# Time series *p*CO<sub>2</sub> at a coastal mooring: Internal consistency, seasonal cycles, and interannual variability

Janet J. Reimer<sup>1</sup>, Wei-Jun Cai<sup>1,2</sup>, Liang Xue<sup>1,3</sup>, Rodrigo Vargas<sup>4</sup>, Scott Noakes<sup>5</sup>, Xinping Hu<sup>6</sup>, Sergio R. Signorini<sup>7</sup>, Jeremy T. Mathis<sup>8</sup>, Richard A. Feely<sup>9</sup>, Adrienne J. Sutton<sup>9,10</sup>, Christopher Sabine<sup>9</sup>, Sylvia Musielewicz<sup>9,10</sup>, Baoshan Chen<sup>1,2</sup> and Rik Wanninkhof<sup>11</sup>

<sup>1</sup>School of Marine Science and Policy, University of Delaware, Newark, DE 19716, USA

<sup>2</sup>Department of Marine Sciences, University of Georgia, Athens, GA 30602, USA

<sup>3</sup>Center for Ocean and Climate Research, First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China

<sup>4</sup>Department of Plant and Soil Sciences, University of Delaware, Newark, DE 19716, USA

<sup>5</sup>Center for Applied Isotope Studies, University of Georgia, Athens, GA 30602, USA

<sup>6</sup>Department of Physical and Environmental Sciences, Texas A&M University - Corpus Christi, Corpus Christi, TX 78412, USA

<sup>7</sup>Science Applications International Corporation, Greenbelt, MD 20771, USA

<sup>8</sup>NOAA Arctic Research Program, Silver Spring, MD 20910, USA

<sup>9</sup>Pacific Marine Environmental Laboratory, NOAA, Seattle, WA 98115, USA

<sup>10</sup>Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98105, USA

<sup>11</sup>Atlantic Oceanographic and Meteorological Laboratory, NOAA, Miami, FL 33149, USA

Continental Shelf Research

#### 1. Abstract

Marine carbonate system monitoring programs often consist of multiple observational methods that include underway cruise data, moored autonomous time series, and discrete water bottle samples. Monitored parameters include all, or some of the following: partial pressure of  $CO_2$  of the water ( $pCO_{2w}$ ) and air, dissolved inorganic carbon (DIC), total alkalinity (TA), and pH. Any combination of at least two of the aforementioned parameters can be used to calculate the others. In this study at the Gray's Reef (GR) mooring in the South Atlantic Bight (SAB) we: examine the internal consistency of pCO<sub>2w</sub> from underway cruise, moored autonomous time series, and calculated from bottle samples (DIC-TA pairing); describe the seasonal to interannual  $pCO_{2w}$  time series variability and air-sea flux (FCO<sub>2</sub>), as well as describe the potential sources of  $pCO_{2w}$  variability; and determine the source/sink for atmospheric  $pCO_2$ . Over the ~8.5 years of GR mooring time series, mooring-underway and mooring-bottle calculated- $pCO_{2w}$ strongly correlate with *r*-values > 0.90.  $pCO_{2w}$  and  $FCO_2$  time series follow seasonal thermal patterns; however, seasonal non-thermal processes, such as terrestrial export, net biological production, and air-sea exchange also influence variability. The linear slope of time series  $pCO_{2w}$  increases by 5.2±1.4 µatm y<sup>-1</sup> with  $FCO_2$  increasing 51 to 70 mmol m<sup>-2</sup> y<sup>-1</sup>. The net  $FCO_2$  sign can switch interannually with the magnitude varying greatly. Non-thermal  $pCO_{2w}$  is also increasing over the time series, likely indicating that terrestrial export and net biological processes drive the long term  $pCO_{2w}$  increase.

**Key words:** ocean nargin CO<sub>2</sub>, east coast, inorganic carbon system, internal consistency

## **Highlights**

- Mooring, underway, and discrete water samples CO<sub>2</sub> is examined
- High interannual CO<sub>2</sub> variability on ocean margins could be due to freshwater inputs
- Non-thermal processes likely contribute to increased CO<sub>2</sub> on an ocean margin

## 2. Introduction

The world's oceans have been estimated to take up ~30% of the atmospheric  $CO_2$  due to fossil fuel burning and other anthropogenic influences since the onset of the Industrial Revolution (Le Quere et al., 2016; Sabine et al., 2004), which has resulted in various regional declines in seawater pH (Feely et al., 2004; Takahashi et al., 2014, 2006). Overall,  $pCO_{2w}$  changes can be caused by at least one, or a combination, of: (1) atmospheric  $CO_2$  uptake (Sabine et al., 2004); (2) sea surface temperature (Jiang et al., 2010; Shadwick et al., 2010; Takahashi et al., 1993); (3) drainage basin hydrologic cycles (Jiang et al., 2013); (4) coastal currents due to wind and/or weather patterns (Huang et al., 2013); (5) upwelling and water mass mixing (Feely et al., 2008); and (6) *in situ* biological production, which may be enhanced by large inputs from organic matter (OM) remineralization (nutrients) in the surface or terrestrial ecosystems, thus creating a positive feedback for biological production in shallow and/or stratified coastal systems (Cai et al., 2011; Shadwick et al., 2011; Sunda and Cai, 2012). Anthropogenically-driven  $pCO_{2w}$ , and subsequently carbonate chemistry, temporal perturbations super-imposed on natural cycles are causes for concern for marine life and are now the focus of numerous studies that drive governmental policy and conservation efforts directed at mitigating the harmful effects in our coastal oceans. Therefore, rigorous *in situ* monitoring programs and modeling efforts are now being employed to track the fate of  $CO_2$  in the world's oceans; consequently, there is also need to ensure that all methods for observing  $CO_2$  are consistent.

Mooring platforms on coastal margins have been identified as key means for monitoring changing ocean  $pCO_{2w}$  on various time scales (Michalak et al., 2011). The coastal zone and marginal seas are expected to have high  $pCO_{2w}$  variability due to terrestrial influences (Bauer et al., 2013; Cai, 2011) and could be prone to enhanced pH declines (acidification) due to large quantities of OM that are remineralized in the shallow, biologically productive waters (Cai et al., 2011). CO<sub>2</sub>, the primary anthropogenically-derived driver of ocean acidification (OA), dynamics and variability are only just beginning to be understood in most coastal regions and margins around the world. Though there is a specific lack of understanding on the U.S. east coast (Bauer et al., 2013; Jiang et al., 2013). The complex nature of CO<sub>2</sub> biogeochemistry and coastal zone heterogeneity, as well as limitations in spatio-temporal coverage, make discerning the time scales and sources of CO<sub>2</sub> variability challenging to identify (Cai, 2011; Jiang et al., 2013, 2008b; Vandemark et al., 2011; Wang et al., 2013; Wanninkhof et al., 2015). Moored coastal time series provide high frequency temporal resolution, while cruise (underway systems and discrete water samples) observations provide complimentary horizontal and vertical spatial patterns needed to assess the aforementioned influences on CO<sub>2</sub> biogeochemistry in coastal waters.

Historically, cruise methods for  $CO_2$  data collection (bottle samples and underway observations) have been in use for several decades to discern spatial patterns and variability, sometimes extrapolating to create higher temporal resolution (Takahashi et al., 1993). While higher frequency moored time series data collection has only been used over the last decade or so (Sutton et al., 2014). Since  $CO_2$  data collection has been, and remains to be, carried out via different methods at the same site, as well as across multiple sites, it is essential to show that consistent results can be reached by all methods. Additionally, discrete water samples for dissolved inorganic carbon (DIC), pH, and total alkalinity (TA) observations are not only used as direct measures themselves, but also to calculate  $pCO_{2w}$ , often in the absence of a direct method for observing pCO<sub>2w</sub>. Calculated values must also be in agreement with mooring and underway observations (Patsavas et al., 2015) not only as an independent method, but also because DIC and TA can be included in biogeochemical models to determine the relative importance and the contributions of different biogeochemical processes to overall pCO<sub>2w</sub> variability (Boehme et al., 1998; Xue et al., 2016). Internal consistency checks must also be periodically made during long-term deployments to ensure that results from all methods continue to agree (Patsavas et al., 2015; Ribas-Ribas et al., 2014) so that combined datasets can increase the spatio-temporal coverage of observed marine  $pCO_{2w}$  as well as allow for the examination of the biogeochemical processes that influence variability.

The first objective is to describe seasonal to interannual variability of the moored GR  $pCO_{2w}$  time series, accompanying variables, and air-sea  $pCO_{2w}$  flux ( $FCO_2$ ), as well as examine how biogeochemical and environmental influences could affect  $pCO_{2w}$  on this ocean margin. The second objective is to test the agreement, or internal consistency, between underway and discrete water bottle-calculated  $pCO_{2w}$ from cruises in the South Atlantic Bight (SAB) with a  $pCO_{2w}$  time series measured by a coastal mooring in the Gray's Reef (GR) National Marine Sanctuary. Internal consistency serves to test the validity and quality of instrumentation as well as show that the same results can be achieved through various methods. Previous validation tests have been carried out on other National Oceanographic and Atmospheric Administration (NOAA) moored autonomous  $pCO_{2w}$  (MAPCO<sub>2</sub>) systems (Schar et al., 2010; Sutton et al., 2014), however, this is the first field internal consistency effort compiled from various exercises over the lifetime of a NOAA MAPCO<sub>2</sub> system for a coastal margin heavily influenced by terrestrial sources of CO<sub>2</sub> (Cai and Wang, 1998; Jiang et al., 2013, 2008a). MAPCO<sub>2</sub> time series are still rather new, with  $\leq$  10 years of observations in the coastal oceans; however, in some cases cruises have been collecting *p*CO<sub>2w</sub> observations for decades longer than moorings. The broader importance of the objectives is that in the future combining results from cruises over the last few decades and moored time series could lay the ground work for studies to discern multi-year seasonal variability. Within the first objective, we also aim to determine if the site is an annual net source or sink for atmospheric CO<sub>2</sub>, as coastal margins represent highly biologically and socio-economically important regions that could be impacted by increased atmospheric CO<sub>2</sub> uptake. The GR mooring is one of the few coastal CO<sub>2</sub> time series that provides an important opportunity to compare *p*CO<sub>2w</sub> measured from a coastal margin MAPCO<sub>2</sub> system to underway and discrete water samples, as well as a describe a continuous time series of this length.

# 3. Methods

#### 3.1 Study site and hydrographic description

The GR mooring (owned and maintained by the National Data Buoy Center as NDBC-41008; http://www.ndbc.noaa.gov/station\_page.php?station=41008) with a National Oceanographic and Atmospheric Administration (NOAA) MAPCO<sub>2</sub> system is located at the ~20 m isobath, ~74 km southeast of Savannah, Georgia and ~40 km east of Altamaha River estuary and coastal marshes around Sapelo Island and the Duplin River, along the Georgia coast (31.400°N, -80.868°W; Figure 1). The GR mooring is located at a previously defined hydrographic boundary of the inner and middle shelf regions of the U.S. SAB (Pomeroy et al., 2000), loosely based on the location of lower saline waters entering the SAB from coastal freshwater sources and Gulf Stream (GS) influences. The SAB is the coastal margin from Cape Hatteras (North Carolina) to Cape Canaveral (Florida). GS intrusions on the shallow inner shelf occur in only ~5% of all reported surface intrusions in the SAB (Castelao, 2011). By the time that upwelled waters from off the shelf reach the inner shelf they have warmed, making them difficult to detect through temperature analysis (Yoder et al., 1985). This ocean margin is also influenced by 10 rivers that empty into the SAB via marsh-dominated estuaries (Figure 1). Historically, river discharge is at a maximum from December through April with a water residence time on the shelf of one to three months, which is partially dependent on regional river stream flow volume (Menzel, 1993). The Altamaha River and its watershed are the largest in the region, with its delivery of inorganic nutrients linked to eutrophication and increased primary production in coastal marshes and its estuary (Sheldon and Burd, 2013). We assume that the Altamaha River represents the general seasonality and biogeochemical characteristics of other rivers in the region, and likely has the greatest influence on the GR mooring of all the regional rivers. There is a southward flowing coastal current (salinity < 34) that is typically held shoreward of the  $\sim 20$  m isobath (Figure 1) and entrains the lower salinity waters (residence time of up to three weeks) coming from the rivers (Menzel, 1993). Seaward of the ~20 m isobath oceanic water influenced by the GS flows north, and then closer to the shelf break the GS itself also flows north, thus creating a complex hydrographic shelf region (Pomeroy et al., 2000, and reference therein). During wet periods (December to May, peaking in the late winter/early spring) when river stream flow is increased, the coastal low salinity frontal zone may extend out beyond the  $\sim 20$  m isobath; conversely, when river stream flow is decreased the low salinity region is greatly reduced and held closer to the coast landward of the ~20 m isobath (Blanton and Atkinson, 1983). The SAB has a relatively weak spring bloom, peaking April through July, as well as small patchy blooms sporadically timed throughout the year (Castelao and He, 2013). The location of the GR mooring allows us to examine terrestrial influences on the  $CO_2$  signal on this ocean margin.

# 3.2 Compilation of *p*CO<sub>2w</sub> observations from cruises

The majority of the cruises that passed close to the GR mooring were not specifically designed for internal consistency measurements of inorganic carbon system parameters, rather they represent years of seasonal cruises that repeated the same transects and stations. A total of 24 cruises were used in this work and are summarized in the Supporting Information Table 1S with more detailed focus on two cruises in 2011 (October) and the July 2015 East Coast Ocean Acidification (ECOA) cruise as specific examples of the ground-truthing exercises carried out at the GR mooring (see Supporting Information). Most cruises collect continuous observations (2 to 3 minute intervals ) of surface  $pCO_2$  (dry mole fraction of CO<sub>2</sub> converted to  $pCO_2$  with the accompanying necessary variables (Jiang et al., 2008b)), as well as discrete water samples for DIC and TA, which are used to calculate  $pCO_2$  via the Matlab version of CO2SYS (Lewis and Wallace, 1998).

