*Thompson and Wallace*, 2000] and the Indian Ocean Dipole [*Saji et al.*, 1999].

The strength of correlation and magnitude of the pCO_2 response to climate indices found in Chapter 4 could be used to develop a new metric for assessing model performance. Many physical models are evaluated in terms of their ability to reproduce the large-scale modes of variability examined in this study [e.g. *Oshima and Tanimoto*, 2009; *Stoner et al.*, 2009]. This effort could be extended to biogeochemical models using the results of this study. It would allow for the first time an observation-based assessment of the performance of the models with respect to the relationships between the climate indices and pCO_2.

Further analysis of the relationship between climate indices and pCO_2 may reveal more regarding the underlying mechanisms that link them. The link between ENSO and pCO_2 in the Equatorial Pacific is easily understood through variations in the upwelling of carbon from the upper ocean. The link between the NAO and pCO_2 is much less obvious. Consistent links have been found in regions of the North Atlantic between NAO and SST [*Visbeck et al.*, 2001] and between SST and pCO_2 [*Takahashi et al.*, 1993], but these do not translate directly into a link between the NAO and pCO_2. The influence of the NAO on pCO_2 is a combination of multiple processes, making the relationship difficult to

quantify. Experiments with computer-based models may prove to be the easiest method of separating the influence of the different processes involved to improve understanding the relationship between the NAO and pCO_2.

5.2.4 Improvements to the interpolation method

The interpolation method developed in Chapter 3 works well over most of the ocean, as evidenced by the successful reproduction of model output: 72% of grid cells have an RMS error of $\leq$20 μatm across the 1990-2008 period, and 89% have an RMS error of $\leq$30 μatm. This is the first purely statistical method for interpolating surface pCO_2 that does not rely on other data sources such as sea surface temperature, chlorophyll or mixed layer depth. From the experience gained through this effort, I highlight below several areas for improvements and further developments.

The spatial resolution of the data set presented in this thesis is relatively coarse at 5°x5°, especially when compared to other attempts to interpolate pCO_2 data. Methods based on neural networks frequently produce results at 1°x1° resolution [*Telszewski et al.*, 2009; *Watson et al.*, 2009], focused on data-rich regions of the North Atlantic. The method developed in this thesis could be altered to produce pCO_2 maps of equally high resolution in the North Atlantic and North Pacific where the measurement density is highest. It is also possible that a higher spatial resolution (e.g. 2°x2°) could be achieved across the global ocean without compromising the accuracy of the output.

The spatial interpolation method used in this technique does not currently take into account the presence of land boundaries. Where strips of land are very narrow, for example in central America near Costa Rica and Panama or the southern tip of South America, it is possible that pCO_2 values from unconnected ocean basins will be combined automatically during the spatial interpolation due to the coarse grid size. Moving to a higher resolution grid would eliminate most of the land-spanning grid cells. The interpolation method should further be adapted through the use of a basin mask to take into account continuous land boundaries and prevent spatial interpolation across basins.

The current method of fitting a uniform seasonal cycle to each grid cell is not ideal because it does not take account of any interannual variability caused by climatic phenomena such as ENSO. In years or seasons where no measurements are available the interpolated

values are taken from the fitted mean seasonal cycle which does not include interannual variability and leads to discontinuities in the time series for a given grid cell. This effect is particularly noticeable in the eastern equatorial Pacific, where the variability related to ENSO is larger than the amplitude of the seasonal cycle. A coarser temporal resolution (weekly or monthly) would help to mitigate this issue as there would be fewer values to be interpolated. A method of deriving a variable seasonal cycle would also be beneficial.

The uncertainties assigned to the interpolated pCO_2 are currently unrealistically large. In the model reconstruction test, the assigned uncertainties were far larger than most of the estimated errors (Figures 3.12 and 3.9, pages 68 and 65). The uncertainties are currently based the autocorrelation characteristics of the pCO_2 data. The ACFs typically resemble an exponential decay, so even values separated by short distances can appear to be unrelated and therefore be given a large uncertainty. A more quantitative technique could be based on empirical calculations of the variability of pCO_2 in space and time measured in μatm instead of autocorrelation coefficients. This would give a much more direct estimate of the likely errors introduced through interpolation.

5.2.5 Intercomparison of interpolation techniques

Several techniques have been developed previously to interpolate surface pCO_2 data [e.g. *Boutin et al.*, 1999; *Jamet et al.*, 2007; *Telszewski et al.*, 2009]. To date these have been restricted to specific ocean regions where measurements are relatively abundant, with few attempts to extend the interpolations into regions of sparse data availability over multiple years. The comparison of errors presented in this study shows that results are robust with respect to the chosen interpolation technique where many measurements are available, and errors are of a similar magnitude for all methods (Table 3.2, page 67). While the interpolation presented here is the first interannual study with global coverage, work is under way to also extend some neural network interpolations globally. Some multi-linear regression techniques could also be similarly expanded with relatively little effort [Schuster, pers. comm.]. As these global interpolation methods are developed, it is critical to establish their relative accuracy. One way forward would be to set up an intercomparison project to establish a set of validation metrics. Such an effort would provide a quantitative assessment of the relative value of the available methods, and may help identify

whether the ensemble mean result of the combined interpolations is more accurate than any individual model, as has been frequently observed by modellers in climate science and other fields [*Fraedrich and Leslie*, 1987; *Palmer et al.*, 2004; *Hagedorn et al.*, 2005; *Knutti et al.*, 2010].

5.2.6 Uses of the interpolated data set

The assessment of the magnitude of the effects of major modes of climate variability presented in Chapter 4 represents just one example of the new types of research that can be undertaken with a temporally and spatially complete data set of surface ocean pCO_2 data. This section describes some of the other possible uses for the data set.

First and foremost, a detailed set of pCO_2 data such as the interpolated set created in this thesis can be used to create equally detailed maps of air-sea CO_2 fluxes with corresponding uncertainties. The fluxes could feed directly into calculations of global and regional carbon budgets. They can be used to verify existing understanding of the spatial and temporal variability of air-sea fluxes, and have the potential to provide new insights in regions where direct observational data are lacking. Discrepancies between the fluxes calculated from the interpolated data and global flux estimates could be used to quantify the likely errors in regions of large uncertainty such as the high latitudes (Figure 3.12, page 68) and Equatorial Pacific (Figure 3.9, page 65). The fluxes in these regions could then be adjusted to account for the discrepancies and therefore produce more realistic estimates. It may also be possible to feed the discrepancies back into the interpolation method to act as a constraint of estimates of pCO_2 values in those regions with the highest uncertainties.

A suite of air-sea flux products could be produced using a variety of input sources such as different wind reanalysis products [*Kalnay et al.*, 1996; *Atlas et al.*, 2011; *Dee et al.*, 2011] and SST data sets [*Reynolds and Chelton*, 2010], as well as different formulations of the gas transfer velocity [*Liss and Merlivat*, 1986; *Wanninkhof*, 1992; *Wanninkhof and McGillis*, 1999; *Nightingale et al.*, 2000; *Ho et al.*, 2006; *Sweeney et al.*, 2007]. These could be compared to provide detailed analysis of the effects of using these different products and methods on global and regional air-sea flux estimates, providing a greater

understanding of the importance of data and method selection. Similar tests of variability between methods will also be useful for understanding the role of non-wind factors affecting gas transfer velocities such as surface film effects [*Frew et al.*, 1990; *Tsai and Liu*, 2003], bubbles [*Asher et al.*, 1996; *McNeil and D'Asaro*, 2007] and precipitation [*Ho et al.*, 1997; *Takagaki and Komori*, 2007] as formulations for these factors are developed [*Wanninkhof et al.*, 2009].

The interpolation method presented here assesses long-term changes in surface pCO_2 only in terms of linear trends. However, the results of the interpolation can be used to better understand how these trends are varying over time, by inspecting the pCO_2 anomalies calculated for the comparison with climate indices, as described in Chapter 4. These anomalies were computed by removing a linear trend and seasonal cycle from the interpolated pCO_2. If the trend in pCO_2 is non-linear in any region, having either accelerating or decelerating growth, this will be visible as a trend in the anomalies after the removal of the long term linear trend. If any such trends are detected they will be indicative of changes in the efficiency of the oceanic carbon sink. Variations in global air-sea flux and export production estimates may also provide some insight into long term pCO_2 variability. A more sophisticated curve fitting technique may be able to detect at least some long term variability in the pCO_2 trend and incorporate it into the interpolated data set, particularly in the North Atlantic and North Pacific.

The interpolated pCO_2 data set could also be useful to support the design of observation network. In addition to providing information on data frequency as explained above, the location of the largest uncertainties calculated from the method (Figure 3.12, page 68) could be used to identify regions where more measurements are required. Similarly, detailed assessment of the errors in reproducing model output (Figure 3.9, page 65) compared to the available measurement density could also be useful, provided those errors can be shown to be due to a lack of measurements rather than a weakness in the method itself.

A complete data set of pCO_2 values will also be very useful to the modelling community. Some work has already begun in assimilating the pCO_2 measurements published in the LDEO database [*Takahashi and Sutherland*, 2009] into biogeochemical models [*Valsala and Maksyutov*, 2010]. The interpolated data set from this study can be used as a similar input, and also provide prior estimates for models utilising the atmospheric inverse

method [e.g. *Gurney et al.*, 2002]. The pCO_2 data set and the analysis of spatio-temporal variability presented in this thesis can also be used to validate the results of biogeochemical models [e.g. *Le Quéré et al.*, 2009], and help determine the degree to which the oceans contribute to the interannual variability in atmospheric CO_2.

