Marine Chemistry 183 (2016) 41–49

Contents lists available at ScienceDirect

Marine Chemistry

journal homepage: www.elsevier.com/locate/marchem

A new statistical approach for interpreting oceanic fCO_2 data

Hongjie Wang [a], Xinping Hu [a,*], Blair Sterba-Boatwright [b,*]

[a] Department of Physical and Environmental Sciences, Texas A&M University —Corpus Christi, Corpus Christi, TX 78412, United States
[b] Department of Mathematics & Statistics, Texas A&M University —Corpus Christi, Corpus Christi, TX 78412, United States

ARTICLE INFO

Article history:
Received 4 November 2015
Received in revised form 18 May 2016
Accepted 23 May 2016
Available online 24 May 2016

Keywords:
Surface ocean fCO_2
Ocean margin
Generalized Additive Mixed Modeling
Heteroscedasticity

ABSTRACT

Surface ocean CO_2 fugacity (fCO_2) often exhibits large fluctuations and heterogeneity because of multiple controlling factors that pose a challenge for trend analysis, especially in the ocean margins. We propose a new statistical approach, Generalized Additive Mixed Modeling (GAMM), to interpret oceanic fCO_2 data in two ocean margins (Japan and Europe) and two open ocean areas. The latter included areas near the Hawaii Ocean Time-series (HOT) and the Bermuda Atlantic Time-series (BATS). This method utilizes day of year, sea surface salinity (SSS), sea surface temperature (SST), and sampling date as predictors. Using this method, we were able to derive multidecadal fCO_2 trends with both improved precision and greater robustness to data gaps compared to an existing deseasonalization method used in the open ocean. The fCO_2 trend derived by our method for the Japanese margin (1992–2013), the European margin (1989–2014), and the open ocean near HOT (1983–2013) were 2.1 ± 0.6, 1.9 ± 0.7, and 2.0 ± 0.5 μatm year^{-1} (mean ± standard deviation of multiple 1° × 1° grids in margins and 5° × 5° grids in open ocean), respectively, and the fCO_2 trends were all close to the atmosphere CO_2 trend (1.7–1.9 μatm year^{-1}). Our analysis produced generally smaller standard errors (paired *t*-test, $p \ll 0.001$) than those obtained using the existing method based on the same dataset. In addition, our method was less sensitive to data gaps compared to this existing method. However, for regularly spaced fCO_2 times-series data, for example, discrete bottle data collected in the BATS station in 1991–2011, this method was not advantageous over the existing method (1.9 ± 0.2 vs. 2.0 ± 0.2 μatm year^{-1}). To test the broader applicability of this method, we compared fCO_2 trends in the Southern Ocean derived using our method with those from a recently reported Markov Chain Monte Carlo method and found no significant difference between the two sets of values. Therefore, we recommend the application of our method in interpreting fCO_2 data in different oceanic environments.

1. Introduction

The concentration of carbon dioxide (CO_2) in the atmosphere has increased from ~280 ppm in preindustrial era to ~399 ppm (annual mean) in 2015 (http://www.esrl.noaa.gov/gmd/ccgg/trends/). If CO_2 emission continues under the "business-as-usual" scenario, atmospheric CO_2 level is expected to exceed 900 ppm by the end of this century (Collins et al., 2013).

The ocean serves as a natural sink (1.4–2.9 Pg C year^{-1}) for atmospheric CO_2 (Khatiwala et al., 2009; Landschützer et al., 2014; Sabine et al., 2004; Takahashi et al., 2009). It takes up CO_2 primarily through air-sea gas exchange, which is a function of both physical conditions (wind speed, salinity, and temperature) and the thermodynamic gradient between seawater and the atmosphere ($\Delta fCO_2 = fCO_2\_ocean - fCO_2\_air$). Surface ocean CO_2 fugacity (fCO_2) is controlled by dissolved inorganic carbon (DIC) concentration, total alkalinity (TA), sea surface salinity (SSS), and sea surface temperature (SST) (Takahashi et al., 1993). Both TA and DIC can also be affected by biogeochemical processes, including respiration, photosynthesis, and carbonate precipitation and dissolution (Zeebe and Wolf-Gladrow, 2001). Note in the literature both fCO_2 and pCO_2 (partial pressure of CO_2) are often used interchangeably. Given that these two parameters are very close to each other ($\Delta < 0.4\%$ (Zeebe and Wolf-Gladrow, 2001)) under normal seawater conditions, trend analysis based on either dataset would produce essentially the same results.