DIC and TA samples were collected and analyzed following standard methods with measurement precision of  $\pm 0.1\%$  (Huang et al., 2012) and accuracy of  $\pm 2 \mu \text{mol kg}^{-1}$ , tested with Dickson Certified Reference Materials (CRM). DIC and TA from bottle samples are used to calculate  $pCO_{2w}$ . We use the equilibrium constants of Mehrbach et al. (1973) refit by Dickson and Millero (1987) and boron constants from Lee et al. (2010), which give the closest match to mooring  $pCO_{2w}$ . DIC samples have also been collected and measured in the Altamaha River at Doctortown, GA (USGS station 02226000) above the head of the tide from November 2000 through the present using the same methods as GR mooring bottle samples. These samples are used to assess the potential for terrestrial sources of DIC to the region represented by the GR mooring.

There were various underway  $pCO_{2w}$  analyzers used during the cruises included over the almost 10 years of field campaigns described in this study, however, the general design and system uncertainty are all  $\pm 2 \mu$  atm, with calibrations using gas standards at set time intervals of several hours, similar to the MAPCO<sub>2</sub> system (Jiang et al., 2008b; Pierrot et al., 2009). When comparing the various  $pCO_{2w}$  values, we align the underway and bottle samples to the closest mooring observation, which typically occur every three hours starting at 00:17 UTC, therefore, the closest temporal sample, for either method, is never more than an hour and a half different. Statistical assessment of internal consistency between methods is done following similar methods employed in previous works, such as assessment of linear correlations (*r*-value), residual values (mean, standard deviation, and median of the difference between the observed value and the predicted value of the linear model), bias (a source of error due to predictor variables), and root mean square errors (RMSE; a measure of the standard deviation between the predicted and observed values of the linear relationship) (Patsavas et al., 2015; Ribas-Ribas et al., 2014). The residual values, RMSE, and bias are all measures of deviations of predicted values of the linear fit, while any one of these estimations of deviations are likely sufficient, since this is a highly variable coastal margin reporting various estimations creates a strict criteria for agreement. We also assess the potential for surface layer and spatial heterogeneity using SSS and SST differences ( $\Delta$ SSS and  $\Delta$ SST, respectively) between the methods. Finally, we explore the potential sources of uncertainty in internal consistency using adjustments made to measured values in CO2SYS.

## 3.3 Time series observations and data processing

The MAPCO<sub>2</sub> system, described in detail in Sutton et al. (2014), has been measuring xCO<sub>2</sub> (the partially dry mole fraction of CO<sub>2</sub>) referenced to water (xCO<sub>2w</sub>; at ~0.6 m depth) and air (xCO<sub>2a</sub>) above the sea surface (~1.5 m), from July 2006 to October 2014 at a sampling interval of three hours. From the same mooring we also have observations of sea surface temperature (SST), sea surface salinity (SSS), other standard MAPCO<sub>2</sub> system variables (e.g., atmospheric pressure, and relative humidity to calculate the partial pressure at 100% humidity), and wind speed from NDBC instrumentation. The anemometer is at ~5 m above sea level; therefore wind speed observations for FCO<sub>2</sub> are adjusted to 10 m (Jiang et al., 2008b). Wind speed does not consider gust speed, but rather is an hourly average of the second moment (reported by NDBC). The MAPCO<sub>2</sub> system is calibrated with standard gasses prior to deployment, has an onboard standard that runs during each measurement cycle, and is also calibrated post deployment, with any necessary corrections applied prior to release of finalized observations (Sutton et al., 2014). pCO<sub>2w</sub> was calculated following that described in Sutton et al. (2014). We applied quality control/quality assessment checks and gap-filling techniques, which are presented in the Supporting Information. Gap-filled values are not used for internal consistency analyses.

To determine possible attributing factors to  $pCO_{2w}$  variability, besides SST and SSS (thermodynamics), we also analyze a time series of Altamaha River stream flow at Doctortown, GA from

2006 through 2014. This site is a USGS hydrological monitoring site above the head of the tide (zero salinity), and accounts for approximately 98.5% of the discharge in the Altamaha River watershed (Sheldon and Burd, 2013). The 15 minute observations are summed to daily values to show the amount of water flowing through the watershed region and into the estuaries. We assume that while there are complex current patterns that will affect the trajectory of water movement, DIC inputs to the SAB from terrestrial sources (aforementioned DIC samples from the USGS site) in the region could be generally represented by Altamaha watershed DIC.

### 3.4 Thermal and non-thermal pCO<sub>2w</sub> calculations

To determine the relative importance of the competing effects of non-thermal and thermal influences on changes in  $pCO_{2w}$  (seasonal cycle), we calculate two operationally defined fractions of  $pCO_{2w}$  using the decomposition by Takahashi et al. (2002). Equation 1 defines non-thermal  $pCO_{2w}$  ( $pCO_{2nt}$ ), which is normalized to a time series mean SST. Deviations from  $pCO_{2nt}$  imply that non-thermal conditions perturb  $pCO_{2w}$ . Equation 2 defines mean  $pCO_{2w}$  adjusted to mean SST, where perturbations of  $pCO_{2t}$  indicate that thermal effects impact  $pCO_{2w}$ :

$$(pCO_2 \text{ at } SST_{mean}) = pCO_{2m} = pCO_{2w} \times exp(0.0412 \times (SST_{mean} - SST))$$
(1)

$$(pCO_2 \text{ at } SST) = pCO_{2t} = pCO_{2mean} \times \exp(0.0412 \times (SST - SST_{mean}))$$
(2)

where  $pCO_{2w}$  is the observed value,  $pCO_{2mean}$  is the time series mean  $pCO_{2w}$ , SST<sub>mean</sub> is the time series mean SST, and SST is the observed value accompanying the  $pCO_{2w}$  measurement. The mean values used in these equations are reported in Section 4.2. The value of the exponent (4.12%) is derived for this specific data set over the SST and SSS range of measured DIC and TA oceanic and near-shore end members following the method of Takahashi et al. (1993) and is within 1 µatm of the Takahashi value of 4.23%. The operationally defined fractions can only distinguish seasonal thermal from non-thermal effects (net biological activity including changes in alkalinity, advection, and air-sea exchange), yet these fractions allow us to distinguish periods when each fraction may be the dominant driver of overall  $pCO_{2w}$  variability.

# 3.5 Air-sea CO<sub>2</sub> flux and gas transfer parameterizations

 $FCO_2$  is calculated for the time series using the parameterization with a recently refit Schmidt number normalized to 660 for a rough surface (Wanninkhof, 2014). The method assumes a dependency on the gas transfer velocity (*k*) derived from wind speed at 10 m above the sea surface ( $k_{u10}$ ), the solubility of CO<sub>2</sub> in sea water (*s*) as a function of SST and SSS, and the air-sea *p*CO<sub>2</sub> difference ( $\Delta pCO_2$ ):

$$FCO_2 = k \times s \times \Delta pCO_2 \tag{3}$$

A negative value for  $FCO_2$  represents a flux into the ocean and a positive value is a flux from the ocean into the atmosphere.

There are numerous parameterizations for characterizing  $k_{u10}$ , though this work does not seek to review all options, we present two representative parameterizations one with a cubic wind speed dependence and non-zero intercept under no/low winds (McGillis et al., 2001) (MG01) and a ploynomial of gas transfer with wind (Wanninkhof et al., 2009) (W09) to encompass a range due to differences in the bulk parameterization equations. The equations used in this work are:

MG01: 
$$k = 3.3 + 0.026 U_{10}^3 \times Sc_{660}^2$$
 (4)

W09: 
$$k = 3.0 + 0.1U_{10} + 0.064U_{10}^2 + 0.011U_{10}^3 \times Sc_{660}^2$$
 (5)

where  $U_{10}$  (m s<sup>-1</sup>) is the wind speed at 10 m above sea level, and *Sc* is the Schmidt number (normalized to 660). The gas transfer velocity parameterization for the MG01 estimate is thought to encompass the effects of white caps at higher wind speeds. The W09 estimate is a hybrid model developed to encompass

other forces that affect turbulence in the surface. It should also be noted that the W09 hybrid model is similar from a quadratic model with a zero intercept for moderate winds  $3 < U_{10} < 15 \text{ m s}^{-1}$  ( $k = 0.24U_{10}^2$ ), which is commonly used in the literature (Wanninkhof et al., 2009). The most recent update to the gas transfer velocity ( $k = 0.251U_{10}^2$ ) also has a zero intercept, and is also not likely to represent the gas transfer well at the GR mooring due to non-wind effects as low winds impacting turbulence and thus k (Wanninkhof, 2014). Uncertainties in  $FCO_2$  due to gap-filled  $pCO_{2w}$  observations are addressed in the Supporting Information section S3. Annual  $FCO_2$  is determined from the sum of the three hour measurements for each day and the daily standard deviations are summed to represent a general range in which the annual sum is most likely to fall.

# 4. Results

# 4.1 Time series of hydrographic and meteorological variables

 $pCO_{2w}$  increases in the warmer months and peaks over the summer (Figure 2B), whereas  $pCO_{2a}$  in the marine boundary layer is higher in the colder months (Figure 2A). The minimum observed  $pCO_{2a}$  is 367 µatm (summer 2006) and the maximum is 427 µatm (winter 2012).  $pCO_{2w}$  largely follows a recurring seasonal thermal pattern ( $pCO_{2t}$ ; blue line Figure 2B, 2C), though excursions are observed, likely indicating influences by non-thermal processes on  $pCO_{2w}$  ( $pCO_{2nt}$ ; pink line in Figure 2B). The daily mean  $pCO_{2w}$  of the time series is 409 ± 73 µatm and ranges from 253 (winter 2010) to 567 µatm (summer 2009). Summer (historically dry season, June through September) mean  $pCO_{2w}$  ranges from 420 ± 20 µatm in 2006 to 515 ± 20 µatm in 2010, while the winter-spring (historically wet season, December through May)  $pCO_{2w}$  is lower than the summer (Figure 2B) with a mean range from 305 ± 37 µatm in 2010 to 368 ± 50 µatm in 2012.  $pCO_{2t}$  peaks in the warmer months, following seasonal warming, whereas  $pCO_{2mt}$  is highest over the late winter when river stream flow (mixing) is typically increased (Figure 2B, 2E). Furthermore, from the late spring to the early winter,  $pCO_{2mt}$  decreases when river stream flow typically decreases.

The mean SST is  $21.5 \pm 5.9$  and ranges from 10 to 30 °C (Figure 2C), and the mean SSS is  $35.3 \pm 1.0$ , however, ranges from 29.7, during high volume river stream flow in 2009, to 36.8 at the end of an extended dry period in early winter of 2012 (Figure 2D). In general, SSS is decreased in the late winter and into the spring when Altamaha River stream flow is increased. Mean U<sub>10</sub> is  $6.0 \pm 3.0$  m s<sup>-1</sup> and ranges from <1 to 21 m s<sup>-1</sup>; with the maximum occurring during Hurricane Andrea in 2007. Daily Altamaha River stream flow ranges from approximately 616 m<sup>3</sup> d<sup>-1</sup> in the dry season (during a drought period) to 233,747 m<sup>3</sup> d<sup>-1</sup> during a high precipitation event in spring of 2009 (Figure 2E). Typically, river stream flow into the SAB is increased in the winter cool periods compared to the summer warm periods (Figure 2E). While there is a large gap in river stream flow data in 2007, it is known to be during a synoptic scale dry period (Sheldon and Alber, 2013). DIC in the Altamaha River is decreased in the winter during increased river stream flow periods, while is it increased during decreased river stream flow (Figure 3A).

Over the ~8.5 years of the time series we fit linear least squares regressions, which minimizes the effects of high variability in data sets, to determine if there are overall increases or decreases in daily mean values for all variables. Using a daily mean value also helps to eliminate the effects of noise by essentially smoothing the daily variations. The linear fit (slope), which is calculated from mid-July 2006 (beginning of the time series) to the end of June 2014), is considered statistically significant if the correlation coefficient (*p*-value) is <0.05 and the 95% confidence interval (CI) does not overlap with zero. There are increases over the time series in  $pCO_{2a}$  and  $pCO_{2w}$  of 2.2 ± 0.2 and 4.9 ± 2.4 µatm y<sup>-1</sup>, respectively, as well as an increase of 7.6 ± 1.9 µatm y<sup>-1</sup> in  $pCO_{2nt}$  (Figure 2A, 2B; Table 1). A linear least squares analysis will be influenced by the start and end points of the data analyzed, however, we attempt to minimize this effect by analyzing full annual cycles starting from the beginning of the time series. We do not consider shorter periods of time over the time series (i.e., multiannual droughts or wet periods) that could be influenced by large episodic variability, which could influence trends. There were no statistically significant changes for,  $pCO_{2n}$ , SST, SSS, or Altamaha River stream flow over the time

series (Table 1 and Figures 2B, 2C, 2D, and 2E, respectively). There is a  $15 \pm 10 \mu$ mol kg<sup>-1</sup> DIC increase (calculated based on bi-weekly sampling, on average) from January 2001 through January 2014, the observations over the GR time series are too few to give a meaningful result, therefore we extended this data analysis. The 95% CI of  $\pm 10 \mu$ mol kg<sup>-1</sup> DIC is almost the same as the estimated increase, likely due a combination of high seasonal variability, as well as data gaps and sampling bias, which was less frequent from 2013 through 2014.