5.3 Concluding remarks

The techniques and analyses provided in this thesis provide a number of new insights into the spatial and temporal variability of surface ocean pCO_2, and by extension air-sea CO_2 fluxes. The results presented also provide a basis for much future research. The efficacy of all interpolation techniques and other studies of global air-sea CO_2 fluxes will continue to be limited by the availability of direct measurements from the ocean. It is critical that measurements continue to be taken on a regular basis and that the measurement network is expanded into those regions of the southern hemisphere where the largest data gaps persist.

References

Ahrens, C. D. (2007), *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 537 pp., Thomson Brooks/Cole, Belmont, CA.

Aiken, C. M., and M. H. England (2008), Sensitivity of the Present-Day Climate to Freshwater Forcing Associated with Antarctic Sea Ice Loss, *Journal of Climate*, *21*, 3936–3946, doi:10.1175/2007JCLI1901.1.

Allan, R., and T. Ansell (2006), A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004, *Journal of Climate*, *19*, 5816–5842, doi:10.1175/JCLI3937.1.

Antonov, J. I., R. A. Locarnini, T. P. Boyer, A. V. Mishonov, H. E. Garcia, and S. Levitus (2006), *World Ocean Atlas 2005 Volume 2: Salinity*, NOAA Atlas NESDIS 62 , 182 pp., U.S. Government Printing Office, Washington, D. C.

Archer, D. (2005), Fate of fossil fuel CO_2 in geologic time, *Journal of Geophysical Research*, *110*, doi:10.1029/2004JC002625.

Archer, D. E., M. Eby, V. Brovkin, A. Ridgwell, L. Cao, U. Mikolajewicz, K. Caldeira, K. Matsumoto, G. Munhoven, A. Montenegro, and K. S. Tokos (2009), Atmospheric lifetime of fossil fuel carbon dioxide, *Annual Review of Earth and Planetary Sciences*, *37*, 117–134, doi:10.1146/annurev.earth.031208.100206.

Arrhenius, S. (1896), On the influence of carbonic acid in the air upon the temperature of the ground, *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science: 5th Series*, *41*.

Asher, W. E., L. M. Karle, B. J. Higgins, P. J. Farley, and E. C. Monahan (1996), The influence of bubble plumes on air-seawater gas transfer velocities, *Journal of Geophysical Research*, *101*, 12,027–12,041, doi:10.1029/96JC00121.

Assmann, K. M., M. Bentsen, J. Segschneider, and C. Heinze (2010), An isopycnic ocean carbon cycle model, *Geoscientific Model Development*, *3*, 143–167, doi:10.5194/gmd-3-143-2010.

Atlas, R., R. N. Hoffman, J. Ardizzone, S. M. Leidner, J. C. Jusem, D. K. Smith, and D. Gombos (2011), A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications, *Bulletin of the American Meteorological Society*, *92*, 157–174, doi:10.1175/2010BAMS2946.1.

Ayers, J. M., and M. S. Lozier (2012), Unraveling dynamical controls on the North Pacific carbon sink, *Journal of Geophysical Research*, *117*, doi:10.1029/2011JC007368.

Barnes, S. L. (1964), A technique for maximizing details in numerical weather map analysis, *Journal of Applied Meteorology*, *3*, 396–409, doi:10.1175/1520-0450(1964)003⟨0396:ATFMDI⟩2.0.CO;2.

Barnston, A. G., and R. E. Livezey (1987), Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Monthly Weather Review*, *115*, 1083–1126, doi:10.1175/1520-0493(1987)115⟨1083:CSAPOL⟩2.0.CO;2.

Bates, N. R., A. F. Michaels, and A. H. Knap (1996), Seasonal and interannual variability of oceanic carbon dioxide species at the U.S. JGOFS Bermuda Atlantic Time-series Study (BATS) site, *Deep-Sea Research Part II*, *43*, 347–383, doi:10.1016/0967-0645(95)00093-3.

Bates, N. R., T. Takahashi, D. W. Chipman, and A. H. Knap (1998), Variability of pCO_2 on diel to seasonal timescales in the Sargasso Sea near Bermuda, *Journal of Geophysical Research*, *103*, 15,567–15,585, doi:10.1029/98JC00247.

Behrenfeld, M. J., R. T. O'Malley, D. A. Siegel, C. R. McClain, J. L. Sarmiento, G. C. Feldman, A. J. Milligan, P. G. Falkowski, R. M. Letelier, and E. S. Boss (2006), Climate-driven trends in contemporary ocean productivity, *Nature*, *444*, 752–755, doi:10.1038/nature05317.

Bindoff, N. L., J. Willebrand, V. Artale, A. Cazenave, J. M. Gregory, S. Gulev, K. Hanawa, C. Le Quéré, S. Levitus, Y. Nojiri, C. K. Shum, L. D. Talley, and A. S. Unnikrishnan (2007), Observations: Oceanic climate change and sea level, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, pp. 385–432, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Boutin, J., J. Etcheto, Y. Dandonneau, D. C. E. Bakker, R. A. Feely, H. Y. Inoue, M. Ishii, R. D. Ling, P. D. Nightingale, N. Metzl, and R. H. Wanninkhof (1999), Satellite sea surface temperature: A powerful tool for interpreting in situ pCO_2 measurements in the equatorial Pacific Ocean, *Tellus B*, *51*, 490–508, doi:10.1034/j.1600-0889.1999.00025.x.

Briffa, K. R., P. D. Jones, F. H. Schweingruber, and T. J. Osborn (1998), Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years, *Nature*, *393*, 2–7, doi:10.1038/30943.

Brix, H., N. Gruber, and C. D. Keeling (2004), Interannual variability of the upper ocean carbon cycle at station ALOHA near Hawaii, *Global Biogeochemical Cycles*, *18*, doi:10.1029/2004GB002245.

Broecker, W. S., J. R. Ledwell, T. Takahashi, R. F. Weiss, L. Merlivat, L. Memery, T.-H. Peng, B. Jahne, and K. O. Munnich (1986), Isotopic versus micrometeorologic ocean CO_2 fluxes: a serious conflict, *Journal of Geophysical Research*, *91*, 10,517–10.

Broecker, W. S., S. C. Sutherland, and T.-H. Peng (1999), A possible 20th-century slowdown of Southern Ocean Deep Water formation, *Science*, *286*, 1132–1135, doi:10.1126/science.286.5442.1132.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *Journal of Geophysical Research*, *111*, doi:10.1029/2005JD006548.

Buitenhuis, E. T., R. B. Rivkin, S. Sailley, and C. Le Quéré (2010), Biogeochemical fluxes through microzooplankton, *Global Biogeochemical Cycles*, *24*, doi:10.1029/2009GB003601.

Caldeira, K., and P. B. Duffy (2000), The role of the Southern Ocean in uptake and storage of anthropogenic carbon dioxide, *Science*, *287*, 620–622, doi:10.1126/science.287.5453.620.

Canadell, J. G., and P. Ciais (in preparation), Regional carbon budgets: Scope, methods and data, *Biogeosciences*.

Canadell, J. G., C. Le Quéré, M. R. Raupach, C. B. Field, E. T. Buitenhuis, P. Ciais, T. J. Conway, N. P. Gillett, R. A. Houghton, and G. Marland (2007), Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 18,866–18,870, doi:10.1073/pnas.0702737104.

Chierici, M., A. Fransson, and Y. Nojiri (2006), Biogeochemical processes as drivers of surface fCO_2 in contrasting provinces in the subarctic North Pacific Ocean, *Global Biogeochemical Cycles*, *20*, doi:10.1029/2004GB002356.

Chiew, F. H. S., T. C. Piechota, J. A. Dracup, and T. A. Mcmahon (1998), El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting, *Journal of Hydrology*, *204*, 138–149, doi:10.1016/S0022-1694(97)00121-2.

Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang (1994), Evidence for interannual variability of the carbon cycle from the NOAA/CMDL Global Air Sampling Network, *Journal of Geophysical Research*, *99*, 22,831–22,855, doi:10.1029/94JD01951.

Cooper, D. J., A. J. Watson, and R. D. Ling (1998), Variation of pCO_2 along a North Atlantic shipping route (U.K. to the Caribbean): A year of automated observations, *Marine Chemistry*, *60*, 147–164, doi:10.1016/S0304-4203(97)00082-0.

Corbière, A., N. Metzl, G. Reverdin, C. Brunet, and T. Takahashi (2007), Interannual and decadal variability of the oceanic carbon sink in the North Atlantic subpolar gyre, *Tellus B*, *59*, 168–178, doi:10.1111/j.1600-0889.2006.00232.x.

Cosca, C. E., R. A. Feely, J. Boutin, J. Etcheto, M. J. McPhaden, F. P. Chavez, and P. G. Strutton (2003), Seasonal and interannual CO_2 fluxes for the central and eastern equatorial Pacific Ocean as determined from fCO_2-SST relationships, *Journal of Geophysical Research*, *108*, doi:10.1029/2000JC000677.

Cressman, G. P. (1959), An operational objective analysis system, *Monthly Weather Review*, *87*, 367–374, doi:10.1175/1520-0493(1959)087⟨0367:AOOAS⟩2.0.CO;2.

Crowley, T. J. (2000), Causes of climate change over the past 1000 years, *Science*, *289*, 270–277, doi:10.1126/science.289.5477.270.

da Silveira, I. C. A., L. B. de Miranda, and W. S. Brown (1994), On the origins of the North Brazil Current, *Journal of Geophysical Research*, *99*, 22,501–22,512, doi:10.1029/94JC01776.