The ocean's uptake of CO_2 has reduced atmospheric buildup of this greenhouse gas and damped associated climate changes. However, CO_2 absorption has also reduced the saturation state of calcium carbonate, known as "ocean acidification" (Doney et al., 2009; Hoegh-Guldberg et al., 2007; Orr et al., 2005), which could hinder carbonate formation in marine organisms (e.g., corals, marine plankton, coralline algae, and shellfish) (Kleypas et al., 1999; Orr et al., 2005; Waldbusser et al., 2015).

The change rate of oceanic fCO_2 relative to that in the atmosphere provides information on how the strength of the ocean as either a CO_2

* Corresponding authors.
E-mail addresses: xinping.hu@tamucc.edu (X. Hu), sterbaboatwright@tamucc.edu (B. Sterba-Boatwright).

http://dx.doi.org/10.1016/j.marchem.2016.05.007

sink or source evolves over time. For example, a region with an oceanic fCO_2 increase rate higher than the atmospheric rate can be interpreted as a decreasing sink or an increasing source depending on the initial seawater fCO_2 relative to atmospheric fCO_2. Conversely, an oceanic fCO_2 increase rate less than the atmospheric rate can be interpreted as an increasing sink (or a decreasing source) (Landschützer et al., 2014; Landschützer et al., 2015; Lenton et al., 2012; Majkut et al., 2014). On the other hand, because of high primary productivity in the ocean margins (Liu et al., 2010), knowledge on the fCO_2 changes there is essential for understanding its role in global carbon cycle.

Getting a more precise fCO_2 trend is crucial to understanding the evolution of carbon sink/source in difference areas of the ocean. In the literature, however, there is no universal method in calculating the fCO_2 trend in surface seawater. The simplest approach is a linear least square regression using observed data. For example, Dore et al. (2009) directly applied a linear regression using observed fCO_2 data in waters at the Hawaii Ocean Time-series (HOT) and found that the trend in the surface ocean is 1.9 ± 0.2 μatm year^{-1}. To avoid biases caused by temporal weighting, some researchers chose the observed data in only selected months for their calculations. For example, Midorikawa et al. (2010) reported that the fCO_2 trend in the western North Pacific (1983–2007) is 1.58 ± 0.12 μatm year^{-1} in winter and 1.37 ± 0.33 μatm year^{-1} in summer, and neither trend is significantly different from their respective air fCO_2 trends (1.65 ± 0.05 μatm year^{-1} in winter and 1.54 ± 0.08 μatm year^{-1} in summer). Similarly, to minimize the biological effects on data interpretation in the Iceland Sea, Olafsson et al. (2009) only selected the first 67 days (presumably with little biological production) of the sampled years (1985–2008) to calculate fCO_2 trend, then with the aid of multivariate linear regression ($y = a \times \text{time} + b \times \text{Temp} + c$), they found that the fCO_2 trend is 2.1 ± 0.2 μatm year^{-1}.

The most commonly used approaches generally obtain linear trends after various deseasonalization techniques. For example, the fCO_2 trend analysis from 1983 to 2010 in Bermuda Atlantic Time-series (BATS) was performed using both raw data and deseasonalized data (Bates et al., 2012). The rate is 1.8 ± 0.1 μatm year^{-1} using the deseasonalized data, but 1.6 ± 0.2 μatm year^{-1} using the raw data. This latter trend may be biased due to more frequent sampling in spring, thus Bates et al. (2012) recommended the deseasonalization method. To fit the seasonal cycle in the European Station for Time series in the ocean (ESTOC) near the Canary Islands, Santana-Casiano et al. (2007) used harmonic functions to decompose the time series into a trend, seasonal variations and errors, and serial correlation was modeled using a second-order autoregressive process. They found that the fCO_2 trend from1995 to 2004 is 1.6 ± 0.4 μatm year^{-1}. In another study, Schuster et al. (2009) fitted a harmonic function in the form of $y = a + b \times t + c \times \cos(2\pi t + d)$ in the North Atlantic to calculate the fCO_2 trend, where t is the year lapse since an arbitrarily defined reference year, and b is the fCO_2 trend in μatm year^{-1}. Their results suggested that sea-surface fCO_2 has closely followed atmospheric fCO_2 in the subtropical regions. McKinley et al. (2011) and Fay and McKinley (2013) also adopted this harmonic function. They found that the fCO_2 trend in the open ocean is sensitive to the chosen start and end years, resulting from climatic oscillations such as El Niño/Southern Oscillation, North Atlantic Oscillation. However, these oscillation signals would fade away as timescales increase (i.e., 25 years), and the fCO_2 trend is parallel to the atmospheric trend (Fay and McKinley, 2013; McKinley et al., 2011).