In Figure 3B we include chlorophyll *a* (Chl *a*; 4 km monthly means centered on the Latitude-Longitude of the GR mooring from the NASA Giovanni server: http://gdata1.sci.gsfc.nasa.gov/daacbin/G3/gui.cgi?instance\_id=ocean\_month) and re-plot  $pCO_{2nt}$  to clearly illustrate the behavior of  $pCO_{2nt}$ (the likely source of longer term change) with respect to Chl *a*. There was no significant trend in monthly Chl *a* over the course of the GR mooring time series, nor from the beginning of NASA time series. The highest Chl *a* events occur in the spring of the wettest years (2009, 2010, and 2013; Figure 3B). The average Chl *a* for the site is  $1.73 \pm 1.09$  mg m<sup>-3</sup>, with a median value of 1.48 mg m<sup>-3</sup> and a range of 0.48 (January 2012) to 7.04 mg m<sup>-3</sup> (April 2010).

# 4.2 FCO<sub>2</sub> time series and annual net sums of FCO<sub>2</sub>

 $FCO_2$  is positive (degassing) over the warmer months, and negative in the cooler months following the seasonal  $\Delta pCO_2$  cycle (Figure 4). The mean U<sub>10</sub> seasonal (November-April versus June-October) variation is < 1 m s<sup>-1</sup>, even though in the summer when U<sub>10</sub> is slightly decreased,  $FCO_2$  still closely follows  $\Delta pCO_2$  (Figure 4).

 $\Delta pCO_2$  becomes more positive over the time series, increasing at a rate of 2.6 ± 2.5 µatm y<sup>-1</sup> (Figure 4D and Table 1). The change in *FCO*<sub>2</sub> from July 2006 through June 2014 is not significant, likely due to the high variability, thus the large 95% CI's (Table 2). However, analyzing the trend from the beginning of the time series through the end in October 2014, results in an increase of 51 ± 25 mmol m<sup>-2</sup> y<sup>-1</sup> to 70 ± 34 mmol m<sup>-2</sup> y<sup>-1</sup> for the W09 and MG01 wind speed parameterizations, respectively. The

two  $k_{u10}$  parameterizations used in this work have similar seasonal patterns (Figures 4A and 4B), however, annual estimates using the different parameterizations do not always agree on whether the site is a net source or sink for atmospheric CO<sub>2</sub> (Figure 5). The difference in source/sink estimate is because while both *k*-values are always positive and  $\Delta pCO_2$  is the same in both cases (but may be negative), the difference in the annual *FCO*<sub>2</sub>-sign must come from high wind periods where the two *k*-parameterizations diverge. The difference between the annual net *FCO*<sub>2</sub> estimates can be up to 0.60 mol m<sup>-2</sup> y<sup>-1</sup> in 2008 (Table 2). The error bars in Figure 5 represent the uncertainty in *FCO*<sub>2</sub> using daily mean values to determine the air-sea exchange, and are a sum of the standard deviations of means for each day with measurements (the number of days with measurements is in the parentheses in Table 2). We do not have observations for every day over the ~8.5 years; the sums represent a total for days with observations including days that are gap-filled.

# 4.3 Internal consistency between the mooring, underway, and DIC-TA calculated pCO<sub>2w</sub>

All internal consistency correlation results and residual values are summarized in Table 3 and Figure 6 and 7. Figure 6 depicts the linear correlations of the measurement methods in terms of measured salinity differences and Figure 7 highlights the correlations in terms of measured SST differences. Over the ~8.5 years of the time series the mooring-bottle and mooring-underway pairs correlate with r > 0.90 (p < 0.001), however, the bottle-underway pair only correlate to r = 0.84 (p < 0.001; Table 1). The bias results for each method suggest that both underway and bottle  $pCO_{2w}$  underestimate (best fit slope in Figures 6A and 6C is less than 1) the mooring value and that underway  $pCO_{2w}$  is less biased (mean bias of  $12.3 \pm 21.3 \mu atm$ ; Table 3) than bottle-calculated  $pCO_{2w}$ . The difference between the mean residual and median residual for the individual methods, as well as the high standard deviations (Table 3), suggests that outliers influence the general relationships, and further examination of the sources of these differences must be assessed.

Since this is a coastal region influenced by freshwater sources, likely with high spatial heterogeneity, we examine method discrepancies in  $pCO_{2w}$  using SSS (freshwater influence) and spatial

distance from the mooring (Figure 6). We initially assume that 1) the distance between where the bottle sample is taken and the closest temporal underway observation (no more than 2 min) is recorded is negligible; and 2) a salinity difference could be due to sensor uncertainty and/or error, as well as vertical differences in the water column caused by stratification (different sampling depths) or horizontal differences caused by advection. Both SSS differences and distance can be a factor, yet there is no one clear pattern (Figure 6). For example, bottle samples taken at the mooring can have 0 to <0.6 SSS unit differences with 20 to 70 µatm  $pCO_{2w}$  discrepancies, while samples taken at a greater distance, with ~2 SSS unit differences, can have only ~10 µatm  $pCO_{2w}$  discrepancies (Figure 6D). There are similar results for underway observations (Figure 6B). Specifically for the difference between mooring and bottle  $pCO_{2w}$  there is a weak dependence explained by the SSS measured by the mooring ( $r^2 = 0.22$ ; p < 0.001; n = 57): at lower SSS mooring  $pCO_{2w}$  is greater than bottle-calculated  $pCO_{2w}$  (Figure 6F). The spatial distribution of SSS around the mooring is discussed in section S4 of Supporting Information (Figure S1).

Differences in SST between the bottle samples, the mooring system, and the underway systems may also contribute to discrepancies (Figure 7A). Adjusting calculated  $pCO_{2w}$  and underway  $pCO_{2w}$  to mooring SST (as well as adjusting using both SST and SSS) did not improve the strength of correlations (Figure 2S). Similar to  $pCO_{2w}$  discrepancies due to SSS, even smaller SST differences can have associated higher  $pCO_{2w}$  discrepancies between methods (Figure 7B), which are not always caused by thermodynamic differences (see Supporting Information section S4). Sources of  $pCO_{2w}$  discrepancies could include measurement error (e.g., system calibrations, contamination, ), measurement uncertainty due to problems with the DIC and TA methods (e.g., measurement precision or inclusion of organic alkalinity; Figure 3S), and/or spatial heterogeneity, and are summarized below and discussed in the Supporting Information section S4.

Using two examples of cruises we illustrate when validation efforts have provided good agreement between all methods and when they did not, in an attempt to improve future sensor validation

efforts in the SAB and other moorings. The first example, one of only a few less successful cruises, in Figure 8A, illustrates the spatial track of the October, 2011 cruise. During this diel period the pCO<sub>2w</sub> at the mooring ranged from 420.7 to 444.9 µatm, with the minimum occurring at approximately 22:00 UTC and the maximum at 15:00 UTC (Figure 8B). Bottle samples ranged from 383.3 to 409.2 µatm (Figure 8B). The mean residual between mooring and underway observations is  $2.5 \pm 20 \,\mu$ atm, with a median of 4.8  $\mu$  atm (n = 51) and between mooring and bottle the mean is -36 ± 11  $\mu$  atm with a median of -37  $\mu$  atm (n = 57). The spatial-temporal differences observed during this cruise represent distance from the mooring (up to ~0.9 km) as well as changes in  $pCO_{2w}$  not captured by the three hour frequency of mooring measurements (Figure 8A). It is also important to note the vertical difference in the water column, which may cause the greater  $pCO_{2w}$  difference between the bottle values at 4 m and those from the underway (2.5 m) and the mooring (0.6 m). In the second, more successful example, in Figure 9A, the average mooring pCO<sub>2w</sub> during ECOA July, 2015 is 446.4  $\pm$  0.4 µatm and the average bottle calculated value is 446.7  $\pm$  3.8 µatm (Figure 9B), which are not statistically different. The mean and median residual mooring-underway are  $-4.9 \pm 5.2 \mu$  atm and the mooring-bottle mean and median residual are  $< 0.1 \pm 3.8 \mu$  atm and 1.9 µ atm, respectively. Regardless of the spatial difference (up to 1.6 km and  $\sim$ 3 m depth), there is good agreement between the three methods during the ECOA 2015 cruise, which likely indicates spatial homogeneity as opposed to the October 2011 cruise, and will be discussed in more detail below.

# 5. Discussion

#### 5.1 *p*CO<sub>2w</sub> internal consistency and validation

We assess  $pCO_{2w}$  internal consistency of mooring, underway, and bottle samples using a combination of correlation results, residual values, and biases (Table 3). At this time we do not offer specific values for allowable agreement since the majority of the cruises were not designed with validation in mind, rather use this discussion not only to address the positive outcomes of internal consistency, but also challenges and how they can be addressed and improved in future work. In general,

internal consistency is comparable among the three methods, as well as to other studies, however, Figures 6 through 8 also illustrate the often problematic spatio-temporal heterogeneity of the system. First, we find that the strongest agreement with the lowest bias is between the mooring-underway pair (Table 3). Using two independent underway  $pCO_2$  systems on the same cruise, similar correlation results (r = 0.96) were obtained by a previous study (Ribas-Ribas et al., 2014). Our result may be slightly weaker than previously published correlations due to high spatial heterogeneity, which would not be expected when comparing two different systems on the same vessel as opposed to locations up to 2 km apart. Second, our mooring-bottle correlation (r = 0.91) is similar to the mooring-underway (r = 0.92), however, the RMSE greater for the mooring-bottle pairing (24.6 µatm) than the mooring-underway pairing (17.6 µatm), suggesting that while the mooring-bottle method is in good agreement, there is an overall potential for differences in these methods of an order of magnitude greater than system uncertainties (Schar et al., 2010; Sutton et al., 2014). Previous work has included Monte Carlo re-sampling to estimate an allowable range of uncertainty (Ribas-Ribas et al., 2014), however, this is not a practical approach for the present work since our sample size is limited and could bias results. In general, internal consistency with respect to the mooring is good and results provide knowledge needed for future work (discussed below).

We identify several potential sources of  $pCO_{2w}$  discrepancies within the data set, which are being used to improve present and future validation efforts at the GR mooring. Detailed analyses of the sources of differences among methods are discussed in the Supporting Information section S4, in summary:

- Poor analytical and storage methods for the DIC-TA pairing could contribute up to 10% relative mean uncertainty in calculated pCO<sub>2w</sub>, a value well above internationally accepted ocean acidification monitoring network best practices (Newton et al., 2014). Poor storage could result in increased pCO<sub>2w</sub> in bottle samples due to respiration, or decreased pCO<sub>2w</sub> due to degassing.
- Spatio-temporal SSS during the two cruises highlighted (October 2011 and ECOA-2015) only had mooring-underway SSS difference of up to  $\pm$  0.2, however, including all cruises the

maximum SSS difference was up to ~2.3 (Figure 6B and 6D). If for every theoretical 1 unit SSS changes, the TA changes ~50  $\mu$ mol kg<sup>-1</sup> (Xue et al., 2016), then the 0.2 or 2.3 unit difference could account for a TA difference ranging from ~10 to 115  $\mu$ mol kg<sup>-1</sup>, respectively. Therefore, potentially affecting the *p*CO<sub>2w</sub> value of one method with respect to the other.

- Assuming that sensor differences or environmental differences represented by differing SST could contribute to discrepancies in *p*CO<sub>2w</sub> we adjust underway and bottle SST to mooring SST (Equation 1 of Supporting Information; Takahashi et al., 1993). This adjustment, however, decreases the overall strength of correlation results, though in some individual instances does decrease the discrepancies between methods (Figure S2). Therefore, environmental or sensor differences could affect some cruise results but in general are not likely the primary source of discrepancies in the present work.
- On a coastal margin influenced by freshwater, SSS spatio-temporal heterogeneity may also be important and indicates if different water masses are sampled. Since SSS only explains up to 87% of TA variability in bottle samples (*p* < 0.001), a portion of TA could be due to organic bases, thus overestimating TA and subsequently underestimating *p*CO<sub>2w</sub> (Yang et al., 2015). Decreasing measured TA in this study 10 to 30 µmol kg<sup>-1</sup> increased the correlation strength for mooring-bottle (Figure 3S) as well as reduced the RMSE and residual values. Therefore, organic bases in the TA method, resulting in systematic biases in calculated *p*CO<sub>2w</sub>, could skew results on coastal margins. Given this conclusion, future validation efforts could include deeper investigation on the role of organic bases in TA measurements at the GR mooring.

While at this time we do not seek to examine environmental details of every individual cruise for sources of internal discrepancies in calculated  $pCO_{2w}$ , we note that these previous validation efforts are an important step in better understanding spatio-temporal CO<sub>2</sub> dynamics at this particular field site. The  $pCO_{2w}$  differences observed at this site could also be due to the influence of tidal or temperature fronts (advection of coastal or oceanic water masses) on very small spatial and temporal time scales (Figures 6-9). Future validation attempts, if they cannot be performed within meters and minutes of the mooring observation, should consider tides and also attempt to identify any different water masses around the mooring that could influence the results. Spatial differences are more likely to occur using mooring-underway validation, whereas, temporal differences are less likely to be responsible for  $pCO_{2w}$  differences given the high temporal frequency of underway observations. Conversely, given that we know the mooring samples at 17 minutes after the hour, though the observation is integrated over a 6 min equilibration time, temporal differences should be less of an issue if samples are taken at the mooring.