Dai, M., Z. Lu, W. Zhai, B. Chen, Z. Cao, K. Zhou, W.-J. Cai, and C.-T. A. Chen (2009), Diurnal variations of surface seawater pCO_2 in contrasting coastal environments, *Marine Geology*, *54*, 735–745.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. W. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597, doi:10.1002/qj.828.

Denman, K. L., G. Brasseur, A. Chidthaisong, P. Ciais, P. M. Cox, R. E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P. L. da Silva Dias, S. C. Wofsy, and X. Zhang (2007), Couplings between changes in the climate system and biogeochemistry, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, pp. 499–587, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Dickson, R. R., and J. Brown (1994), The production of North Atlantic Deep Water: Sources, rates, and pathways, *Journal of Geophysical Research*, *99*, 12,319–12,341, doi:10.1029/94JC00530.

Doney, S. C., N. Mahowald, I. Lima, R. A. Feely, F. T. Mackenzie, J.-F. Lamarque, and P. J. Rasch (2007), Impact of anthropogenic atmospheric nitrogen and sulfur deposition on ocean acidification and the inorganic carbon system., *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 14,580–14,585, doi:10.1073/pnas.0702218104.

Doney, S. C., I. D. Lima, R. A. Feely, D. M. Glover, K. Lindsay, N. Mahowald, J. K. Moore, and R. H. Wanninkhof (2009), Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air–sea CO_2 fluxes: Physical climate and atmospheric dust, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 640–655, doi:10.1016/j.dsr2.2008.12.006.

Edson, J. B., A. A. Hinton, K. E. Prada, J. E. Hare, and C. W. Fairall (1998), Direct Covariance Flux Estimates from Mobile Platforms at Sea, *Journal of Atmospheric and Oceanoc Technology*, *15*, 547–562, doi:10.1175/1520-0426(1998)015⟨0547:DCFEFM⟩2.0.CO;2.

Etcheto, J., J. Boutin, Y. Dandonneau, D. C. E. Bakker, R. A. Feely, R. D. Ling, P. D. Nightingale, and R. H. Wanninkhof (1999), Air–sea CO_2 flux variability in the equatorial Pacific Ocean near 100°W, *Tellus B*, *51*, 734–747, doi:10.1034/j.1600-0889.1999.t01-1-00013.x.

Fabry, V. J., B. A. Seibel, R. A. Feely, and J. C. Orr (2008), Impacts of ocean acidification on marine fauna and ecosystem processes, *ICES Journal of Marine Science*, *65*, 414–432, doi:10.1093/icesjms/fsn048.

Falkowski, P. G., R. M. Greene, and R. J. Geider (1992), Physiological limitations on phytoplankton productivity in the ocean, *Oceanography*, *5*, 84–91.

Fangohr, S., and D. K. Woolf (2007), Application of new parameterizations of gas transfer velocity and their impact on regional and global marine CO_2 budgets, *Journal of Marine Systems*, *66*, 195–203, doi:10.1016/j.jmarsys.2006.01.012.

Feely, R. A., R. H. Gammon, B. A. Taft, P. E. Pullen, L. S. Waterman, T. J. Conway, J. F. Gendron, and D. P. Wisegarver (1987), Distribution of chemical tracers in the Eastern Equatorial Pacific during and after the 1982–1983 El Niño/Southern Oscillation event, *Journal of Geophysical Research*, *92*, 6545–6558, doi:10.1029/JC092iC06p06545.

Feely, R. A., R. H. Wanninkhof, C. Goyetj, and D. E. Archer (1997), Variability of CO_2 distributions and sea-air fluxes in the central and eastern equatorial Pacific during the 1991-1994 El Niño, *Deep Sea Research*, *44*, 1851–1867, doi:doi:10.1016/S0967-0645(97)00061-1.

Feely, R. A., R. H. Wanninkhof, T. Takahashi, and P. P. Tans (1999), Influence of El Niño on the equatorial Pacific contribution to atmospheric CO_2 accumulation, *Nature*, *398*, 597–601, doi:10.1038/19273.

Feely, R. A., J. Boutin, C. E. Cosca, Y. Dandonneau, J. Etcheto, H. Y. Inoue, M. Ishii, C. Le Quéré, D. J. Mackey, and M. J. McPhaden (2002), Seasonal and interannual variability of CO_2 in the equatorial Pacific, *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*, 2443–2469, doi:10.1016/S0967-0645(02)00044-9.

Feely, R. A., T. Takahashi, R. H. Wanninkhof, M. J. McPhaden, C. E. Cosca, S. C. Sutherland, and M.-E. Carr (2006), Decadal variability of the air-sea CO_2 fluxes in the equatorial Pacific Ocean, *Journal of Geophysical Research*, *111*, doi:10.1029/2005JC003129.

Flatau, M. K., L. D. Talley, and P. P. Niiler (2003), The North Atlantic Oscillation, Surface Current Velocities, and SST Changes in the Subpolar North Atlantic, *Journal of Climate*, *16*, 2355–2369, doi:10.1175/2787.1.

Fourier, J.-B. J. (1827), On the temperatures of the terrestrial sphere and interplanetary space, *Mémoires de l'Académie Royales des Sciences*, *7*, 569–604.

Fraedrich, K., and L. M. Leslie (1987), Combining predictive schemes in short-term forecasting, *Monthly Weather Review*, *115*, 1640–1644, doi:10.1175/1520-0493(1987)115⟨1640:CPSIST⟩2.0.CO;2.

Fransson, A., M. Chierici, and Y. Nojiri (2009), New insights into the spatial variability of the surface water carbon dioxide in varying sea ice conditions in the Arctic Ocean, *Continental Shelf Research*, *29*, 1317–1328, doi:10.1016/j.csr.2009.03.008.

Frew, N. M., J. C. Goldman, M. R. Dennett, and A. Sherwood Johnson (1990), Impact of phytoplankton-generated surfactants on air-sea gas exchange, *Journal of Geophysical Research*, *95*, 3337–3352, doi:10.1029/JC095iC03p03337.

Friedlingstein, P., P. M. Cox, R. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. C. Doney, M. Eby, I. Y. Fung, G. Bala, J. John, C. D. Jones, F. Joos, T. Kato, M. Kawayima, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. J. Rayner, C. Reick, E. Roeckner, K.-G. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng (2006), Climate–carbon cycle feedback analysis : Results from the C4MIP model intercomparison, *Journal of Climate*, *19*, 3337–3353, doi:10.1175/JCLI3800.1.

Garçon, V. C., F. Thomas, C. S. Wong, and J.-F. Minster (1992), Gaining insight into the seasonal variability of CO_2 at ocean station P using an upper ocean model, *Deep Sea Research Part A*, *39*, 921–938, doi:10.1016/0198-0149(92)90032-O.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, and J. I. Antonov (2006), *World Ocean Atlas 2005 Volume 4: Nutrients (phosphate, nitrate, silicate)*, NOAA Atlas NESDIS 64 , 396 pp., U.S. Government Printing Office, Washington, D. C.

GLOBALVIEW-CO2 (2008), Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

Gloor, M., J. L. Sarmiento, and N. Gruber (2010), What can be learned about carbon cycle climate feedbacks from the CO_2 airborne fraction?, *Atmospheric Chemistry and Physics*, *10*, 7739–7751, doi:10.5194/acp-10-7739-2010.

González-Dávila, M., J. M. Santana-Casiano, M.-J. Rueda, O. Llinás, and E.-F. González-Dávila (2003), Seasonal and interannual variability of sea-surface carbon dioxide species at the European Station for Time Series in the Ocean at the Canary Islands (ESTOC) between 1996 and 2000, *Global Biogeochemical Cycles*, *17*, doi:10.1029/2002GB001993.

Graven, H. D., N. Gruber, R. M. Key, S. Khatiwala, and X. Giraud (submitted), New insights on shallow-to-deep ocean exchange and anthropogenic CO_2 uptake, *Journal of Geophysical Research*.

Gruber, N., C. D. Keeling, and N. R. Bates (2002), Interannual variability in the North Atlantic Ocean carbon sink, *Science*, *298*, 2374–2378, doi:10.1126/science.1077077.

Gurney, K. R., R. M. Law, A. S. Denning, P. J. Rayner, D. Baker, P. Bousquet, L. M. P. Bruhwiler, Y.-H. Chen, P. Ciais, S. Fan, I. Y. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, T. Maki, S. Maksyutov, K. A. Masarie, P. Peylin, M. Prather, B. C. Pak, J. T. Randerson, J. L. Sarmiento, S. Taguchi, T. Takahashi, and C.-W. Yuen (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–30, doi:10.1038/415626a.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept, *Tellus*, *57*, 219–233, doi:10.1111/j.1600-0870.2005.00103.x.

Hansell, D. A., C. A. Carlson, D. J. Repeta, and R. Schlitzer (2009), Dissolved organic matter in the ocean: A controversy stimulates new insights, *Oceanography*, *22*, 202–211, doi:10.5670/oceanog.2009.109.

Hegerl, G. C., F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. Marengo Orsini, N. Nicholls, and J. E. Penner (2007), Understanding and attributing climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, pp. 663–745, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Held, I. M., and B. J. Soden (2000), Water vapour feedback and global warming, *Annual Reviews of Energy and the Environment*, *25*, 441–475, doi:10.1146/annurev.energy.25.1.441.

