A Multiple Linear Regression of pCO_2 against Sea-Surface Temperature, Salinity, and Chlorophyll a at Station ALOHA and its Potential for Estimate pCO_2 from Satellite Data

Ashley D. Berryman
Undergraduate Research Experience
Elizabeth City State University
Elizabeth City, NC
adberryman@mail.ecsu.edu

Diaminatou Goudiaby
Undergraduate Research Experience
Elizabeth City State University
Elizabeth City, NC
dagoudiaby@mail.ecsu.edu

Phillip Moore
Undergraduate Research Experience
Saint Augustine's College
Raleigh, NC
phillipmoore115@yahoo.com

Abstract—**Ocean is one of the major reservoirs of carbon and can be a major sink of anthropogenic carbon dioxide. Together with pH, alkalinity, and total dissolved inorganic carbon (DIC), partial pressure of carbon dioxide (pCO2) is one of the four essential parameters for determining aquatic CO_2 system. These four CO_2 parameters are interrelated through chemical equilibrium and the determination of any two is sufficient for calculating the other two parameters. Ship-based oceanographic research cruise, that is expensive to operate and inefficient to provide global coverage, has long been the main source of data for characterizing oceanic CO_2 system. Recently, Lohrenz and Cai (2006) conducted a field study of partial pressure of carbon dioxide, temperature, salinity, and chlorophyll *a* in surface waters of the Northern Gulf of Mexico and developed a correlation method for estimating carbon dioxide distribution from the Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing data. Although it showed great potential, the correlation is based on field data with a small temperature variation and atypical salinity, and it is not clear whether it can be applied elsewhere. Here, we propose to extend the applicability of the method by conducting a data analysis study of field observations conducted at station ALOHA (A Long-term Oligotrophic Habitat Assessment; 22° 45'N, 158° 00'W)**

Specifically we: (1) Obtained field data of alkalinity, DIC, temperature, salinity, and Chlorophill *a* determined at station ALOHA in the last two decades; (2) Calculated pCO_2 from alkalinity and DIC; (3) Applied the correlation method to test the applicability of the method in the central North Pacific Ocean; (4) Applied the correlation method and predict the distributions of partial pressure and air-sea fluxes of carbon dioxide in the central Pacific Ocean from MODIS data.

I. Introduction

Ocean is a major sink of human produced carbon dioxide. Scientists are becoming more and more interested in partial pressure of carbon dioxide (pCO_2) because it is a key parameter in measuring carbon dioxide (CO_2), one of the most powerful greenhouse gases and one that has a determining effect on climate. It is very difficult to measure sea-surface CO_2 in the field; it is even harder to measure it from satellite. Research cruise has been the method of choice for many decades. But those cruises are very expensive to conduct and time consuming, resulting in scientists working on alternative methods. Satellite remote sensing is a promising alternative method for measuring pCO_2. Although satellite measurements have been proposed [Lohrenz and Cai 2006] on the Mississippi shelf, its application in the other ocean regions is questionable. We set out to expand that method on other parts of the ocean. Our research will be based on monthly measurements of temperature, salinity, chlorophyll a, PC, DOC, and pCO_2 measurements at Hawaii.

II. Methods

A- Early Methods

Previous researches used different methods for the different parameters. Total Alkalinity (TAlk), dissolved inorganic carbon (DIC), temperature, salinity and chl *a* were obtained by field measurements. Surface water samples were obtained with a CTD rosette system equipped with 10-L Niskin bottles. Wang and Cai [2004] provided the formulas to calculate DIC, Talk, and pH, and also to calculate pCO_2 from DIC and Talk. They filtered surface water samples with GF/F filter and fluorometric assay of 90% acetone extracts to get chlorophyll measurements.

Remote sensing methods for sea-surface temperature have been in routine use for two decades [*McClain et al., 1985*]. Sea-surface chlorophyll *a* is determined using OC4 algorithms [*O'Reilly et al., 1998]*, dissolved organic carbon (DOC) and particular organic carbon (POC) together with sea-surface chlorophyll *a* can also be estimated with a algorithm proposed recently [*Maritorena et al., 2002]*.

Lohrenz and Cai [2006] proposed a new method to calculate pCO_2 from satellite. Their method was a multiple linear correlation using parameters that can be measured with satellite. The correlation was developed using ship-based measurements.

B- Our Methods

We used field data of Temperature, Salinity, TAlk, Dissolved Organic Carbon (DOC), Dissolved Inorganic Carbon (DIC), Chlorophyll *a*, particular carbon (PC), and pH collected from ALOHA Station (Figure 1) from the past two decades. Raw data were sorted and all unuseful measurements (i.e. incomplete) were removed. The remaining data were put into the data analysis software in Microsoft Excel. That information was then used to formulate the regression. The results of the regression analysis allowed us to compare original pCO_2 data and pCO_2 calculated with the model 1 multiple linear regression equation.