The magnitude of the mooring-bottle results are similar to those determined for a MAPCO<sub>2</sub> system in the Hood Canal (Washington, USA), however, in that instance calculated pCO<sub>2w</sub> was greater than the moored value (-12  $\pm$  30  $\mu$ atm), and the difference between mooring and underway pCO<sub>2w</sub> was better (-9  $\pm$  8 µatm) (Schar et al., 2010). Consistent with the aforementioned study we find that there is better agreement between mooring-underway than mooring-bottle using the DIC-TA pairing, which could be due to analytical challenges using calculated  $pCO_{2w}$ . A reasonable understanding of the bottlemooring pair is needed in order to use SSS-derived TA and  $pCO_{2w}$  to accurately estimate the rest of the carbonate system variables at the GR mooring (Sutton et al., 2016) in future works. Once we are able to calculate all carbonate system variables, with a good sense of validity through internal consistency, then we can apply biogeochemical and/or more sophisticated physical-biological coupled models in the future to better understand the sources of pCO<sub>2w</sub> variability on this coastal margin. This initial validation is a first step in showing that reasonable internal consistency has been achieved at this mooring and that GR mooring time series measurements are good quality and can be used to assess  $pCO_{2w}$  at this site. Given how measurement methods have changed over the last decade, as OA and carbonate system variable observations have become crucial, this is an important milestone in carbonate system studies. The next step in future time series research is to compare OA monitoring site results from present studies to those conducted prior to the initiation of our present MAPCO<sub>2</sub> systems as well as to historic observations.

#### 5.2 Seasonal to interannual *p*CO<sub>2w</sub> and *F*CO<sub>2</sub> variability

Intra-annual  $pCO_{2w}$  variation follows a seasonal thermal cycle, with increased partial pressure in the summer compared to the winter. When  $pCO_{2w}$  is decomposed,  $pCO_{2t}$  and  $pCO_{2nt}$  exhibit out-of-phase seasonal harmonics that roughly coincide with SST and river stream flow cycles, respectively (Figures 2 and 3). The almost identical phases of  $pCO_{2w}$  and  $pCO_{2t}$  (or SST), as well as the greater seasonal amplitude of  $pCO_{2t}$  compared to  $pCO_{2w}$ , support a SST-dominated seasonal cycle (Figure 3B and 3C). Rather than examining well defined  $pCO_{2w}$  seasonal thermal cycles in the SAB, as has recently been done (e.g., Signorini et al., 2013; Xue et al., 2016), we assess the potential contributions to seasonal  $pCO_{2nt}$ variability and interannual changes.

 $pCO_{2nt}$  variability, which is out of phase with observed  $pCO_{2w}$ , reflects the overall influence of non-thermal processes: vertical and/or horizontal mixing (river stream flow), net biological reactions, and air-sea exchange (Xue et al., 2016). GR mooring  $pCO_{2nt}$ , which is increased in the winter and early spring with the river discharge, decreases as spring time Chl *a* increases (Figure 3). The timing of  $pCO_{2nt}$ variability strongly suggest that the initial  $pCO_{2nt}$  increase is a response to terrestrial CO<sub>2</sub> exported from the Altamaha River (Figure 3), as a proxy for all freshwater sources to the SAB. Even though riverine DIC concentration decreases (dilution) with increased river stream flow (Figure 3A), not only would the overall lateral flux out of the watershed be enhanced but, the CO<sub>2</sub> to DIC ratio would increase as well (Jiang et al., 2013) at higher stream flow. Increased river discharge also mobilizes material (OM, CO<sub>2</sub>, DIC, and inorganic nutrients) export fluxes from salt marshes via tidal flushing (Cai, 2011), thereby enhancing lateral fluxes across the SAB inner shelf. Consequently, exported OM likely provides a sustainable source for CO<sub>2</sub> via respiration on the shelf (Cai et al., 2003; Jiang et al., 2013, 2010; Pomeroy et al., 2000) possibly even after river stream flow decreases.

Another consequence of increased river discharge in the winter and early spring is an increase in nutrient export from the watershed, estuaries, and marshes (Figure 3) (Cai et al., 2003; Jiang et al., 2013; Schaefer and Alber, 2007). We suggest that nutrient export is the cause of increased Chl *a*, an indicator of

biomass, during the spring and early summer at the GR mooring starting from when spring SST warming begins. As a result in the spring-early summer  $pCO_{2nt}$  decreases following the Chl *a* peak as biological production consumes CO<sub>2</sub> (Figure 3B). Overall  $pCO_{2w}$ , however, increases due to warming  $(pCO_{2t})$ , leading to degassing and positive  $FCO_2$  (Figure 4), therefore,  $pCO_{2nt}$  continues to decrease in the warm summer months due to the switch from negative to positive  $FCO_2$ . Finally, in the fall-winter  $pCO_{2nt}$ increases once again due to lateral transport and respiration of OM exported from estuaries earlier in the season, as well as heterotrophy on the shelf (Pomeroy et al., 2000). When SST cools,  $FCO_2$  returns to negative, which also contributes to the  $pCO_{2nt}$  increase via atmospheric CO<sub>2</sub> uptake (Figures 2 and 4). While the aforementioned processes all contribute to overall  $pCO_{2w}$  and  $pCO_{2nt}$  their relative importance likely varies season-to-season and interannually, as was reported by *Xue et al.* (2016).

# 5.3 Sub-decadal CO<sub>2</sub> trends and larger scale implications

pCO<sub>2w</sub> increases over the time series, though still on a sub-decadal scale, by 5.2 ± 1.4 µatm y<sup>-1</sup>. The large data loss in 2009 is in the middle of the time series, and therefore does not greatly influence the long-term trend. The pCO<sub>2w</sub> increase is more than double the pCO<sub>2a</sub> increase at this coastal site, as well as the rate at the Bermuda Atlantic Time Series (BATS) station (1.7 ± 0.09 µatm y<sup>-1</sup>) (Bates et al., 2014, 2012) and the global mean pCO<sub>2a</sub> ~2.0 µatm y<sup>-1</sup> (Wanninkhof et al., 2013). Such increases in pCO<sub>2w</sub> have also been noted in the northern South China Sea, which is increasing at a rate slightly higher than that of the atmosphere (Tseng et al., 2007). Also, pCO<sub>2w</sub> in the North Sea between two summers is increasing at a rate 5-6 times the atmospheric, thus decreasing the net regional sink (Thomas et al., 2007), which could also be the case at the GR mooring and SAB inner shelf. Furthermore, at the GR mooring there is a 7.1 ± 1.9 µatm y<sup>-1</sup> pCO<sub>2nt</sub> increase. Consistent with previous studies in the North Atlantic basin (Fay and McKinley, 2013; McKinley et al., 2011), we find that GR mooring pCO<sub>2w</sub> change over the course of our time series does not reflect an increase in SST. While SST influences daily and seasonal pCO<sub>2w</sub> variability in the SAB (Signorini et al., 2013; Xue et al., 2016), we find no statistically significant change over the time series in SST or pCO<sub>2t</sub> (Figure 2B and 2C), therefore the change is likely due only to

 $pCO_{2nt}$ . Since neither of the variables associated with mixing, Altamaha River stream flow and SSS, exhibit statistically significant increases, we can conclude from the available observations that the increasing river DIC concentration, and possibly the DIC flux from rivers and marshes to the SAB, is a potential source for increasing  $pCO_{2w}$  at the GR mooring (Figure 3A; Table 1). There is no evidence of changes in biological production (Chl *a*) at the GR mooring, however, without an estimate for biological respiration or net primary production, we cannot estimate the potential role of DIC produced on the inner shelf.

As a result of increased  $pCO_{2w}$ , which is greater than the atmospheric increase,  $\Delta pCO_2$  and  $FCO_2$ are also increasingly more positive and could result in a decreased annual net sink. These results are consistent with findings on the Scotia Shelf, North America, which is an increasing net source of  $CO_2$  to the atmosphere (Shadwick et al., 2010). The  $pCO_{2a}$  increase at the GR mooring likely contributes up to ~2.0 µatm y<sup>-1</sup> of the overall  $pCO_{2w}$  increase, however, other sources must account for the remaining portion. Based on the elimination of factors, such as a lack of increasing trends in SSS, SST, and Chl *a*, we suggest that further investigation into the shelf DIC and net biological processes in the coastal marshes and rivers could help resolve the source of increased GR mooring  $pCO_{2w}$ .

Compared with open ocean  $pCO_{2w}$ , coastal  $pCO_{2w}$ , and its variability, are more strongly influenced by terrestrial export and episodic events, such as storm induced vertical mixing over a shallow water column or mixed layer, than in the open ocean (Gledhill et al., 2015). The confidence interval associated with the  $pCO_{2w}$  increase over ~8.5 years at the GR mooring is greater than those reported for open ocean sites in the Atlantic (Bates et al., 2014, 2012; Fay and McKinley, 2013; McKinley et al., 2011), and could be interpreted as higher coastal zone variability compared with the open ocean. The observed trend variability is also likely influenced by the start and end dates of the time series; that is the time series starts during a dry period and ends during a wet period. Episodic events, such as wet periods (e.g., 2009-2010), drought periods (e.g., 2012), and tropical cyclones (surface water cooling or mixing) in the southeast USA (Crosswell et al., 2014) can all contribute to large trend variability. In Section 5.1 we assessed the influence of lateral materials flux via river stream flow, though apart from river stream flow, other extreme mixing variations could transport upwelled oceanic cold intrusions onshore (see Section 5.3). Episodic events increase variability and make it challenging to determine the underlying mechanisms that control sub-decadal trends; however, a greater understanding of mixing at the GR mooring could constrain this particular uncertainty. Finally, carrying the time series out to over two decades could also decrease the overall trend and its variability (Fay and McKinley, 2013), which highlights the necessity of continued observations at this and other coastal moorings.

## 5.4 Net annual FCO<sub>2</sub> source-sink variability

Considering the data loss during most years, if we only compare years with minimal data loss (2008, 2010, and 2012; Table 2), then there is still no consensus on whether this site is consistently a source or a sink. Rather, the site is highly variable and there is still a great amount of uncertainty surrounding the annual  $FCO_2$  (Figure 5). As we find for year 2012, a previous assessment of the SAB in 2005 and 2006 also determined that the inner shelf could be a source of  $CO_2$  to the atmosphere (Jiang et al., 2008b). The observed interannual variability and uncertainty in FCO<sub>2</sub> in this work, and differences with estimates from previous studies of the SAB, could be due the complex hydrographic regime at the boundary of the inner and middle shelf where the GR mooring is located (Castelao, 2011; Menzel, 1993; Signorini and McClain, 2007). Whether the region represented by the GR mooring is an annual source or sink is determined by the hydrography, including extent of the freshwater plume, the location of the Gulf Stream, and/or Gulf Stream-induced upwellings. For example, decreased river stream flow (year 2012) could contribute by increasing the residence time of water on the inner shelf (Menzel, 1993), thus allowing for a longer period of time over which laterally transported OM, and subsequently respiration, could remain on the inner shelf near the GR mooring and enhance pCO<sub>2nt</sub> (Figure 3). Another example, during 2007 transport/Gulf Stream-induced upwelling of lower SST waters with higher TA and nutrient concentrations than shelf waters, could have decreased  $pCO_{2w}$ , which could lead to negative  $FCO_2$ , while providing nutrients for Chl a increases (Signorini et al., 2013). A short term return to negative FCO<sub>2</sub>

accompanied by a decrease in SST during the later spring in 2007 could be evidence of the effects of onshelf transport, thus the observed decreases in  $pCO_{2t}$  and  $pCO_{2nt}$  with a several month increase in Chl *a* (Figures 2B & 2C, 3B, and 4A, 4B, & 4C). Even short-term hydrographic changes could impact annual sums and uncertainty, though a more thorough understanding of the mixing regime at the GR mooring is need to better assess the quantitative impacts on interannual *FCO*<sub>2</sub> and the alternating source/sink.

Another factor that could contribute to large uncertainties in annual FCO<sub>2</sub> is the use of various gas transfer velocity estimates. We find that our estimates using two different  $FCO_2$  parameterizations do not always agree on the sign of  $FCO_2$  (Figure 5). For example, the W09 hybrid model closely resembles a quadratic function, which shows larger differences at higher winds speeds compared with the MG01 cubic function, and there are also slight differences in the intercepts (Equations 4 and 5). For years with a greater number of days with higher U<sub>10</sub>, this may be an important distinction and is likely a key source for daily to interannual uncertainty since at higher wind speeds there are greater differences between the resulting fluxes (Figure 4A & 4B). The large error bars (annual uncertainty) in Figure 5 are due to accumulation of uncertainty (daily variability in  $pCO_{2w}$ ). Furthermore, from a methodological standpoint, larger data gaps in 2007, 2009, 2011, 2013, and 2014 (see Table 2 for the number of missing days in each year) will also affect annual fluxes. For comparison, the Jiang et al. (2008a) study presented estimates using a quadratic wind speed relationship with a zero intercept (Wanninkhof, 1992), thus, the results of the present study should generally be comparable; therefore methodologies between studies are not likely contributing to differences in our source/sink estimates. In general, the alternating annual source-sink and disagreement between the  $FCO_2$  parameterizations at the site has implications for estimating annual carbon budgets in this region and could contribute to overall regional budget uncertainties.