Heywood, K. J., A. C. Naveira Garabato, and D. P. Stevens (2002), High mixing rates in the abyssal Southern Ocean, *Nature*, *415*, 1011–1014, doi:10.1038/4151011a.

Ho, D. T., L. F. Bliven, R. H. Wanninkhof, and P. Schlosser (1997), The effect of rain on air-water gas exchange, *Tellus B*, *49*, 149–158, doi:10.1034/j.1600-0889.49.issue2.3.x.

Ho, D. T., C. S. Law, M. J. Smith, P. Schlosser, M. Harvey, and P. Hill (2006), Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations, *Geophysical Research Letters*, *33*, doi: 10.1029/2006GL026817.

Hofmann, D. J., J. H. Butler, E. J. Dlugokencky, J. W. Elkins, K. A. Masarie, S. A. Montzka, and P. P. Tans (2006), The role of carbon dioxide in climate forcing from 1979 to 2004: Introduction of the Annual Greenhouse Gas Index, *Tellus B*, *58*, 614–619, doi:10.1111/j.1600-0889.2006.00201.x.

Holzer, M., and F. W. Primeau (2008), The path-density distribution of oceanic surface-to-surface transport, *Journal of Geophysical Research*, *113*, doi:10.1029/2006JC003976.

Houghton, R. A. (2007), Balancing the global carbon budget, *Annual Review of Earth and Planetary Sciences*, *35*, 313–347, doi:10.1146/annurev.earth.35.031306.140057.

Hurrell, J. W. (1995), Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation., *Science*, *269*, 676–9, doi:10.1126/science.269.5224.676.

Imawaki, S., H. Uchida, H. Ichikawa, M. Fukasawa, and S.-I. Umatani (2001), Satellite altimeter south of Japan monitoring the Kuroshio transport south of Japan, *Geophysical Research Letters*, *28*, 17–20, doi:10.1029/2000GL011796.

Indermühle, A., T. F. Stocker, F. Joos, H. Fischer, H. J. Smith, M. Wahlen, B. L. Deck, D. Mastroianni, J. Tschumi, T. Blunier, R. Meyer, and B. Stauffer (1999), Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antarctica, *Nature*, *398*, 121–126, doi:10.1038/18158.

Indermühle, A., E. Monnin, B. Stauffer, T. F. Stocker, and M. Wahlen (2000), Atmospheric CO_2 concentration from 60 to 20 kyr BP from the Taylor Dome Ice Core, Antarctica, *Geophysical Research Letters*, *27*, 735–738, doi:10.1029/1999GL010960.

Ingleby, B., and M. Huddleston (2007), Quality control of ocean temperature and salinity profiles–Historical and real-time data, *Journal of Marine Systems*, *65*, 158–175.

Inoue, H. Y., and Y. Sugimura (1992), Variations and distributions of CO_2 in and over the equatorial Pacific during the period from the 1986/88 El Niño event to the 1988/89 La Niña event, *Tellus B*, *44*, 1–22, doi:10.1034/j.1600-0889.1992.00001.x.

Inoue, H. Y., Y. Sugimura, and K. Fushimi (1987), pCO_2 and $\delta^{13}C$ in the air and surface sea water in the western North Pacific, *Tellus B*, *39*, 228–242, doi:10.1111/j.1600-0889.1987.tb00285.x.

Inoue, H. Y., H. Matsueda, M. Ishii, K. Fushimi, M. Hiroto, I. Asanuma, and Y. Takasugi (1995), Long-term trend of the partial pressure of carbon dioxide (pCO_2) in surface waters of the western North Pacific, 1984–1993, *Tellus B*, *47*, 391–413, doi:10.1034/j.1600-0889.47.issue4.2.x.

Inoue, H. Y., M. Ishii, H. Matsueda, M. Ahoyama, and I. Asanuma (1996), Changes in longitudinal distribution of the partial pressure of CO_2 (pCO_2) in the central and western equatorial Pacific, west of 160°W, *Geophysical Research Letters*, *23*, 1781–1784, doi:10.1029/96GL01674.

Inoue, H. Y., M. Ishii, H. Matsueda, S. Saito, T. Midorikawa, and K. Nemoto (1999), MRI measurements of partial pressure of CO_2 in surface waters of the Pacific during 1968-70: Re-evaluation and comparison of data with those of the 1980s and 1990s, *Tellus B*, *51*, 830–848, doi:10.1034/j.1600-0889.1999.t01-3-00007.x.

Ishii, M., H. Y. Inoue, T. Midorikawa, S. Saito, T. Tokieda, D. Sasano, A. Nakadate, K. Nemoto, N. Metzl, C. S. Wong, and R. A. Feely (2009), Spatial variability and decadal trend of the oceanic CO_2 in the western equatorial Pacific warm/fresh water, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 591–606, doi:10.1016/j.dsr2.2009.01.002.

Jamet, C., C. Moulin, and N. Lefèvre (2007), Estimation of the oceanic pCO_2 in the North Atlantic from VOS lines in-situ measurements: Parameters needed to generate seasonally mean maps, *Annales Geophysicae*, *25*, 2247–2257, doi:10.5194/angeo-25-2247-2007.

Jiang, L.-Q., W.-J. Cai, R. H. Wanninkhof, Y. Wang, and H. Lüger (2008), Air-sea CO_2 fluxes on the U.S. South Atlantic Bight: Spatial and seasonal variability, *Journal of Geophysical Research*, *113*, 1–17, doi:10.1029/2007JC004366.

Jiao, N., G. J. Herndl, D. A. Hansell, R. Benner, G. Kattner, S. W. Wilhelm, D. L. Kirchman, M. G. Weinbauer, T. Luo, F. Chen, and F. Azam (2010), Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nature reviews: Microbiology*, *8*, 593–599, doi:10.1038/nrmicro2386.

Johnson, K. S., W. M. Brelson, E. S. Boss, Z. Chase, H. Claustre, S. R. Emerson, N. Gruber, A. Körtzinger, M. J. Perry, and S. C. Riser (2009), Observing biogeochemical cycles at global scales with profiling floats and gliders, *Oceanography*, *22*, 216–225, doi:10.5670/oceanog.2009.81.

Jones, C. D., and P. M. Cox (2005), On the significance of atmospheric CO_2 growth rate anomalies in 2002–2003, *Geophysical Research Letters*, *32*, 2001–2004, doi:10.1029/2005GL023027.

Joseph, P. V., J. K. Eischeid, and R. J. Pyle (1994), Interannual variability of the onset of the Indian summmer monsoon and its association with atmospheric features, El Niño, and sea surface temperature anomalies, *Journal of Climate*, *7*, 81–105.

Kalnay, E., M. Kanamitsu, R. Kistler, W. D. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. W. Reynolds, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, *77*, 437–471.

Kaplan, A., Y. Kushnir, M. A. Cane, and M. B. Blumenthal (1997), Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures, *Journal of Geophysical Research*, *102*, 835–27.

Karakurt, I., G. Aydin, and K. Aydiner (2012), Sources and mitigation of methane emissions by sectors: A critical review, *Renewable Energy*, *39*, 40–48, doi:10.1016/j.renene.2011.09.006.

Karl, D. M., and R. Lukas (1996), The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation, *Deep Sea Research II*, *43*, 129–156, doi:10.1016/0967-0645(96)00005-7.

Kawabe, M. (1995), Variations of Current Path, Velocity, and Volume Transport of the Kuroshio in Relation with the Large Meander, *Journal of Physical Oceanography*, *25*, 3103–3117, doi:10.1175/1520-0485(1995)025⟨3103:VOCPVA⟩2.0.CO;2.

Keeling, C. D., S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer (2001), Exchanges of Atmospheric CO_2 and $^{13}CO_2$ with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global Aspects, *Tech. rep.*, Scripps Institte of Oceanography, UC San Diego.

Keeling, R. F., R. B. Bacastow, A. E. Bainbridge, C. A. Ekdahl Jr, P. R. Guenther, and L. S. Waterman (1976), Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus*, *28*, 538–551.

Kelly, K. A. (1991), Surface transport, position, and velocity variance from 73° to 46°W, *Journal of Geophysical Research*, *96*, doi:10.1029/91JC01380.

Khalil, M. A. K., R. A. Rasmussen, and M. J. Shearer (2002), Atmospheric nitrous oxide: patterns of global change during recent decades and centuries., *Chemosphere*, *47*, 807–821, doi:10.1016/S0045-6535(01)00297-1.

Kiehl, J. T., and K. E. Trenberth (1997), Earth's annual global mean energy budget, *Bulletin of the American Meteorological Society*, *78*, 197–208, doi:10.1175/1520-0477(1997)078⟨0197:EAGMEB⟩2.0.CO;2.

Klein, S. A., B. J. Soden, and N.-C. Lau (1999), Remote Sea Surface Temperature Variations during ENSO: Evidence for a Tropical Atmospheric Bridge, *Journal of Climate*, *12*, 917–932, doi:10.1175/1520-0442(1999)012⟨0917:RSSTVD⟩2.0.CO;2.

Knorr, W. (2009), Is the airborne fraction of anthropogenic CO_2 emissions increasing?, *Geophysical Research Letters*, *36*, doi:10.1029/2009GL040613.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010), Challenges in Combining Projections from Multiple Climate Models, *Journal of Climate*, *23*, 2739–2758, doi:10.1175/2009JCLI3361.1.

Kondo, F., and O. Tsukamoto (2007), Air-Sea CO_2 Flux by Eddy Covariance Technique in the Indian Ocean, *Journal of Oceanography*, *63*, 449–456.