Takahashi et al. (2009) proposed a simple yet effective deseasonalization method and used it to obtain global surface pCO_2 climatology. In their study, all historical data from regularly spaced "grids" (latitude × longitude of either 4° × 5° or 5° × 10°) were used to obtain the rate in each grid. First, seasonal changes were calculated on the basis of the monthly mean values computed from a 4-year subsample of the entire time series. Then pCO_2 values for months with no measurements were estimated by a linear interpolation using two adjacent monthly means. The difference between a monthly mean and the annual mean represents the correction to be applied to deseasonalize the monthly mean. Finally, the deseasonalized monthly mean values are regressed against time (year) using least square method to obtain the mean rate of change (Lenton et al., 2012; Takahashi et al., 2009). Using their method, surface water pCO_2 in the North Atlantic, North and South Pacific and Southern Oceans increases at a mean rate of 1.5 μatm year^{-1} from 1970 to 2007 (Takahashi et al., 2009). Overall, different independent studies suggested that surface fCO_2 trend in the open ocean has increased more or less the same as the atmospheric fCO_2 has. These trends indicate that the driving force for air-sea CO_2 gas exchange has not changed significantly over the last three decades (Bates et al., 2012).

Recently, Majkut et al. (2014) developed a Markov Chain Monte Carlo (MCMC) method to calculate long-time pCO_2 trends in the open ocean. They found that the pCO_2 trend in the Southern Ocean is 1.4 ± 0.5 μatm year^{-1} (based on the Lamont-Doherty Earth Observatory LDEO V2010 database) in the 1995–2008 period, and they suggested a global increase in the CO_2 uptake of 0.4 ± 0.1 Pg C year^{-1} decade^{-1}, because surface pCO_2 is increasing more slowly than the atmospheric value (but not significantly different from the latter). This method resulted a smaller value than linear regression method (2.2 ± 0.2 μatm year^{-1}) reported in Lenton et al. (2012) based on the same dataset. Using a neural network approach, Landschützer et al. (2015) reported the weakening carbon sink trend in the Southern ocean (south of 35°S) stopped around 2002 and the carbon sink there increasing from ~0.6 Pg C year^{-1} in 2002 to ~1.2 Pg C year^{-1} in 2011.

One of the assumptions required for inference with regression models is that residuals have constant variance throughout the range of the predictors and the fitted values ("homoscedasticity"). Unfortunately, in most cases this statistical test was not performed or explicitly demonstrated, even in the sophisticated models published recently, for example, the MCMC approach (Majkut et al., 2014) or the neural network approach (Landschützer et al., 2013; Landschützer et al., 2015). Given the spatial heterogeneity of biogeochemical reactions in marginal areas, which are heavily modulated by terrestrial influences, biological activities, and physical processes (such as upwelling) (Liu et al., 2010), it is unknown whether the above deseasonalization method can be used in these areas, where the cyclic (or seasonal) behavior may not be as stable as that in the open sea (Cai, 2011; Liu et al., 2010). The assumption of homoscedasticity is not necessarily met because of heterogeneity of the marginal ocean and large fCO_2 fluctuations. Violation of this assumption suggests that linear least squares may not necessarily result in the best unbiased estimates of model parameters. Therefore, the regression analysis obtained from heteroscedastic variables could produce inaccurate standard errors, too large for some values and too small in others, and potentially biased regression coefficients.

In this work, we used a statistical approach that can fit the seasonal cycle more precisely and reduce the impact of heteroscedasticity on the calculated trend, and we compared results from our method with those from the Takahashi method (hereafter T0) that effectively generated open ocean pCO_2 climatology (Takahashi et al., 2009). We also compared fCO_2 trend derived using our method with recently reported results in the Southern Ocean using the MCMC method to test a broader applicability of our method.

2. Method

2.1. Data

We chose two marginal areas and two open ocean areas in this study. The two marginal areas include the northwestern North Pacific margin (east of Japan, 28–45°N, 130–150°E) and northeastern North Atlantic offshore Europe (44–50°N, 3–14°W). The two open ocean areas are near the two time series stations, HOT (10–30°N, 140–165°W) and BATS (31°40′N, 64°10′W), respectively. The grid size for our model