# 6. Conclusions

We have shown that MAPCO<sub>2</sub> system, underway, and calculated  $pCO_{2w}$  from discrete water samples at the GR mooring are in reasonably good agreement, even though the majority of the cruise observations compiled over the time series of the GR mooring were not collected using rigorous internal consistency methods. While there are some discrepancies in the validation results, general uncertainties associated with the method can be quantified, thus improving data synthesis. We recommend that samples collected for internal consistency should be collected as close to the mooring equilibrator intake as possible and within minutes of the MAPCO<sub>2</sub> system observation. If samples cannot be analyzed immediately while the cruise is underway, then they should be treated and stored according to established protocols and analyzed as soon as possible (Dickson et al., 2007). While DIC and TA are the most commonly and easily collected and analyzed carbonate system variables, this pairing may not be the most accurate or precise method for calculating  $pCO_{2w}$ , especially on highly variable ocean margins influenced by lower SSS waters. The broader implications for validation and time series assessments will allow future studies to merge the methods for determining  $pCO_{2w}$  variability explored here, and can also contribute to discerning the biogeochemical drivers of CO<sub>2</sub> on highly variable coastal margins.

SST is an important source of daily to seasonal  $pCO_{2w}$  and  $FCO_2$  variability; however, longer term increases in  $pCO_{2w}$  are likely driven by  $pCO_{2nt}$  at the GR mooring. The combined effects of all non-thermal processes (net biological processes, air-sea exchange, and transport) likely contribute to the increase in  $pCO_{2nt}$ , thus  $pCO_{2w}$ . Further work is needed to calculate time series of DIC and TA at the GR mooring, which can be used in biogeochemical models to determine the contributions of mixing, biological processes, and gas exchange.

Depending on the parameterization used, the net annual  $FCO_2$  at the GR mooring can have large differences in the source-sink term and can also vary greatly year-to-year. According to the estimates presented here, the majority of the time the GR mooring site is a sink for atmospheric CO<sub>2</sub>, though the mixing regime likely plays an important role in annual estimates, which we have yet to discern. The results of the present study differ from previous estimates, which were based on seasonal cruises and interpolation rather than a continuous time series, assuming that the GR mooring is a good representation of the SAB inner shelf. There appears to be a clear need to re-evaluate  $FCO_2$  on the entire shelf, as well as determine the biogeochemical processes that influence  $pCO_{2w}$ .

# Acknowledgments

The Gray's Reef  $CO_2$  mooring effort has been supported initially by the National Oceanic and Atmospheric Administration's Global Carbon Cycle (NOAA-GCC) and now NOAA–Ocean Acidification Program (OAP) via the South East Coastal Ocean Observing Regional Association (SECOORA); cooperative agreement number: NA11NOS0120033. JJR's time has been supported by both NOAA OAP and a University of Delaware internal fund (to WJC). XH is supported by Texas Research Development Fund. LX thanks the China Scholarship Council for providing the state scholarship fund (No. 201404180014). We thank the Georgia Coastal Ecosystems Long Term Ecological Research network for assistance in ground-truthing the mooring. Our late colleague Yongchen Wang was instrumental in the initial deployment of the mooring sensors and for ground-truthing the sensor data. Finally, we thank Sarah Fangman, Jared Halonen, Todd Recicar, and Randy Rudd at the Gray's Reef National Marine Sanctuary for providing vessel and diver support to help maintain the mooring operation. The authors also thank Dr. Todd Martz for very useful comments on an early version of this manuscript. This is PMEL contribution 4289 and JISAO contribution 2414. WJC would also like to thank NASA support on coastal carbon synthesis efforts (NNX11AD47G).

All observations used in this research are available to the public through the Carbon Dioxide Information Analysis Center (CDIAC; http://cdiac.ornl.gov/), the NASA Giovanni Server (http://disc.sci.gsfc.nasa.gov/giovanni), and the Georgia Coastal Ecosystems Long Term Research Network (http://gce-lter.marsci.uga.edu/portal/usgs\_doctortown/historic/data/index.xml).

# References

- Bates, N., Astor, Y., Church, M., Currie, K., Dore, J., González-Davila, M., Lorenzoni, L., Muller-Karger, F., Olafsson, J., Santana-Casiano, J., 2014. A Time-Series View of Changing Ocean Chemistry Due to Ocean Uptake of Anthropogenic CO2 and Ocean Acidification. Oceanography 27, 126–141.
- Bates, N.R., Best, M.H.P., Neely, K., Garley, R., Dickson, A.G., Johnson, R.J., 2012. Detecting anthropogenic carbon dioxide uptake and ocean acidification in the North Atlantic Ocean. Biogeosciences 9, 2509–2522.
- Bauer, J.E., Cai, W., Raymond, P.A., Bianchi, T.S., Hopkinson, C.S., Regnier, P.A.G., 2013. The changing carbon cycle of the coastal ocean. Nature 504, 61–70.
- Blanton, J., Atkinson, P., 1983. Transport and Fate of River Discharge on the Continental Shelf of the Southeastern United States. J. Geophys. Res. 88, 4730–4738.
- Boehme, S.E., Sabine, C.L., Reimers, C.E., 1998. CO2 fluxes from a coastal transect: a time-series approach. Mar. Chem. 63, 49–67.
- Cai, W.-J., 2011. Estuarine and Coastal Ocean Carbon Paradox: CO 2 Sinks or Sites of Terrestrial Carbon Incineration? Ann. Rev. Mar. Sci. 3, 123–145.
- Cai, W.-J., Hu, X., Huang, W.-J., Jiang, L.-Q., Wang, Y., Peng, T.-H., Zhang, X., 2010. Alkalinity distribution in the western North Atlantic Ocean margins. J. Geophys. Res. 115.
- Cai, W.-J., Hu, X., Huang, W.-J., Murrell, M.C., Lehrter, J.C., Lohrenz, S.E., Chou, W.-C., Zhai, W., Hollibaugh, J.T., Wang, Y., Zhao, P., Guo, X., Gundersen, K., Dai, M., Gong, G.-C., 2011. Acidification of subsurface coastal waters enhanced by eutrophication. Nat. Geosci. 4, 766–770.
- Cai, W.-J., Wang, Y., 1998. The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. Limnol. Oceanogr. 43, 657–668.
- Cai, W.-J., Wang, Y., Hodson, R.E., 1998. Acid-Base Properties of Dissolved Organic Matter in the Estuarine Waters of Georgia, USA. Geochim. Cosmochim. Acta 62, 473–483.
- Cai, W.-J., Wang, Z.A., Wang, Y., 2003. The role of marsh-dominated heterotrophic continental margins in transport of  $CO_2$  between the atmosphere, the land-sea interface and the ocean. Geophys. Res. Lett. 30, 1–4.
- Castelao, R., 2011. Intrusions of Gulf Stream waters onto the South Atlantic Bight shelf. J. Geophys. Res. 116.
- Castelao, R.M., 2014. Mesoscale eddies in the South Atlantic Bight and the Gulf Stream Recirculation region: Vertical structure. J. Geophys. Res. Ocean. 119, 2048–2065.
- Castelao, R.M., He, R., 2013. Mesoscale eddies in the South Atlantic Bight. J. Geophys. Res. Ocean. 118, 5720–5731.
- Crosswell, J.R., Wetz, M.S., Hales, B., Paerl, H.W., 2014. Extensive CO<sub>2</sub> emissions from shallow coastal waters during passage of Hurricane Irene (August 2011) over the Mid-Atlantic Coast of the U.S.A. 59, 1651–1665.

- Dickson, A.G., Millero, F.J., 1987. A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media. Deep Sea Res. Part A. Oceanogr. Res. Pap. 34, 1733–1743.
- Dickson, A.G., Sabine, C.L., Christian, J.R., 2007. Guide to Best Practices for Ocean CO<sub>2</sub> Measurements. PICES Special Publication 3.
- Fay, A.R., McKinley, G.A., 2013. Global trends in surface ocean p CO<sub>2</sub> from in situ data. Global Biogeochem. Cycles 27, 541–557.
- Feely, R. a, Sabine, C.L., Lee, K., Berelson, W., Kleypas, J., Fabry, V.J., Millero, F.J., 2004. Impact of anthropogenic CO<sub>2</sub> on the CaCO<sub>3</sub> system in the oceans. Science 305, 362–6.
- Feely, R.A., Sabine, C.L., Hernandez-ayon, J.M., Ianson, D., Hales, B., 2008. Evidence for upwelling of corrosive "acidified" water onto the continental shelf. Science 320, 1490–1492.
- Gledhill, D.K., White, M.M., Salisbury, J., Thomas, H., Mlsna, I., Liebman, M., Mook, B., Grear, J., Candelmo, A.C., Chambers, R.C., Gobler, C.J., Hunt, C.W., King, A.L., Price, N.N., Signorini, S.R., Stancioff, E., Stymiest, C., Wahle, R.A., Waller, J.D., Rebuck, N.D., Wang, Z.A., 2015. Ocean and coastal acidification off New England and Nova Scotia. Oceanography 28, 182–197.
- Huang, W.-J., Cai, W.-J., Castelao, R.M., Wang, Y., Lohrenz, S.E., 2013. Effects of a wind-driven crossshelf large river plume on biological production and CO<sub>2</sub> uptake on the Gulf of Mexico during spring. Limnol. Oceanogr. 58, 1727–1735.
- Huang, W.-J., Wang, Y., Cai, W.-J., 2012. Assessment of sample storage techniques for total alkalinity and dissolved inorganic carbon in seawater. Limnol. Oceanogr. Methods 10, 711–717.
- Jiang, L.-Q., Cai, W.-J., Wang, Y., 2008a. A comparative study of carbon dioxide degassing in river- and marine-dominated estuaries. Limnol. Oceanogr. 53, 2603–2615.
- Jiang, L.-Q., Cai, W.-J., Wang, Y., Bauer, J.E., 2013. Influence of terrestrial inputs on continental shelf carbon dioxide. Biogeosciences 10, 1–22.
- Jiang, L.-Q., Cai, W.-J., Wang, Y., Diaz, J., Yager, P.L., Hu, X., 2010. Pelagic community respiration on the continental shelf off Georgia, USA. Biogeochemistry 98, 101–113.
- Jiang, L.-Q., Cai, W.-J., Wanninkhof, R., Wang, Y., Lüger, H., 2008b. Air-sea CO<sub>2</sub> fluxes on the U.S. South Atlantic Bight: Spatial and seasonal variability. J. Geophys. Res. 113, C07019.
- Le Quere, C., Andrew, R.M., Canadell, J.G., Sitch, S., Ivar Korsbakken, J., Peters, G.P., Manning, A.C., Boden, T.A., Tans, P.P., Houghton, R.A., Keeling, R.F., Alin, S., Andrews, O.D., Anthoni, P., Barbero, L., Bopp, L., Chevallier, F., Chini, L.P., Ciais, P., Currie, K., Delire, C., Doney, S.C., Friedlingstein, P., Gkritzalis, T., Harris, I., Hauck, J., Haverd, V., Hoppema, M., Klein Goldewijk, K., Jain, A.K., Kato, E., Kortzinger, A., Landschutzer, P., Lefevre, N., Lenton, A., Lienert, S., Lombardozzi, D., Melton, J.R., Metzl, N., Millero, F., Monteiro, P.M.S., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.I., O'Brien, K., Olsen, A., Omar, A.M., Ono, T., Pierrot, D., Poulter, B., Rodenbeck, C., Salisbury, J., Schuster, U., Schwinger, J., Seferian, R., Skjelvan, I., Stocker, B.D., Sutton, A.J., Takahashi, T., Tian, H., Tilbrook, B., Van Der Laan-Luijkx, I.T., Van Der Werf, G.R., Viovy, N., Walker, A.P., Wiltshire, A.J., Zaehle, S., 2016. Global Carbon Budget 2016. Earth Syst. Sci. Data 8, 605–649.

- Lee, K., Kim, T.W., Byrne, R.H., Millero, F.J., Feely, R.A., Liu, Y.M., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochim. Cosmochim. Acta 74, 1801– 1811.
- Lewis, E., Wallace, D., 1998. Program developed for CO<sub>2</sub> system calculations. ORNL/CDIAC 105, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory US Department of Energy, Oak Ridge, Tennessee.
- Lueker, T.J., Dickson, A.G., Keeling, C.D., 2000. Ocean pCO<sub>2</sub> calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO<sub>2</sub> in gas and seawater at equilibrium. Mar. Chem. 70, 105–119.
- McGillis, W.R., Edson, J.B., Hare, J.E., Fairall, C.W., 2001. Direct covariance air-sea CO<sub>2</sub> fluxes. J. Geophys. Res. 106, 16729–16745.
- McKinley, G. a., Fay, A.R., Takahashi, T., Metzl, N., 2011. Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales. Nat. Geosci. 4, 606–610.
- Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicz, R.M., 1973. Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure. Limnol. Oceanogr. 18, 897–907.
- Menzel, D., 1993. Ocean processes: US southeast continental shelf. US Dep. Energy, Off. Tech. Inf. DOE/OSTI-1, 121.
- Michalak, A.M., Jackson, R.B., Marland, G., Sabine, C.L., Group, C.C.W., 2011. A U.S. Carbon Cycle Science Plan. Nova Science Publishers, Inc.
- Newton, J.A., Feely, R.A., Jewett, E.B., Williamson, P., Mathis, J., 2014. Global Ocean Acidification Observing Network: Requirements and Governance Plan.
- Patsavas, M.C., Byrne, R.H., Wanninkhof, R., Feely, R. a., Cai, W.-J., 2015. Internal Consistency of Marine Carbonate System Measurements and Assessments of Aragonite Saturation State: Insights from Two U.S. Coastal Cruises. Mar. Chem. 176, 9–20.
- Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T., Olsen, A., Feely, R. a., Cosca, C.E., 2009. Recommendations for autonomous underway pCO<sub>2</sub> measuring systems and data-reduction routines. Deep Sea Res. Part II Top. Stud. Oceanogr. 56, 512–522.
- Pomeroy, L.R., Sheldon, J.E., Sheldon, W.M., Blanton, J.O., Amft, J., Peters, F., 2000. Seasonal Changes in Microbial Processes in Estuarine and Continental Shelf Waters of the South-eastern U.S.A. Estuar. Coast. Shelf Sci. 51, 415–428.
- Ribas-Ribas, M., Rerolle, V.M.C., Bakker, D.C.E., Kitidis, V., Lee, G.A., Brown, I., Achterberg, E.P., Hardman-Mountford, N.J., Tyrrell, T., 2014. Intercomparison of carbonate chemistry measurements on a cruise in northwestern European shelf seas. Biogeosciences 11, 4339–4355.
- Sabine, C.L., Feely, R. a, Gruber, N., Key, R.M., Lee, K., Bullister, J.L., Wanninkhof, R., Wong, C.S., Wallace, D.W.R., Tilbrook, B., Millero, F.J., Peng, T.-H., Kozyr, A., Ono, T., Rios, A.F., 2004. The oceanic sink for anthropogenic CO<sub>2</sub>. Science 305, 367–71.
- Schaefer, S.C., Alber, M., 2007. Temporal and spatial trends in nitrogen and phosphorus inputs to the watershed of the Altamaha River, Georgia, USA. Biogeochemistry 86, 231–249.