Körtzinger, A., U. Send, D. W. R. Wallace, J. Karstensen, and M. D. DeGrandpre (2008), Seasonal cycle of O_2 and pCO_2 in the central Labrador Sea: Atmospheric, biological, and physical implications, *Global Biogeochemical Cycles*, *22*, GB1014, doi:10.1029/2007GB003029.

Krasakopoulou, E., S. Rapsomanikis, A. Papadopoulos, and E. Papathanassiou (2009), Partial pressure and air–sea CO_2 flux in the Aegean Sea during February 2006, *Continental Shelf Research*, *29*, 1477–1488, doi:10.1016/j.csr.2009.03.015.

Latif, M., and N. S. Keenlyside (2009), El Niño/Southern Oscillation response to global warming, *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 20,578–20,583, doi:10.1073/pnas.0710860105.

Lau, N.-C., and M. J. Nath (1996), The Role of the "Atmospheric Bridge" in Linking Tropical Pacific ENSO Events to Extratropical SST Anomalies, *Journal of Climate*, *9*, 2036–2057, doi:10.1175/1520-0442(1996)009\%3C2036:TROTBI\%3E2.0.CO;2.

Le Quéré, C., and N. Metzl (2004), Natural processes regulating the uptake of CO_2, in *The Global Carbon Cycle*, edited by C. B. Field and M. R. Raupach, pp. 243–255, Island Press, Washington, D. C.

Le Quéré, C., J. C. Orr, P. Monfray, O. Aumont, and G. Madec (2000), Interannual variability of the oceanic sink of CO_2 from 1979 through 1997, *Global Biogeochemical Cycles*, *14*, 1247–1266, doi:10.1029/1999GB900049.

Le Quéré, C., C. Rödenbeck, E. T. Buitenhuis, T. J. Conway, R. L. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metzl, N. P. Gillett, and M. Heimann (2007), Saturation of the southern ocean CO_2 sink due to recent climate change., *Science*, *316*, 1735–8, doi:10.1126/science.1136188.

Le Quéré, C., M. R. Raupach, J. G. Canadell, G. Marland, L. Bopp, P. Ciais, T. J. Conway, S. C. Doney, R. A. Feely, P. Foster, P. Friedlingstein, K. R. Gurney, R. A. Houghton, J. I. House, C. Huntingford, P. E. Levy, M. R. Lomas, J. Majkut, N. Metzl, J. P. Ometto, G. P. Peters, I. C. Prentice, J. T. Randerson, S. W. Running, J. L. Sarmiento, U. Schuster, S. Sitch, T. Takahashi, N. Viovy, G. R. van der Werf, and F. I. Woodward (2009), Trends in the sources and sinks of carbon dioxide, *Nature Geoscience*, *2*, 831–836, doi:10.1038/ngeo689.

Le Quéré, C., T. Takahashi, E. T. Buitenhuis, C. Rödenbeck, and S. C. Sutherland (2010), Impact of climate change and variability on the global oceanic sink of CO_2, *Global Biogeochemical Cycles*, *24*, 1–10, doi:10.1029/2009GB003599.

Le Texier, H., S. Solomon, and R. R. Garcia (1988), The role of molecular hydrogen and methane oxidation in the water vapour budget of the stratosphere, *Quarterly Journal of the Royal Meteorological Society*, *114*, 281–295.

Lean, J., J. Beer, and R. Bradley (1995), Reconstruction of solar irradiance since 1610: Implications for climate change variability anomalies, *Geophysical Research Letters*, *22*, 3195–3198, doi:10.1029/95GL03093.

Lee, K., R. H. Wanninkhof, T. Takahashi, S. C. Doney, and R. A. Feely (1998), Low interannual variability in recent oceanic uptake of atmospheric carbon dioxide, *Nature*, *396*, 155–159, doi:10.1038/24139.

Lefèvre, N., and A. Taylor (2002), Estimating pCO_2 from sea surface temperatures in the Atlantic gyres, *Deep-Sea Research Part I*, *49*, 539–554, doi:10.1016/S0967-0637(01)00064-4.

Lefèvre, N., A. J. Watson, A. Olsen, A. F. Ríos, F. F. Pérez, and T. Johannessen (2004), A decrease in the sink for atmospheric CO_2 in the North Atlantic, *Geophysical Research Letters*, *31*, 2–5, doi:10.1029/2003GL018957.

Leinweber, A., N. Gruber, H. Frenzel, G. E. Friederich, and F. P. Chavez (2009), Diurnal carbon cycling in the surface ocean and lower atmosphere of Santa Monica Bay, California, *Geophysical Research Letters*, *36*, 3–7, doi:10.1029/2008GL037018.

Lenton, A., and R. J. Matear (2007), Role of the Southern Annular Mode (SAM) in Southern Ocean CO_2 uptake, *Global Biogeochemical Cycles*, *21*, doi:10.1029/2006GB002714.

Lenton, A., R. J. Matear, and B. Tilbrook (2006), Design of an observational strategy for quantifying the Southern Ocean uptake of CO_2, *Global Biogeochemical Cycles*, *20*, 1–11, doi:10.1029/2005GB002620.

Lenton, A., L. Bopp, and R. J. Matear (2009), Strategies for high-latitude northern hemisphere CO_2 sampling now and in the future, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 523–532, doi:10.1016/j.dsr2.2008.12.008.

Levitus, S. (1982), *Climatological atlas of the world ocean*, United States Government Printing.

Li, Z., D. Adamec, T. Takahashi, and S. C. Sutherland (2005), Global autocorrelation scales of the partial pressure of oceanic CO_2, *Journal of Geophysical Research*, *110*, doi:10.1029/2004JC002723.

Limpasuvan, V., and D. L. Hartmann (1999), Eddies and the annular modes of climate variability, *Geophysical Research Letters*, *26*, 3133, doi:10.1029/1999GL010478.

Liss, P. S., and L. Merlivat (1986), Air-sea gas exchange rates: Introduction and synthesis, in *The role of air-sea exchange in geochemical cycling*, pp. 113–127, D. Reidel Publishing Company, Dordrecht, Netherlands.

Litt, E. J., N. J. Hardman-Mountford, J. C. Blackford, G. Mitchelson-Jacob, A. Goodman, G. E. Moore, D. G. Cummings, and M. Butenschon (2010), Biological control of pCO_2 at station L4 in the Western English Channel over 3 years, *Journal of Plankton Research*, *32*, 621–629, doi:10.1093/plankt/fbp133.

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, and H. E. Garcia (2006), *World Ocean Atlas 2005 Volume 1: Temperature*, NOAA Atlas NESDIS 61 , 182 pp., U.S. Government Printing Office, Washington, D. C.

Lohrenz, S. E., and W.-J. Cai (2006), Satellite ocean color assessment of air-sea fluxes of CO_2 in a river-dominated coastal margin, *Geophysical Research Letters*, *33*, doi:10.1029/2005GL023942.

Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill (1995), An estimate of global primary production in the ocean from satellite radiometer data, *Journal of Plankton Research*, *17*, 1245–1271, doi:10.1093/plankt/17.6.1245.

Lüger, H., D. W. R. Wallace, A. Körtzinger, and Y. Nojiri (2004), The pCO_2 variability in the midlatitude North Atlantic Ocean during a full annual cycle, *Global Biogeochemical Cycles*, *18*, doi:10.1029/2003GB002200.

Manabe, S., and R. T. Wetherland (1967), Thermal equilibrium of the atmosphere with a given distribution of relative humidity, *Journal of the Atmospheric Sciences*, *24*, 241–258.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis (1997), A Pacific Interdecadal Climate Oscillation with impacts on salmon production, *Bulletin of the American Meteorological Society*, *78*, 1069–1079, doi:10.1175/1520-0477(1997)078⟨1069:APICOW⟩2.0.CO;2.

Martin, J. H., S. E. Fitzwater, and R. M. Gordon (1990), Iron deficiency limits phytoplankton growth in Antarctic waters, *Global Biogeochemical Cycles*, *4*, 5–12, doi:10.1029/GB004i001p00005.

Masarie, K. A., and P. P. Tans (1995), Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *Journal of Geophysical Research*, *100*, 11,593–11,610, doi:10.1029/95JD00859.

McCartney, M. S., and L. D. Talley (1984), Warm-to-cold water conversion in the northern North Atlantic Ocean, *Journal of Physical Oceanography*, *14*, 922–935, doi:10.1175/1520-0485(1984)014⟨0922:WTCWCI⟩2.0.CO;2.

McGillis, W. R., J. B. Edson, J. E. Hare, and C. W. Fairall (2001), Direct covariance air-sea CO_2 fluxes, *Journal of Geophysical Research*, *106*, 16,729–16,745, doi:10.1029/2000JC000506.

McKinley, G. A., T. Takahashi, E. T. Buitenhuis, F. Chai, J. R. Christian, S. C. Doney, M.-S. Jiang, K. Lindsay, J. K. Moore, C. Le Quéré, I. D. Lima, R. Murtugudde, L. Shi, and P. Wetzel (2006), North Pacific carbon cycle response to climate variability on seasonal to decadal timescales, *Journal of Geophysical Research*, *111*, doi:10.1029/2005JC003173.

McKinley, G. A., A. R. Fay, T. Takahashi, and N. Metzl (2011), Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales, *Nature Geoscience*, *4*, 606–610, doi:10.1038/ngeo1193.

McNeil, C., and E. D'Asaro (2007), Parameterization of air–sea gas fluxes at extreme wind speeds, *Journal of Marine Systems*, *66*, 110–121, doi:10.1016/j.jmarsys.2006.05.013.