- Schar, D., Atkinson, M., Johengen, T., Pinchuk, A., Purcell, H., Robertson, C., Smith, G.J., Tamburri, M., 2010. Performance demonstration statement PMEL MAPCO2/Battelle Seaology pCO<sub>2</sub> monitoring system.
- Shadwick, E.H., Thomas, H., Azetsu-Scott, K., Greenan, B.J.W., Head, E., Horne, E., 2011. Seasonal variability of dissolved inorganic carbon and surface water pCO<sub>2</sub> in the Scotian Shelf region of the Northwestern Atlantic. Mar. Chem. 124, 23–37.
- Shadwick, E.H., Thomas, H., Comeau, A., Craig, S.E., Hunt, C.W., Salisbury, J.E., 2010. Air-sea CO<sub>2</sub> fluxes on the Scotian Shelf: seasonal to multi-annual variability. Biogeosciences 7, 3851–3867.
- Sheldon, J.E., Alber, M., 2013. Effects of Climate Signals on River Discharge to Ossabaw, St. Andrew, and Cumberland Sounds. Georgia Coastal Reseach Council.
- Sheldon, J.E., Burd, A.B., 2013. Alternating effects of climate drivers on Altamaha River discharge to coastal Georgia, USA. Estuaries and Coasts 37, 772–788.
- Signorini, S.R., Mannino, A., Najjar, R.G., Friedrichs, M. a. M., Cai, W.-J., Salisbury, J., Wang, Z.A., Thomas, H., Shadwick, E., 2013. Surface ocean pCO<sub>2</sub> seasonality and sea-air CO<sub>2</sub> flux estimates for the North American east coast. J. Geophys. Res. Ocean. 118, 5439–5460.
- Signorini, S.R., McClain, C.R., 2007. Large-scale forcing impact on biomass variability in the South Atlantic Bight. Geophys. Res. Lett. 34, L21605.
- Sunda, W.G., Cai, W.-J., 2012. Eutrophication induced CO<sub>2</sub>-acidification of subsurface coastal waters: interactive effects of temperature, salinity, and atmospheric PCO<sub>2</sub>. Environ. Sci. Technol. 46, 10651–9.
- Sutton, A.J., Sabine, C.L., Feely, R.A., Cai, W.-J., Cronin, M.F., McPhaden, M.J., Morell, J.M., Newton, J.A., Noh, J.-H., Olafsdottir, S.R., Salisbury, J.E., Send, U., Vandemark, D.C., Weller, R.A., 2016. Using present-day observations to detect when anthropogenic change forces surface ocean carbonate chemistry outside pre-industrial bounds. Biogeosciences, 13, 5065–5083.
- Sutton, A.J., Sabine, C.L., Maenner-Jones, S., Lawrence-Slavas, N., Meinig, C., Feely, R. a., Mathis, J.T., Musielewicz, S., Bott, R., McLain, P.D., Fought, J., Kozyr, A., 2014. A high-frequency atmospheric and seawater pCO<sub>2</sub> data set from 14 open ocean sites using a moored autonomous system. Earth Syst. Sci. Data Discuss. 7, 385–418.
- Takahashi, T., Olafsson, J., Goddard, J., Chipman, D., Sutherland, S., 1993. Seasonal variation of CO<sub>2</sub> and nutrients in the high-latitude surface oceans: A comparative study. Global Biogeochem. Cycles 7, 843–878.
- Takahashi, T., Sutherland, S.C., Chipman, D.W., Goddard, J.G., Cheng, H., Newberger, T., Sweeney, C., Munro, D.R., 2014. Climatological distributions of pH, pCO<sub>2</sub>, total CO<sub>2</sub>, alkalinity, and CaCO<sub>3</sub> saturation in the global surface ocean, and temporal changes at selected locations. Mar. Chem. 164, 95–125.
- Takahashi, T., Sutherland, S.C., Feely, R. a., Wanninkhof, R., 2006. Decadal change of the surface water pCO<sub>2</sub> in the North Pacific: A synthesis of 35 years of observations. J. Geophys. Res. Ocean. 111, 1-20.

- Takahashi, T., Sutherland, S.C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. a., Sabine, C., Olafsson, J., Nojiri, Y., 2002. Global sea–air CO<sub>2</sub> flux based on climatological surface ocean pCO<sub>2</sub>, and seasonal biological and temperature effects. Deep Sea Res. Part II Top. Stud. Oceanogr. 49, 1601–1622.
- Thomas, H., Friederike Prowe, A.E., van Heuven, S., Bozec, Y., de Baar, H.J.W., Schiettecatte, L.-S., Suykens, K., Koné, M., Borges, A. V., Lima, I.D., Doney, S.C., 2007. Rapid decline of the CO<sub>2</sub> buffering capacity in the North Sea and implications for the North Atlantic Ocean. Global Biogeochem. Cycles 21, 1–13.
- Tseng, C.-M., Wong, G.T.F., Chou, W.-C., Lee, B.-S., Sheu, D.-D., Liu, K.-K., 2007. Temporal variations in the carbonate system in the upper layer at the SEATS station. Deep Sea Res. Part II Top. Stud. Oceanogr. 54, 1448–1468.
- Uppstrom, L.R., 1974. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. Deep. Res. Oceanogr. Abstr. 21, 161–162.
- Vandemark, D., Salisbury, J.E., Hunt, C.W., Shellito, S.M., Irish, J.D., McGillis, W.R., Sabine, C.L., Maenner, S.M., 2011. Temporal and spatial dynamics of CO<sub>2</sub> air-sea flux in the Gulf of Maine. J. Geophys. Res. 116, C01012.
- Wang, Z.A., Wanninkhof, R., Cai, W.-J., Byrne, R.H., Hu, X., Peng, T.-H., Huang, W.-J., 2013. The marine inorganic carbon system along the Gulf of Mexico and Atlantic coasts of the United States : Insights from a transregional coastal carbon study. Limnol. Oceanogr. 58, 325–342.
- Wanninkhof, R., 1992. Relationship between wind speed and gas exchange over the ocean. J. Geophys. Res. Oceans. 97, 7373–7382.
- Wanninkhof, R., 2014. Relationship between wind speed and gas exchange over the ocean revisited. Limnol. Oceanogr. Methods 12, 351–362.
- Wanninkhof, R., Asher, W.E., Ho, D.T., Sweeney, C., McGillis, W.R., 2009. Advances in quantifying air-sea gas exchange and environmental forcing. Ann. Rev. Mar. Sci. 1, 213–244.
- Wanninkhof, R., Barbero, L., Byrne, R., Cai, W.J., Huang, W.J., Zhang, J.Z., Baringer, M., Langdon, C., 2015. Ocean acidification along the Gulf Coast and East Coast of the USA. Cont. Shelf Res. 98, 54-71.
- Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S.C., McKinley, G. a., Lenton, a., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., Khatiwala, S., 2013. Global ocean carbon uptake: magnitude, variability and trends. Biogeosciences 10, 1983-2000.
- Xue, L., Cai, W.J., Hu, X., Sabine, C., Jones, S., Sutton, A.J., Jiang, L.Q., Reimer, J.J., 2016. Sea surface carbon dioxide at the Georgia time series site (2006-2007): Air-sea flux and controlling processes. Prog. Oceanogr. 140, 14–26.
- Yang, B., Byrne, R.H., Lindemuth, M., 2015. Contributions of organic alkalinity to total alkalinity in coastal waters: A spectrophotometric approach. Mar. Chem. 176, 199–207.
- Yoder, J.A., Atkinson, L.P., Bishop, S.S., Blanton, J.O., Lee, T.N., Pietrafesa, L.J., 1985. Phytoplankton dynamics within Gulf Stream intrusions on the southeastern United States continental shelf during summer 1981. Cont. Shelf Res. 4, 611–635.

**Table 1.** Summary of  $pCO_{2w}$  internal consistency. GR Mooring observations are from 0.6 m, whereas underway are from 2.5 to 4 m, and bottle samples from 0 to 4 m. All units are µatm and in each instance the residual value is the first method minus the second method. The root mean squared error (RMSE) gives a statistical result for the error in the correlation model.

Pairing	<b>Correlation Statistics</b>	Mean, St dev, and median residual	Best fit slope and intercept	Mean Bias
Mooring-Underway	r = 0.92; p < 0.001; n = 51;	5.54 ± 32.4, 8.2	0.76, 102	$12.3 \pm 21.3$
Mooring-Bottle	RMSE = $1/.6$ r = 0.91; p < 0.001; n = 57; RMSE = $24.6$	24.4 ± 29.0, 25.0	0.79, 70.4	$24.6\pm30.7$
Underway-Bottle	r = 0.84; p < 0.001; n = 41; RMSE = 28.3	24.8 ± 29.2, 12.2	0.76, 80.7	$21.2 \pm 35.5$

**Table 2.** Linear least squares fit results for time series of variables with statistically significant results. In the least squares fit equation "t" is the unit of time (per day). The units of annual change are  $\mu$  atm y<sup>-1</sup> for CO<sub>2</sub> fractions, mmol m<sup>-2</sup> y<sup>-1</sup> for *F*CO<sub>2</sub>, and  $\mu$ mol kg<sup>-1</sup> for DIC. \*These results are for July 2006 through October 2014, the linear model is not significant July 2006 through July 2014.

	Regression	95% CI	Least squares fit	Annual change
pCO <sub>2a</sub>	$r^2 = 0.37; p < 0.001; n = 2,636$	5.8 to $6.4 \times 10^{-3}$	$pCO_{2a} = (6.1 \times 10^{-3}) \times t + 384.1$	$2.2\pm0.2$
$pCO_{2w}$	$r_{1}^{2} = 0.02; p < 0.001; n = 2,636$	$1.0 \text{ to } 1.7 \times 10^{-2}$	$pCO_{2w} = (1.3 \times 10^{-2}) \times t + 386.0$	$4.9\pm2.4$
<i>p</i> CO <sub>2nt</sub>	$r^2 = 0.08; p < 0.001; n = 2,612$	$1.8 \text{ to } 2.3 \times 10^{-2}$	$pCO_{2nt} = (1.9 \times 10^{-3}) \times t + 376.3$	$7.6\pm1.9$
$\Delta p CO_2$	$r^2 = 0.006; p < 0.001; n = 2,636$	5.3 to $12 \times 10^{-3}$	$\Delta p CO_2 = (8.5 \times 10^{-3}) \times t + 0.39$	$3.1 \pm 2.4$
MG01 FCO <sub>2</sub> *	$r^2 = 0.03; p < 0.001; n = 2,486$	$1.5 \text{ to } 2.4 \times 10^{-6}$	$FCO_2 = (1.9 \times 10^{-6}) \times t - (2.6 \times 10^{-3})$	$70 \pm 34$
W09 FCO <sub>2</sub> *	$r^2 = 0.03; p < 0.001; n = 2,486$	1.1 to $1.7 \times 10^{-6}$	$FCO_2 = (1.4 \times 10^{-6}) \times t - (1.9 \times 10^{-3})$	$51 \pm 25$
DIC	$r^2 = 0.05; p < 0.001; n = 455$	0.02 to 0.005	$DIC = 0.04 \times t + 611$	$15\pm16$

**Table 3.** Annual net  $FCO_2$  at the GR mooring and the sum of the daily standard deviations. The number of days for which there are missing observations after interpolation and gapfilling is in parenthesis following the year in the first column. W09 and MG01 are the two estimates for  $FCO_2$ , respectively. Since the time series begins in July of 2006, we only have data for the second half of the year, and because the deployment for the mooring ends in October of 2014, observations are not yet available for use for the end of 2014 through the present. Units are in mol m<sup>-2</sup>. \*The time series began in July of 2006. \*\*53 days are missing from the time series through the end of the deployment, which was October 24, 2014; however, the total missing for the 2014 calendar year is 122.

Year	W09 (mol m <sup>-2</sup> y <sup>-1</sup> )	MG01(mol m <sup>-2</sup> y <sup>-1</sup> )	MG01-W09
<b>2006</b> (198)*	$-0.39 \pm 0.26$	$\textbf{-0.35} \pm 0.44$	0.02
<b>2007</b> (49)	$-0.31 \pm 0.68$	$0.08 \pm 1.1$	0.39
2008 (9)	$\textbf{-0.07} \pm 0.85$	$0.53 \pm 1.4$	0.60
<b>2009</b> (73)	$-0.37 \pm 0.63$	$\textbf{-0.19} \pm 0.93$	0.18
<b>2010</b> (13)	$\textbf{-0.60} \pm 0.97$	$-0.04 \pm 1.4$	0.56
<b>2011</b> (100)	$-0.33 \pm 0.69$	$0.05 \pm 1.0$	0.38
<b>2012</b> (10)	$0.73 \pm 0.81$	$1.3 \pm 1.4$	0.57
<b>2013</b> (64)	$-0.28 \pm 0.55$	$0.01{\pm}0.96$	0.29
<b>2014</b> (53)**	$-0.66 \pm 0.68$	$-0.43 \pm 1.1$	0.23



**Figure 1.** Map of the SAB showing the location of the GR mooring, the major rivers in the region, and Sapelo Island at the mouth of the Altamaha River. The "X" marks the GR mooring at ~18 m. The blue arrows show the general location and directional flow of the low salinity counter current. The northward flowing mean position of the Gulf Stream core is the thick red line (see Castelao, 2014).



**Figure 2.** Time series of variables with slopes of the linear least squares fit. From the top: (A)  $pCO_{2a}$  with least squares (red line), (B)  $pCO_{2w}$  (black line and red slope line),  $pCO_{2t}$  (blue line), and  $pCO_{2nt}$  (pink line and light blue slope line) (C) SST, (D) SSS, and (E) Altamaha River stream flow. The trends (slopes) shown are significant at p < 0.001 and are significantly different than zero. The values of the trends are also presented in Table 1. Note the difference in scale of  $pCO_{2a}$  and  $pCO_{2w}$ .



**Figure 3.** This figure shows some of the different biogeochemical components of the  $pCO_{2nt}$  (plotted in panel B): in panel (A) mixing, represented by Altamaha River stream flow (smoothed over a 30 day period) and biological contributions from DIC (at the Doctortown, GA USGS site) in the watershed. In panel (B) monthly mean Chl *a* as a proxy for local biological consumption of CO<sub>2</sub> and  $pCO_{2nt}$  (daily mean time series and monthly means).



**Figure 4.** Time series of (A) the MG01  $FCO_2$  and (B) W09  $FCO_2$  estimates; (C) U<sub>10</sub>; and (D) the air-sea  $pCO_2$  difference ( $\Delta pCO_2$ ). The green lines in panels A and B and the red line in panel D are the linear least squares best fit lines that indicate a statistically significant increase in the given variable, based on observations that include values calculated with climatological mean SST and SSS. Negative values indicate a flux into the sea surface and positive values indicate a flux into the atmosphere.



**Figure 5.** Annual  $FCO_2$  using the two gas transfer parameterizations. The error bars are the sum of the daily standard deviations for W09 and MG01. Given considerable data loss in 2014, which was gap filled at a daily frequency, thus we do not have daily standard deviations for the whole year as we do for other years. Therefore, the error for year 2014 is the mean error from all other years. Larger data gaps occurred in 2007, 2009, 2011, 2013, and 2014 (end of deployment not finalized); see Table 3 for the number of missing days.



**Figure 6.** Internal consistency correlations from the three different  $pCO_{2w}$  observations methods using SSS and distance from the mooring to examine sources of discrepancies. Panels A, C, and E are the results of correlation analyses, and B and D are comparisons of spatial distance and SSS differences between the mooring and the other method. In this graphic  $\Delta pCO_2$  is the difference between  $pCO_{2w}$  from the x-axis minus  $pCO_{2w}$  from the y-axis. Panel F is the  $pCO_{2w}$  difference between the mooring and that calculated from DIC-TA of the bottle with respect to the SSS measured at the mooring. The black lines in all panels are the 1:1 lines and the red lines are the best-fit lines (slopes and intercepts are in Table 2).



**Figure 7.** Internal consistency correlation for results between mooring and underway  $pCO_{2w}$  using SST to examine sources of  $pCO_{2w}$  discrepancies. Panel A is the result of correlation analysis illustrating the same  $pCO_{2w}$  relationship in Figure 2A, however, with the color bar representing the difference in SST between the mooring and underway observations. Panel B is the comparison of spatial distance and SST differences between the mooring and underway observations. In this graphic  $\Delta pCO_2$  is the difference between  $pCO_{2w}$  from the x-axis minus  $pCO_{2w}$  from the y-axis. The black line is the 1:1 line and the red line is the best fit-line.



**Figure 8.** October 2011 internal consistency of (A) spatial underway  $pCO_{2w}$ , bottle-calculated  $pCO_{2w}$ , and GR mooring  $pCO_{2w}$  and (B) the observations plotted over time. The star in panel A is the mean from the mooring over the time period of the ground-truthing pass and the squares are the bottle samples. The gaps in underway observations in panel B either represent instrument calibration periods (typically on the order of 25 minutes), when the system was shut down to clean the filters (this region can be highly turbid, especially during wet periods), or are points that were omitted when the ship moved too far from the mooring.



**Figure 9.** July 2015 internal consistency of (A) spatial underway  $pCO_{2w}$ , bottle-calculated  $pCO_{2w}$ , and GR mooring  $pCO_{2w}$  and (B) the observations plotted over time. The star in panel A is the mean from the mooring over the time period of the ground-truthing pass and the squares are the bottle samples. The gaps in underway observations either represent instrument calibration periods or were omitted when the ship moved too far from the mooring. Note that the scale of the color bar in panel (A) is different than the scale in panel (A) of Figure 4. Note the difference in spatial scale between panel A of this figure and panel 4A.

## **Supporting Information**

# S1. Compilation of *p*CO<sub>2w</sub> observations from cruises

There have been a total of 24 different cruises ranging in duration of several hours to a full diel cycle on several ships with varying sea water intake depths for underway systems (Table 1 of the main text). Depending on the size of the ship, internal-consistency passes can get as close as tying up to the GR mooring and as far as up to 2 km away. Most cruises collect continuous observations as well as discrete water samples for DIC and TA. Over the last two years discrete water samples have also been collected during routine servicing of the GR mooring (zero meters distance), but are not accompanied by underway  $pCO_{2w}$  observations. Though cruises cover the whole SAB, we include primarily cruise observations (underway and discrete water samples) that were collected within less than 1 km from the GR mooring due to high offshore heterogeneity (Jiang et al., 2013, 2008b), however, in some years due to the limited number of passes, when samples less than 1 km from the mooring were not collected we do extend the radius to include cruises that came within 2 km of the site after assessing the spatial heterogeneity. Samples are matched to the mooring three-hour frequency as closely in time as possible, additionally, during some of the cruises the frequency of the mooring observation is increased to one measurement per hour, however, for the majority of cruises this did not occur. The 11 samples collected during the April 2011 cruise, with  $\Delta$ SSS up to 2.87, had subsequently large pCO<sub>2w</sub> differences (up to  $\sim$ 150 µatm; data not shown) and were omitted from internal consistency validation. While we cannot exactly say that spatial heterogeneity is the only cause of the  $pCO_{2w}$  discrepancies during April 2011, we are unable to identify any other issue with this cruise at this time.

Unfortunately, at this time we have too few high precision spectrophotometric pH samples to include this variable in a robust examination of  $pCO_{2w}$  internal consistency. DIC and TA are typically the most widely collected carbonate system samples (Patsavas et al., 2015), therefore in this study we focus on this pairing for calculating  $pCO_{2w}$  using CO2SYS (Lewis and Wallace, 1998). We briefly compared the CO2SYS options for  $K_1$  and  $K_2$  constants from Lueker et al. (2000) to Dickson and Millero (1987), as

well as Uppström (1974) or Lee (2010) for the Boron constants. We find that the closest calculated match (data not shown) to mooring  $pCO_{2w}$  is Dickson and Millero (1987) and Lee (2010).

## S2. Time series observations, processing, and gap-filling

Prior to calculating  $pCO_{2w}$  quality control/quality assurance was performed on the CO<sub>2</sub> data set as well as the other variables needed to calculate  $pCO_{2w}$ , including de-spiking, to detect observations potentially affected by biofouling or other sources of contamination or potential instrument errors. In this work we do not address the potential sources errors in the MAPCO<sub>2</sub> system, however, it should be noted that the system uncertainty in field tests is  $\pm 2 \mu atm$  (Schar et al., 2010; Sutton et al., 2014).

Gap-filling is often an essential and necessary component of time series analysis as sensors can have failures or have omitted observations that must be replaced. Due to the complexities of  $pCO_2$ calculations from multiple variables, several gap-filling methods were used at high (three hour) and low (daily mean) observation frequencies. Atmospheric pressure, SST, SSS, and  $xCO_2$  gaps did not always occur simultaneously because they are measured by independent sensors. Linear interpolation for gaps up to six hours (two measurements) was applied in the high (three hour) frequency time series; this method accounted for less than 3% of all observations in the high frequency time series. Once this most basic level of gap-filling was completed, we calculated daily mean values for all variables. Missing observations in the daily mean time series were linearly interpolated if they were two days or less, or filled using climatological values (ensemble means of the time series) for gaps three to seven days long (except two periods in 2013-2014 see below). At the daily time step, gap-filling was primarily accomplished through insertion of climatological means for SST, SSS, and atmospheric pressure to calculate  $pCO_{2w}$  for instances when  $xCO_{2w}$  was measured but the other variables were not. Approximately 25% of daily mean SST and SSS observations were gap-filled with either linear interpolation or climatological means through the end of the October 2013 deployment. We did not gapfill pCO<sub>2w</sub> with climatological means as this would introduce a high level of uncertainty and defeat the purpose of identifying  $pCO_{2w}$  variability itself. After gap-filling, for the time period between July 2006 and October 2013 we achieved 91% of daily  $pCO_2$  values. Gaps in summer of 2009 (May 19 to July 20), 2011 (May 19 to August 28), and 2014 (April 30 to May 13 and then June 16 to July 9) were not gapfilled since they were due to complete system malfunctions or only loss of  $xCO_{2w}$ . A longer gap in 2008 also occurred, due to suspected biofouling, and was also not gap-filled. It is important to note that we did not use any gap-filled values for the internal consistency analyses.

Due to longer term data loss during the 2013-2014 deployment of measured SST and/or SSS caused by system malfunctions, we gap-filled these sections (approximately November 1, 2013 to March 1, 2014 and August 13 to October 24, 2014)  $pCO_{2w}$  using an adjusted  $xCO_{2w}$  value since fewer observation of  $xCO_{2w}$  were lost. This adjustment uses the mean residual value between  $pCO_{2w}$  and  $xCO_{2w}$  from correlation analysis, using all of the high frequency  $xCO_{2w}$  and calculated  $pCO_{2w}$  values before gap-filling. The mean residual value from the correlation is  $-1.1 (\pm 0.9)$  and  $-0.6 (\pm 0.7)$  to adjust from units of ppm to µatm for the water and air, respectively. This correction is applied to 244 days of three hour frequency observations from November 2013 through October 2014. These corrections are within the MAPCO<sub>2</sub> system uncertainty for  $pCO_{2w}$  measurements, and therefore likely do not substantially affect  $pCO_{2w}$  uncertainty.

One of the greatest sources of  $pCO_{2w}$  and  $pCO_{2a}$  variability is temperature due to  $CO_2$  dependence on solubility in the water as well as humid air. We assess the uncertainty related to gap-filling with climatological mean SST in order to show that it does not greatly influence our final results, which are based on calculated daily mean  $pCO_{2w}$ . We find that the mean daily SST variation in this time series is  $\pm \sim 1 (\pm 0.45)$  °C. Then, based on the Takahashi et al. (1993) relationship of temperature change to  $pCO_{2w}$ we calculate the influence of a  $\pm 1$  °C temperature uncertainty using the time series mean  $pCO_{2w}$ (409 µatm). The mean SST uncertainty results in an adjustment of  $\pm 17$  µatm  $pCO_2$ . Using the minimum and maximum  $pCO_2$  (253 and 567 µatm), we find that the uncertainty can range from  $\pm 10$  to 23 µatm. Daily  $pCO_2$  standard deviations range from  $\sim 0$  to 47 µatm, therefore, the  $pCO_2$  uncertainty associated with using climatological means is well within the range of daily SST variability, though likely higher than the moored  $pCO_2$  system uncertainty.

## S3. Gap-filling FCO<sub>2</sub> time series and related uncertainty

Due to the loss of SST and SSS 2013-2014 (244 days) we also assess uncertainty using climatological means for calculating  $FCO_2$ . According to the bulk equation, the primary factors that regulate  $FCO_2$  are the air-sea differential ( $\Delta pCO_2$ ) and gas transfer velocity, as well as solubility though to a lesser extent than the first two variables (Wanninkhof, 2014). Even if the thermodynamic effects of SST and SSS on  $FCO_2$  could be minimal via the bulk equation, ultimately  $FCO_2$  is also affected by  $pCO_{2w}$ and  $pCO_{2a}$  via the water vapor pressure from the conversion of dry mole fraction  $CO_2$  to the wet partial pressure (Sutton et al., 2014) as well. Similar to the test of SST uncertainty for pCO<sub>2w</sub> calculations (~4% °C<sup>-1</sup>), we use upper and lower  $pCO_{2w}$  examples from the 2013-2014 missing values to estimate uncertainty propagated through calculated  $FCO_2$ . Step-wise, we find that the  $pCO_{2w}$  and  $pCO_{2a}$  could be under/overestimated by ~1  $\mu$ atm (xCO<sub>2</sub> - pCO<sub>2</sub> adjustment described above), then including the 1  $\mu$ atm propagated through the FCO<sub>2</sub> calculation and the effects of solubility in the bulk equation (up to  $a \pm 1^{\circ}C$ mean daily variation), the final result could only be up to  $\pm$  3.3% relative difference between FCO<sub>2</sub> calculated with and without a climatological value. Therefore, we assume this contribution of uncertainty to daily mean values is relatively small compared to other uncertainties associated to the gas transfer velocity (Wanninkhof, 2014). The effects of gap-filling on FCO<sub>2</sub> annual sums can only give a general sum of annual uncertainty because true daily variations are smoothed using climatological means, thus reducing the up to 2.5°C day-night SST variation to the ~1°C mean daily variation. Therefore, future work should further seek to constrain uncertainties related to gap-filling FCO<sub>2</sub>. The biogeochemical mechanistic effects of SST and SSS on  $FCO_2$  are impossible to discern within the gap-filled time period without observations, and as of yet there is no established method for gap-filling three-hour frequency SST and SSS at this site.