Meehl, G. A., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. G. Watterson, A. J. Weaver, and Z.-C. Zhai (2007), Global climate projections, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, pp. 747–845, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Mémery, L., M. Lévy, S. Vérant, and L. Merlivat (2002), The relevant time scales in estimating the air-sea CO_2 exchange in a mid-latitude region, *Deep Sea Research Part II*, *49*, 2067–2092, doi:10.1016/S0967-0645(02)00028-0.

Metzl, N. (2009), Decadal increase of oceanic carbon dioxide in Southern Indian Ocean surface waters (1991–2007), *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 607–619, doi:10.1016/j.dsr2.2008.12.007.

Metzl, N., C. Brunet, A. Jabaud-Jan, A. Poisson, and B. Schauer (2006), Summer and winter air-sea CO_2 fluxes in the Southern Ocean, *Deep Sea Research Part I: Oceanographic Research Papers*, *53*, 1548–1563, doi:10.1016/j.dsr.2006.07.006.

Michaels, A. F., and A. H. Knap (1996), Overview of the U.S. JGOFS Bermuda Atlantic Time-series Study and the Hydrostation S program, *Deep Sea Research Part II: Topical Studies in Oceanography*, *43*, 157–198, doi:10.1016/0967-0645(96)00004-5.

Midorikawa, T., K. Ogawa, K. Nemoto, H. Kamiya, T. Umeda, N. Hiraishi, A. Wada, and M. Ishii (2003), Interannual variations of net community production and air–sea CO_2 flux from winter to spring in the western subarctic North Pacific, *Tellus*, *55*, 466–477, doi:10.1034/j.1600-0889.2003.00009.x.

Midorikawa, T., M. Ishii, K. Nemoto, H. Kamiya, A. Nakadate, S. Masuda, H. Matsueda, T. Nakano, and H. Y. Inoue (2006), Interannual variability of winter oceanic CO_2 and air-sea CO_2 flux in the western North Pacific for 2 decades, *Journal of Geophysical Research*, *111*, doi:10.1029/2005JC003095.

Mikaloff Fletcher, S. E., N. Gruber, A. R. Jacobson, S. C. Doney, S. Dutkiewicz, M. Gerber, M. J. Follows, F. Joos, K. Lindsay, D. Menemenlis, A. Mouchet, S. A. Müller, and

J. L. Sarmiento (2006), Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean, *Global Biogeochemical Cycles*, *20*, doi:10.1029/2005GB002530.

Minnett, P. J. (2001), Satellite remote sensing of sea surface temperatures, in *Encyclopedia of Ocean Sciences*, edited by J. H. Steele, S. A. Thorpe, and K. K. Turekian, pp. 91–102, Elsevier Science.

Morales, C., and C. Lange (2004), Oceanographic studies in the Humboldt current system off Chile: An introduction, *Deep Sea Research Part II: Topical Studies in Oceanography*, *51*, 2345–2348, doi:10.1016/j.dsr2.2004.08.008.

Moran, P. A. P. (1950), Notes on continuous stochastic phenomena, *Biometrika*, *37*, 17–23.

Nemoto, K., T. Midorikawa, A. Wada, K. Ogawa, S. Takatini, H. Kimoto, M. Ishii, and H. Y. Inoue (2009), Continuous observations of atmospheric and oceanic CO_2 using a moored buoy in the East China Sea: Variations during the passage of typhoons, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 542–553, doi:10.1016/j.dsr2.2008.12.015.

Newman, M., G. P. Compo, and M. A. Alexander (2003), ENSO-Forced Variability of the Pacific Decadal Oscillation, *Journal of Climate*, *16*, 3853–3857, doi:10.1175/1520-0442(2003)016⟨3853:EVOTPD⟩2.0.CO;2.

Nightingale, P. D., G. Malin, C. S. Law, A. J. Watson, P. S. Liss, M. I. Liddicoat, J. Boutin, and R. C. Upstill-Goddard (2000), In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, *Global Biogeochemical Cycles Cycles*, *14*, 373–387.

Olsen, A., A. M. Omar, A. C. Stuart-Menteth, and J. A. Triñanes (2004), Diurnal variations of surface ocean pCO_2 and sea-air CO_2 flux evaluated using remotely sensed data, *Geophysical Research Letters*, *31*, 2–5, doi:10.1029/2004GL020583.

Olsen, A., K. R. Brown, M. Chierici, T. Johannessen, and C. Neill (2008), Sea-surface CO_2 fugacity in the subpolar North Atlantic, *Biogeosciences*, *5*, 535–547, doi:10.5194/bg-5-535-2008.

Ono, T., T. Saino, N. Kurita, and K. Sasaki (2004), Basin-scale extrapolation of shipboard pCO_2 data by using satellite SST and Chla, *International Journal of Remote Sensing*, *25*, 3803–3815, doi:10.1080/01431160310001657515.

Oshima, K., and Y. Tanimoto (2009), An evaluation of reproducibility of the Pacific Decadal Oscillation in the CMIP3 simulations, *Journal of the Meteorological Society of Japan*, *87*, 755–770, doi:10.2151/jmsj.87.755.

Padin, X. A., M. Vázquez-Rodríguez, M. Castaño, A. Velo, F. Alonso-Pérez, J. Gago, M. Gilcoto, M. Álvarez, P. C. Pardo, M. de la Paz, A. F. Ríos, and F. F. Pérez (2010), Air-Sea CO_2 fluxes in the Atlantic as measured during boreal spring and autumn, *Biogeosciences*, *7*, 1587–1606, doi:10.5194/bg-7-1587-2010.

Padin, X. A., C. G. Castro, A. F. Ríos, and F. F. Pérez (2011), Oceanic CO_2 uptake and biogeochemical variability during the formation of the Eastern North Atlantic Central Water under two contrasting NAO scenarios, *Journal of Marine Systems*, *84*, 96–105, doi:10.1016/j.jmarsys.2010.10.002.

Palmer, T. N., A. Alessandri, U. Anderson, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Díez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. S. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J.-M. Terres, and M. C. Thomson (2004), Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER), *Bulletin of the American Meteorological Society*, *85*, 853–872, doi:10.1175/BAMS-85-6-853.

Park, G.-H., K. Lee, R. H. Wanninkhof, and R. A. Feely (2006), Empirical temperature-based estimates of variability in the oceanic uptake of CO_2 over the past 2 decades, *Journal of Geophysical Research*, *111*, C07S07, doi:10.1029/2005JC003090.

Park, G.-H., R. H. Wanninkhof, S. C. Doney, T. Takahashi, K. Lee, R. A. Feely, C. L. Sabine, J. A. Triñanes, and I. D. Lima (2010), Variability of global net sea-air CO_2 fluxes over the last three decades using empirical relationships, *Tellus B*, *62*, 352–368, doi:10.1111/j.1600-0889.2010.00498.x.

Peters, G. P., G. Marland, C. Le Quéré, T. Boden, J. G. Canadell, and M. R. Raupach (2012), Rapid growth in CO_2 emissions after the 2008–2009 global financial crisis, *Nature Climate Change*, *2*, 2–4, doi:10.1038/nclimate1332.

Pfeil, B., A. Olsen, D. C. E. Bakker, S. Hankin, H. Koyuk, A. Kozyr, J. Malczyk, A. Manke, N. Metzl, C. L. Sabine, J. Akl, S. R. Alin, R. G. J. Bellerby, A. Borges, J. Boutin, P. J. Brown, W.-J. Cai, P. F. Chavez, A. Chen, C. Cosca, A. J. Fassbender, R. A. Feely, M. González-Dávila, C. Goyet, N. Hardman-Mountford, C. Heinze, M. Hood, M. Hoppema, C. W. Hunt, D. Hydes, M. Ishii, T. Johannessen, S. D. Jones, R. M. Key, A. Körtzinger, P. Landschützer, S. K. Lauvset, N. Lefèvre, A. Lenton, A. Lourantou, L. Merlivat, T. Midorikawa, L. Mintrop, C. Miyazaki, A. Murata, A. Nakadate, Y. Nakano, S. Nakaoka, Y. Nojiri, A. M. Omar, X. A. Padin, G.-H. Park, K. Paterson, F. F. Perez, D. Pierrot, A. Poisson, A. F. Rìos, J. M. Santana-Casiano, J. Salisbury, V. V. S. S. Sarma, R. Schlitzer, B. Schneider, U. Schuster, R. Sieger, I. Skjelvan, T. Steinhoff, T. Suzuki, T. Takahashi, K. Tedesco, M. Telszewski, H. Thomas, B. Tilbrook, J. Tjiputra, D. Vandemark, T. Veness, R. Wanninkhof, A. J. Watson, R. Weiss, C. S. Wong, and H. Yoshikawa-Inoue (submitted), A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT), *Earth System Science Data*, doi:10.5194/essdd-5-1-2012.

Philipona, R., B. Dürr, A. Ohmura, and C. Ruckstuhl (2005), Anthropogenic greenhouse forcing and strong water vapor feedback increase temperature in Europe, *Geophysical Research Letters*, *32*, doi:10.1029/2005GL023624.

Pierrot, D., C. Neill, K. F. Sullivan, R. Castle, R. H. Wanninkhof, H. Lüger, T. Johannessen, A. Olsen, R. A. Feely, and C. E. Cosca (2009), Recommendations for autonomous underway pCO_2 measuring systems and data-reduction routines, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 512–522, doi:10.1016/j.dsr2.2008.12.005.