# S4. Discussion of internal consistency validation method discrepancies

Discrepancies between validation methods can come from either environmental heterogeneity and/or methodological errors. For example, during three large scale cruises, Gulf of Mexico and East Coast Carbon I and II (July 2007 and July 2012, respectively), and ECOA 2015, all DIC and TA samples were analyzed at sea without preservation, while the rest of the cruises preserved (addition of saturated Mercuric Chloride [HgCl<sub>2</sub>]) samples and analyzed them upon return, days to weeks after collection. In some cases samples were collected using small boats (hand sampled from surface water) and were not preserved but analyzed within several days of collection, and other times HgCl<sub>2</sub> was added up to 8 hours after sample collection due to various logistical limitations. Therefore DIC in some samples could be altered due to biological processes that occurred due to improper and/or inconsistent preservation of the sample. Samples with very large, often inexplicable, excursions were not included in the data analyses.

Analytical uncertainty, storage errors, and human errors are also introduced into the DIC and TA analysis process (Huang et al., 2012); therefore there is a possibility of introduced error as well as environmental differences in discrete bottle samples that arguably is not an issue with autonomous and underway systems that are calibrated prior to use and consistently throughout the deployment. While all methods have internal checks for their own functions, human error in laboratory samples without replicate samples can be difficult to detect. If we assume that the analysis uncertainty in discrete water samples can be approximated by the  $\pm 2 \mu \text{mol kg}^{-1}$  analytical uncertainty, then adjusting the DIC value plus  $2 \mu \text{mol kg}^{-1}$ , and TA minus  $2 \mu \text{mol kg}^{-1}$ , could result in a mean  $p\text{CO}_{2w}$  uncertainty of 7.40  $\pm$  1.25  $\mu \text{atm}$  (n = 57) for each bottle sample or a mean relative difference of  $1.77 \pm 2.20\%$  of the original calculated  $p\text{CO}_{2w}$  from DIC-TA. Whereas, the results for a combined  $\pm 4 \mu \text{mol kg}^{-1}$  TA-DIC change would cause a greater discrepancy of up to  $42 \pm 33 \mu \text{atm } p\text{CO}_{2w}$ , or a mean relative uncertainty of 10%, well beyond acceptable internationally recognized OA monitoring network standards (Newton et al., 2014). Furthermore, both the MAPCO<sub>2</sub> and underway systems have an uncertainty of  $\leq 2 \mu \text{atm}$ , which can then increase or decrease the bottle-mooring or bottle-underway discrepancy. The maximum calculated analytical uncertainty for the combined  $\pm 4 \mu \text{mol kg}^{-1}$  in DIC/TA could propagate through calculations

and account for a substantial portion of the median residual  $pCO_{2w}$  between methods (Table 2), especially if TA and DIC uncertainties are in opposite directions. Even the seemingly very large and consistent offset between the mooring and bottle  $pCO_{2w}$  (and underway-bottle) in October 2011 (Figure 4B) is within this relatively accepted uncertainty of  $\pm 4 \mu mol kg^{-1}$  in the DIC-TA combination. Other studies have also noted that  $pCO_{2w}$  calculated from DIC and TA had larger discrepancies at higher concentrations, which could also be the case in the present internal consistency effort (Figures 2A, 2C, 2E, and 3A), though to date it is unclear why (Lueker et al., 2000; Ribas-Ribas et al., 2014). Even though the correlation value for the mooring-bottle pair is high, the residual values are out of the range of analytical uncertainty (Table 2), suggesting that another source of for  $pCO_{2w}$  discrepancy exists.

Environmental vertical and horizontal spatial heterogeneity (based on assessment of SST and SSS) is another factor for inconsistencies in  $pCO_{2w}$ . For example, if the water column is stratified, then the difference in sampling depth between the mooring (0.6 m) and either bottle (typically < 2.5 m) or underway (2.5 to 4 m) measurement may account for high residual  $pCO_{2w}$  and higher bias values. Since SSS greatly influences the TA at this site (Cai et al., 2010) different values between the methods can indicate differences in TA, thus calculated pCO<sub>2w</sub>. For example, previous work in the SAB determined that for every theoretical horizontal 1 unit SSS change there can be an average  $\sim 50 \ \mu mol \ kg^{-1}$  change in TA (Xue et al., 2016), which would account for residual values for the mooring-bottle pair out of the range of analytical uncertainty (Patsavas et al., 2015) when SSS differences are larger. During the two cruise examples detailed in this work the spatio-temporal SSS differences were  $\leq 0.2$  units (Figure S1) but in general were up to  $\pm \sim 2.3$  units (Figure 2B and 2D), however, more detailed assessment of spatial heterogeneity should be addressed in the future. Assuming water masses are the same within the cruise track and the water layer samples around the mooring then  $pCO_{2w}$  discrepancies between the methods could be due to differences in SST and SSS measurements of the three methods, thus conversions of dry mole  $CO_2$  to  $pCO_{2w}$  could cause the discrepancies. Therefore, we correct  $pCO_{2w}$  to the *in situ* SST of the mooring (Schar et al., 2010) using:

$$pCO_{2TC} = pCO_{2w}^{(0.0412 \times (mSST - xSST))}$$
(1)

where  $pCO_{2TC}$  is the SST-corrected  $pCO_{2w}$ , mSST is mooring SST and xSST is either the bottle or underway SST (after Takahashi et al. [1993], and using a coefficient derived for the range of SST and SSS of the SAB inner shelf). The resulting correlation coefficients decrease as do the slopes of the best fit lines (Figure S2), therefore, the measured SST differences, whether due to environmental variability or sensor differences, are not likely the cause of  $pCO_{2w}$  discrepancies.

TA overestimation, due to organic bases, could cause misrepresentations of calculated pCO<sub>2w</sub> (Yang et al., 2015). Since the SSS for the 57 bottle samples only explains 87% of TA variability  $(r^2 = 0.87; p < 0.001)$ , TA may be influenced by a source other than SSS, thus potentially underestimating  $pCO_{2w}$  in bottle samples. Previous studies have identified estimates for organic alkalinity (org-Alk) from organic bases on coastal margins as contributors to TA, which can be as high as 25  $\mu mol~kg^{\text{-1}}$  near the GR mooring (Patsavas et al., 2015) to 40 µmol kg<sup>-1</sup> in the eastern Gulf of Mexico (Yang et al., 2015). Furthermore, org-Alk may be as high as ~50 µmol kg<sup>-1</sup> in the Altamaha River (Cai et al., 1998). Overestimation of TA relative to the same DIC could artificially decrease calculated  $pCO_{2w}$ , and therefore could explain the overall underestimated calculated pCO<sub>2w</sub> values for the mooring-bottle relationship (Figure 2) (Patsavas et al., 2015). Upon closer examination of the  $pCO_{2w}$  discrepancies in the present studies, adjusting the TA concentration of the bottles for the potential effects of org-Alk, we find that subtracting 10 to 30 µmol kg<sup>-1</sup> for samples with greater discrepancies can improve the correlation to mooring  $pCO_{2w}$  from r = 0.91 to r = 0.96 (Figure 3S) as well as reducing the RMSE from 24.6 to 17.6 µatm. Adjusting for org-Alk also reduces the residual values to a mean of  $< 1 \pm 17.5$  µatm pCO<sub>2w</sub> and median of 0.78 µatm pCO<sub>2w</sub>. Additionally, up to 22% of the mooring-bottle pCO<sub>2w</sub> discrepancy could be explained by lower SSS (Figure 2F). Coastal lower SSS waters could have greater org-Alk than higher SSS oceanic waters, therefore bottle  $pCO_{2w}$  could be underestimated due to overestimated TA. The correlation between underway and bottles improves from r = 0.84 to r = 0.91 and reduced the RMSE

from 28.3 to 26.5 µatm as well as improves the mean residual to  $5.8 \pm 26$  µatm and the median to 3.7 µatm. Specifically for the October 2011 cruise, adjusting for Org-Alk the mooring-bottle mean residual reduces to  $5.9 \pm 5.7$  µatm (median 5.7 µatm) and the underway-bottle mean also reduces to  $5.8 \pm 26$  µatm (median 3.7 µatm). It should be noted that this adjustment is a rudimentary analysis based on estimates from previous studies and not from laboratory testing of the samples used in this study, however, it may also illustrate the importance of determining the org-Alk in future water samples obtained for mooring ground-truthing in coastal waters.

**Table 1.** Summary of cruises conducted in the SAB with stations at or near the GR mooring. The number of bottles column refers to surface samples from the cruises that were collected from stations no more than 2 km from the GR mooring. All bottle samples in this table are time matched to the closest mooring measurement. Underway and bottle samples were collected during each cruise unless otherwise stated. Mooring maintenance trips were conducted on various small boats provided by the Gray's Reef National Marine Sanctuary (GRNMS) where a hose run over the side of the boat was placed ~0.5 m below the water line with bottles sampled off a splitter.

	Sea water	Number of	
Date at GR mooring	intake depth	bottle	
(UTC)	( <b>m</b> )	observations	Comments
July 20, 2007	5	1	<i>GOMECC I</i> ; DIC and TA collected from different Niskin bottles
August 29, 2008	0.5	2	R/V Joe Ferguson
December 19, 2008	0.5	5	GRNMS
May 28, 2009	0.5	1	R/V Joe Ferguson
November 16, 2009	0.5	1	R/V Joe Ferguson
February 9, 2010	0.5	1	GRNMS
April 21, 2010	0.5	1	R/V Joe Ferguson
August 2, 2010	0.5	1	R/V Joe Ferguson
October 12, 2010	0.5	4	R/V Joe Ferguson
March 22, 2011	0.5	1	R/V Joe Ferguson
April 19 and 21 to 22, 2011	2.5	11 <sup>A</sup>	R/V Savannah; 24 hours: 21st to 22nd
October 3 and 6 to 7, 2011	2.5	9	R/V Savannah; 24 hours: 6 <sup>th</sup> to 7 <sup>th</sup>
December 15, 2011	0.5	1	R/V Joe Ferguson
February 2, 2012	0.5	3	R/V Joe Ferguson
May 1, 2012	0.5	3	R/V Joe Ferguson
August 2, 2012	5	3	<b>GOMECC II</b> ; One sample each from of two stations around the meering
September 5, 2013	0.5	4	R/V Joe Ferguson
May 18, 2014	2.5	1	R/V Savannah
July 19, 2014	2.5	1	R/V Savannah
September 29, 2014	2.5	1	R/V Savannah
November 4, 2014	2.5	1	R/V Savannah
December 7, 2014	2.5	1	R/V Savannah
April 22, 2015 <sup>B</sup>	N/A	3 <sup>C</sup>	GRNMS
Jun 23, 2015 <sup>B</sup>	N/A	3 <sup>C</sup>	GRNMS
July 19, 2015 <sup>D</sup>	4	3 <sup>C</sup>	ECOA-2015

<sup>A</sup>Discrete water sample data from this cruise were omitted from correlation analyses results, see explanation in the text. <sup>B</sup>Only bottle samples were collected at 0 km from the mooring. <sup>C</sup>Nine bottles total were collected as triplicates and averaged to one value for comparison to the mooring and underway observations with accuracy within  $\pm 2.0 \ \mu$ mol kg<sup>-1</sup>. <sup>D</sup>Note that while the full MAPCO<sub>2</sub> 2014 to 2016 deployment data have not been finalized, the 24 hour period used for internal consistency during ECOA-2015 has been made available for this work.



**Figure 1S.** Spatial distribution of SSS during (A) the October 2011 cruise and (B) ECOA-2015, July 2015. Note the spatial difference in the area covered by the cruise. The stars are the mean SSS and location of the mooring, the squares are the location of the bottle samples (assuming the same body of water as the underway sample), and the circles are the underway samples. Note that there is less than an approximately  $\pm 0.2$  unit difference in SSS between the mooring and location of either underway or bottle samples during both cruises.



**Figure 2S.** Correlations for (A) underway  $pCO_{2w}$  adjusted to mooring *in situ* SST and (B) calculated  $pCO_{2w}$  from bottle samples adjusted to mooring  $pCO_{2w}$ . The stars (and blue best fit line) in each panel are the adjusted values and the circles (and red best fit lines) represent the original values. The blue and red text in each panel are the correlation statistics for their respective best fit lines. The best fit slope and intercept, respectively, for adjusted (blue line) underway values are 0.69 and 131.6 and for bottle values are 0.67 and 110.5; original values (red line) are 0.76 and 102 for mooring-underway and 0.79 and 70.4 for mooring-bottle (Table 2 of the main text).



**Figure 3S.** Correlation of mooring  $pCO_{2w}$  and bottle-calculated  $pCO_{2w}$  after reducing the TA values for the potential influences of organic alkalinity due to organic bases. The black line is the 1:1 line and the red line is the best fit line with a slope of 0.92 and intercept of 31.4.