Primeau, F. W., and M. Holzer (2006), The ocean's memory of the atmosphere: Residence-time and ventilation-rate, *Journal of Physical Oceanography*, *36*, 1439–1456, doi:10.1175/JPO2919.1.

Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G. Y. Shi, and S. Solomon (2001), Radiative forcing of climate change, in *Climate Change 2001: The Scientific Basis*, edited by J. T. Houghton, Y. Ding, D. J. Griggs,

M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, pp. 349–416, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Rasmusson, E. M., and J. M. Wallace (1983), Meteorological Aspects of the El Niño / Southern Oscillation, *Science*, *222*, 1195–1202.

Raupach, M. R., J. G. Canadell, and C. Le Quéré (2008), Anthropogenic and biophysical contributions to increasing atmospheric CO_2 growth rate and airborne fraction, *Biogeosciences*, *5*, 1601–1613, doi:10.5194/bg-5-1601-2008.

Raven, J., K. Caldeira, H. Elderfield, O. Hoegh-Guldberg, P. S. Liss, U. Riebesell, J. Shepard, C. Turley, and A. J. Watson (2005), *Ocean acidification due to increasing atmospheric carbon dioxide*, 60 pp., The Clyvedon Press Ltd., Cardiff, UK.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research*, *108*, doi:10.1029/2002JD002670.

Reynolds, R. W., and D. B. Chelton (2010), Comparisons of daily sea surface temperature analyses for 2007–08, *Journal of Climate*, *23*, 3545–3562, doi:10.1175/2010JCLI3294.1.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang (2002), An Improved In Situ and Satellite SST Analysis for Climate, *Journal of Climate*, *15*, 1609–1625, doi:10.1175/1520-0442(2002)015⟨1609:AIISAS⟩2.0.CO;2.

Richardson, P. L. (2005), Caribbean Current and eddies as observed by surface drifters, *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*, 429–463, doi:10.1016/j.dsr2.2004.11.001.

Röckmann, T., J. U. Grooß, and R. Müller (2004), The impact of anthropogenic chlorine emissions, stratospheric ozone change and chemical feedbacks on stratospheric water, *Atmospheric Chemistry and Physics*, *4*, 693–699, doi:10.5194/acp-4-693-2004.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmospheric Chemistry and Physics Discussions*, *3*, 2575–2659, doi:10.5194/acpd-3-2575-2003.

Roe, G. H., and M. B. Baker (2007), Why is climate sensitivity so unpredictable?, *Science*, *318*, 629–32, doi:10.1126/science.1144735.

Ropelewski, C. F., and P. D. Jones (1987), An extension of the Tahiti–Darwin Southern Oscillation Index, *Monthly Weather Review*, *115*, 2161–2165, doi:10.1175/1520-0493(1987)115⟨2161:AEOTTS⟩2.0.CO;2.

Sabine, C. L., R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. H. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tilbrook, F. J. Millero, T.-H. Peng, A. Kozyr, T. Ono, and A. F. Rios (2004), The oceanic sink for anthropogenic CO_2, *Science*, *305*, 367–71, doi:10.1126/science.1097403.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata (1999), A dipole mode in the tropical Indian Ocean, *Nature*, *401*, 360–3, doi:10.1038/43854.

Santana-Casiano, J. M., M. González-Dávila, M.-J. Rueda, O. Llinás, and E.-F. González-Dávila (2007), The interannual variability of oceanic CO_2 parameters in the northeast Atlantic subtropical gyre at the ESTOC site, *Global Biogeochemical Cycles*, *21*, doi: 10.1029/2006GB002788.

Santana-Casiano, J. M., M. González-Dávila, and I. R. Ucha (2009), Carbon dioxide fluxes in the Benguela upwelling system during winter and spring: A comparison between 2005 and 2006, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 533–541, doi:10.1016/j.dsr2.2008.12.010.

Sarma, V. V. S. S. (2003), Monthly variability in surface pCO_2 and net air-sea CO_2 flux in the Arabian Sea, *Journal of Geophysical Research*, *108*, 3255, doi:10.1029/2001JC001062.

Schmitz, W. J. (1995), On the interbasin-scale thermohaline circulation, *Reviews of Geophysics*, pp. 151–173, doi:10.1029/95RG00879.

Schmitz, W. J. (1996), On the world ocean circulation. Volume I, some global features/North Atlantic circulation, *Tech. rep.*, Woods Hole Oceanographic Institute.

Schneider, T., P. A. O'Gorman, and X. J. Levine (2010), Water vapor and the dynamics of climate changes, *Reviews of Geophysics*, *48*, doi:10.1029/2009RG000302.

Schuster, U., and A. J. Watson (2007), A variable and decreasing sink for atmospheric CO_2 in the North Atlantic, *Journal of Geophysical Research*, *112*, doi: 10.1029/2006JC003941.

Schuster, U., A. J. Watson, N. R. Bates, A. Corbière, M. González-Dávila, N. Metzl, D. Pierrot, and J. M. Santana-Casiano (2009), Trends in North Atlantic sea-surface fCO_2 from 1990 to 2006, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 620–629, doi:10.1016/j.dsr2.2008.12.011.

Seinfeld, J. H. (2011), Insights on global warming, *Envrironmental and Energy Engineering*, *57*, 3259–3284, doi:10.1002/aic.12780.

Shadwick, E. H., H. Thomas, K. Azetsu-Scott, B. J. W. Greenan, E. Head, and E. Horne (2011), Seasonal variability of dissolved inorganic carbon and surface water pCO_2 in the Scotian Shelf region of the Northwestern Atlantic, *Marine Chemistry*, *124*, 23–37, doi:10.1016/j.marchem.2010.11.004.

Sheu, D. D., W.-C. Chou, C.-L. Wei, W.-P. Hou, G. T. F. Wong, and C.-W. Hsu (2010), Influence of El Niño on the sea-to-air CO_2 flux at the SEATS time-series site, northern South China Sea, *Journal of Geophysical Research*, *115*, doi:10.1029/2009JC006013.

Shim, J., D. Kim, Y. C. Kang, J. H. Lee, S.-T. Jang, and C.-H. Kim (2007), Seasonal variations in pCO_2 and its controlling factors in surface seawater of the northern East China Sea, *Continental Shelf Research*, *27*, 2623–2636, doi:10.1016/j.csr.2007.07.005.

Simmons, A. J., S. M. Uppala, D. P. Dee, and S. Kobayashi (2007), ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF Newsletter*, *111*, 25–35.

Smith, S. D., and E. P. Jones (1985), Evidence for Wind-Pumping of Air-Sea Gas Exchange Based on Direct Measurements of CO_2 Fluxes, *Journal of Geophysical Research*, *90*, 869–875, doi:10.1029/JC090iC01p00869.

Stoner, A. M. K., K. Hayhoe, and D. J. Wuebbles (2009), Assessing General Circulation Model simulations of atmospheric teleconnection patterns, *Journal of Climate*, *22*, 4348–4372, doi:10.1175/2009JCLI2577.1.

Sweeney, C., T. Takahashi, and A. Gnanadesikan (2002), Spatial and temporal variability of surface water pCO_2 and sampling strategies, in *A Large-Scale CO_2 Observing Plan: In Situ Oceans and Atmosphere (LSCOP)*, pp. 155–176, National Technical Information Service, Springfield, VA.

Sweeney, C., M. Gloor, A. R. Jacobson, R. M. Key, G. A. McKinley, J. L. Sarmiento, and R. H. Wanninkhof (2007), Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements, *Global Biogeochemical Cycles*, *21*, doi:10.1029/2006GB002784.

Taft, B. A., A. R. Robinson, and W. J. Schmitz (1973), Current path and bottom velocity of the Kuroshio, *Journal of Physical Oceanography*, *3*, 347–350.

Takagaki, N., and S. Komori (2007), Effects of rainfall on mass transfer across the air-water interface, *Journal of Geophysical Research*, *112*, 1–11, doi:10.1029/2006JC003752.

Takahashi, T., and S. C. Sutherland (2009), Global ocean surface water partial pressure of CO_2 database: Measurements performed during 1968-2006 (Version 1.0), *ORNL/CDIAC-152, NDP-088. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN*, *37831*, 20.

Takahashi, T., J. Olafsson, J. G. Goddard, D. W. Chipman, and S. C. Sutherland (1993), Seasonal variation of CO_2 and nutrients in the high-latitude surface oceans: a comparative study, *Global Biogeochemical Cycles*, *7*, 843–878, doi:10.1029/93GB02263.

Takahashi, T., R. A. Feely, R. F. Weiss, R. H. Wanninkhof, D. W. Chipman, S. C. Sutherland, and T. T. Takahashi (1997), Global air-sea flux of CO_2: an estimate based on measurements of sea-air pCO_2 difference., *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 8292–9.

Takahashi, T., S. C. Sutherland, C. Sweeney, A. Poisson, N. Metzl, B. Tilbrook, N. R. Bates, R. H. Wanninkhof, R. A. Feely, C. L. Sabine, J. Olafsson, and Y. Nojiri (2002), Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*, 1601–1622, doi:10.1016/S0967-0645(02)00003-6.

Takahashi, T., S. C. Sutherland, R. A. Feely, and C. E. Cosca (2003), Decadal variation of the surface water pCO_2 in the western and central equatorial Pacific., *Science*, *302*, 852–6, doi:10.1126/science.1088570.

Takahashi, T., S. C. Sutherland, R. A. Feely, and R. H. Wanninkhof (2006), Decadal change of the surface water pCO_2 in the North Pacific: A synthesis of 35 years of observations, *Journal of Geophysical Research*, *111*, doi:10.1029/2005JC003074.

Takahashi, T., S. C. Sutherland, R. H. Wanninkhof, C. Sweeney, R. A. Feely, D. W. Chipman, B. Hales, G. E. Friederich, F. P. Chavez, C. L. Sabine, A. J. Watson, D. C. E. Bakker, U. Schuster, N. Metzl, H. Y. Inoue, M. Ishii, T. Midorikawa, Y. Nojiri, A. Körtzinger, T. Steinhoff, M. Hoppema, J. Olafsson, T. S. Arnarson, B. Tilbrook, T. Johannessen, A. Olsen, R. G. J. Bellerby, C. S. Wong, B. Delille, N. R. Bates, and H. J. W. de Baar (2009), Climatological mean and decadal change in surface ocean

pCO_2, and net sea–air CO_2 flux over the global oceans, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*, 554–577, doi:10.1016/j.dsr2.2008.12.009.

Takamura, T. R., H. Y. Inoue, T. Midorikawa, M. Ishii, and Y. Nojiri (2010), Seasonal and inter-annual variations in pCO_2^{sea} and air-sea CO_2 fluxes in mid-latitudes of the western and eastern North Pacific during 1999-2006: Recent results utilizing Voluntary Observation Ships, *Journal of the Meteorological Society of Japan*, *88*, 883–898, doi: 10.2151/jmsj.2010-602.

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

Thomas, H., A. E. Friederike Prowe, I. D. Lima, S. C. Doney, R. H. Wanninkhof, R. J. Greatbatch, U. Schuster, and A. Corbière (2008), Changes in the North Atlantic Oscillation influence CO_2 uptake in the North Atlantic over the past 2 decades, *Global Biogeochemical Cycles*, *22*, doi:10.1029/2007GB003167.

Thompson, D. W. J., and J. M. Wallace (2000), Annular modes in the extratropical circulation. Part I: Month-to-month variability, *Journal of Climate*, *13*, 1000–1016, doi: 10.1175/1520-0442(2000)013⟨1000:AMITEC⟩2.0.CO;2.

Tjiputra, J. F., K. M. Assmann, and C. Heinze (2010), Anthropogenic carbon dynamics in the changing ocean, *Ocean Science*, *6*, 605–614, doi:10.5194/os-6-605-2010.

Tjiputra, J. F., A. Olsen, K. M. Assmann, B. Pfeil, and C. Heinze (2012), A model study of the seasonal and long–term North Atlantic surface pCO_2 variability, *Biogeosciences*, *9*, 907–923, doi:10.5194/bg-9-907-2012.

Tortell, P. D., and M. C. Long (2009), Spatial and temporal variability of biogenic gases during the Southern Ocean spring bloom, *Geophysical Research Letters*, *36*, doi:10. 1029/2008GL035819.

Trenberth, K. E. (1997), The definition of El Niño, *Bulletin of the American Meteorological Society*, *78*, 2771–2777, doi:10.1175/1520-0477(1997)078⟨2771:TDOENO⟩2. 0.CO;2.

Trenberth, K. E., P. D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. E. Parker, F. Rahimzadeh, J. A. Renwick, M. Rusticucci, B. Soden, and P. Zhai (2007), Observations: Surface and atmospheric climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, pp. 235–336, Cambridge University Press, Cambridge, United Kingdom and New York, N.Y., USA.

Tsai, W.-T., and K.-K. Liu (2003), An assessment of the effect of sea surface surfactant on global atmosphere-ocean CO_2 flux, *Journal of Geophysical Research*, *108*, doi:10. 1029/2000JC000740.

Tyndall, J. (1861), On the absorption and radiation of heat by gases and vapours, and on the physical connexion of radiation, absorption, and conduction, *Philosophical Magazine: Series 4*, *22*, 169–94, 273–85.

Ullman, D. J., G. A. McKinley, V. Bennington, and S. Dutkiewicz (2009), Trends in the North Atlantic carbon sink: 1992–2006, *Global Biogeochemical Cycles*, *23*, doi: 10.1029/2008GB003383.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen (2005), The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society*, *131*, 2961–3012, doi:10.1256/qj.04.176.

Valsala, V., and S. Maksyutov (2010), Simulation and assimilation of global ocean pCO_2 and air-sea CO_2 fluxes using ship observations of surface ocean pCO_2 in a simplified biogeochemical offline model, *Tellus B*, *62*, 821–840, doi:10.1111/j.1600-0889.2010.00495.x.

Valsala, V., S. Maksyutov, M. Telszewski, S.-I. Nakaoka, Y. Nojiri, M. Ikeda, and R. Murtugudde (2011), Climate impacts on the structures of the North Pacific air-sea CO_2 flux variability, *Biogeosciences Discussions*, *8*, 4239–4280, doi:10.5194/bgd-8-4239-2011.

Van Sebille, E., L. M. Beal, and A. Biastoch (2010), Sea surface slope as a proxy for Agulhas Current strength, *Geophysical Research Letters*, *37*, doi:10.1029/2010GL042847.

Visbeck, M. H., J. W. Hurrell, L. Polvani, and H. M. Cullen (2001), The North Atlantic Oscillation: Past, present, and future, *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 12,876–12,877, doi:10.1073/pnas.231391598.

Wanninkhof, R. H. (1992), Relationship between wind speed and gas exchange over the ocean, *Journal of Geophysical Research*, *97*, 7373–7382, doi:10.1029/92JC00188.

Wanninkhof, R. H., and W. R. McGillis (1999), A cubic relationship between air-sea CO_2 exchange and wind speed, *Geophysical Research Letters*, *26*, 1889, doi:10.1029/1999GL900363.

Wanninkhof, R. H., R. A. Feely, H. Chen, C. E. Cosca, and P. P. Murphy (1996), Surface water fCO_2 in the eastern equatorial Pacific during the 1992–1993 El Niño, *Journal of Geophysical Research*, *101*, 16,333–16,343, doi:10.1029/96JC01348.

Wanninkhof, R. H., W. E. Asher, D. T. Ho, C. Sweeney, and W. R. McGillis (2009), Advances in quantifying air-sea gas exchange and environmental forcing, *Annual Review of Marine Science*, *1*, 213–244, doi:10.1146/annurev.marine.010908.163742.

Watson, A. J., and A. C. Naveira Garabato (2005), The role of Southern Ocean mixing and upwelling in glacial-interglacial atmospheric CO_2 change, *Tellus B*, *58*, 73–87.

Watson, A. J., U. Schuster, D. C. E. Bakker, N. R. Bates, A. Corbière, M. González-Dávila, T. Friedrich, J. Hauck, C. Heinze, T. Johannessen, A. Körtzinger, N. Metzl, J. Olafsson, A. Olsen, A. Oschlies, X. A. Padin, B. Pfeil, J. M. Santana-Casiano, T. Steinhoff, M. Telszewski, A. F. Rios, D. W. R. Wallace, and R. H. Wanninkhof (2009), Tracking the variable North Atlantic sink for atmospheric CO_2, *Science*, *326*, 1391–1393, doi:10.1126/science.1177394.

Weare, B. C., A. R. Navato, and R. E. Newell (1976), Empirical orthogonal analysis of Pacific sea surface temperatures, *Journal of Physical Oceanography*, *6*, 671–678, doi: 10.1175/1520-0485(1976)006⟨0671:EOAOPS⟩2.0.CO;2.

Weiss, R. F. (1974), Carbon dioxide in water and seawater: The solubility of a non-ideal gas, *Marine Chemistry*, *2*, 203–215.

Wesely, M. L. (1986), Response to "Isotopic Versus Micrometeorologic Ocean CO_2 Fluxes: A Serious Conflict" by W . Broecker et al ., *Journal of Geophysical Research*, *91*, 1982–1984, doi:10.1029/JC091iC09p10533.

Wichura, M. J. (1988), The percentage points of the normal distribution, *Applied Statistics*, *37*, 477–484.

Wong, C. S., J. R. Christian, S.-K. Emmy Wong, J. Page, L. Xie, and S. Johannessen (2010), Carbon dioxide in surface seawater of the eastern North Pacific Ocean (Line P), 1973–2005, *Deep Sea Research Part I: Oceanographic Research Papers*, *57*, 687–695, doi:10.1016/j.dsr.2010.02.003.

Yelland, M. J., R. W. Pascal, P. K. Taylor, and B. I. Moat (2009), AutoFlux: an autonomous system for the direct measurement of the air-sea fluxes, *Journal of Operational Oceanography*, *2*, 15–23.

Yulaeva, E., and J. M. Wallace (1994), The signature of ENSO in global temperature and precipitation fields derived from the microwave sounding unit, *Journal of Climate*, *7*, 1719–1736.

Zhang, L., L. Xue, M. Song, and C. Jiang (2010), Distribution of the surface partial pressure of CO_2 in the southern Yellow Sea and its controls, *Continental Shelf Research*, *30*, 293–304, doi:10.1016/j.csr.2009.11.009.

Zhang, X. (2007), Comment on "Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations" by D. T. Ho et al., *Geophysical Research Letters*, *34*, 2–4, doi:10.1029/2007GL029983.

Zhang, Y., J. M. Wallace, and D. S. Battisti (1997), ENSO-like Interdecadal Variability : 1900 – 93, *Journal of Climate*, *10*, 1004–1020, doi:10.1175/1520-0442(1997) 010⟨1004:ELIV⟩2.0.CO;2.