Map of the Hawaiian Islands

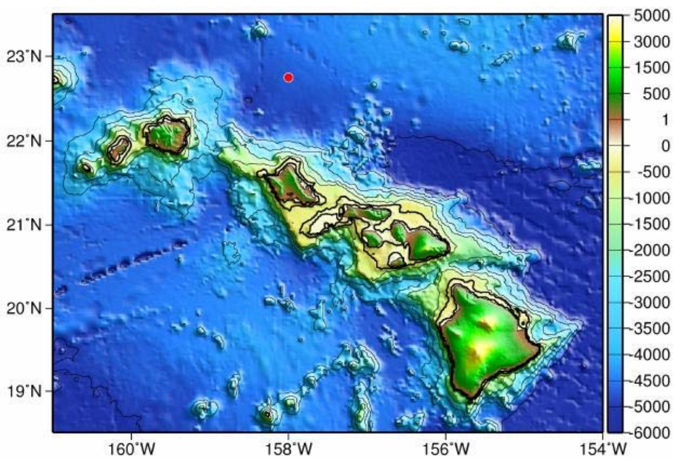

Figure 1. Red circle indicates the location of Station ALOHA (From the ALOHA Observatory website)

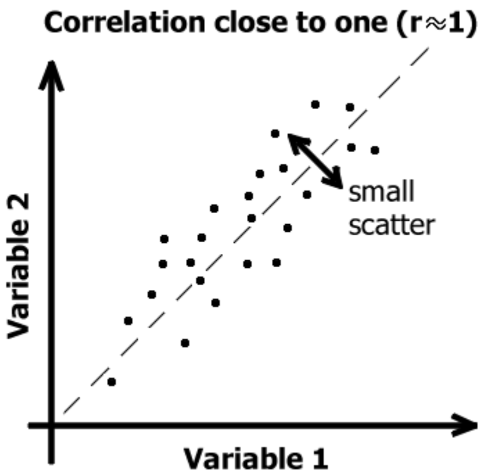

Figure 2. The closer the correlation coefficient is to 1, the more the points will fall along the line

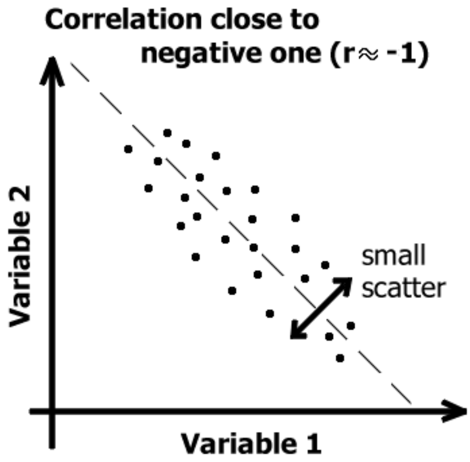

Figure 3. The closer the correlation coefficient is to -1, the more the points will fall along the line

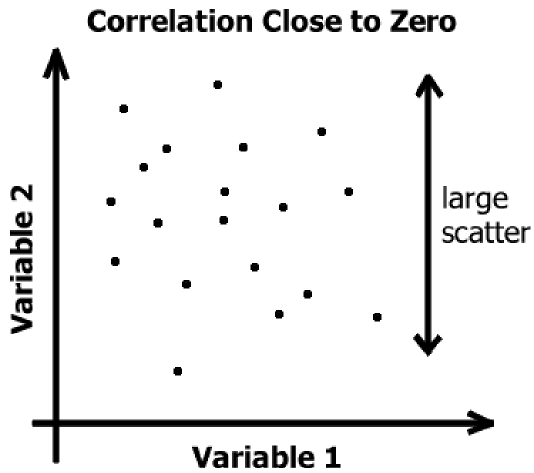

Figure 4.The closer the correlation coefficient is to 0, the less the points fall on the straight line

P-value allows us to know the significance of the parameters we obtained in the correlation. The closer the P-value is to 0, the more significant the parameter is to the correlation.

III. Results

A- Simple Linear Regression

Simple Linear Regression between pCO_2 and temperature, salinity, and chlorophyll *a* yield moderate to poor correlations. pCO_2 versus Temperature has a moderate correlation with a R Square of 0.5926 (Figure 5), same as pCO_2 versus salinity (Figure 6). However, there was a poor correlation between Pco2 and Chlorophyll a, with a R Square of less than 0.5 (Figure 7). Although the first two Simple Linear correlations can be used to predict pCO_2 since R Squares are relatively small, the accuracy of predicted pCO_2 will be poor. And therefore we set out to try the multiple linear correlation to improve the correlation.

B- Multiple Linear Regression

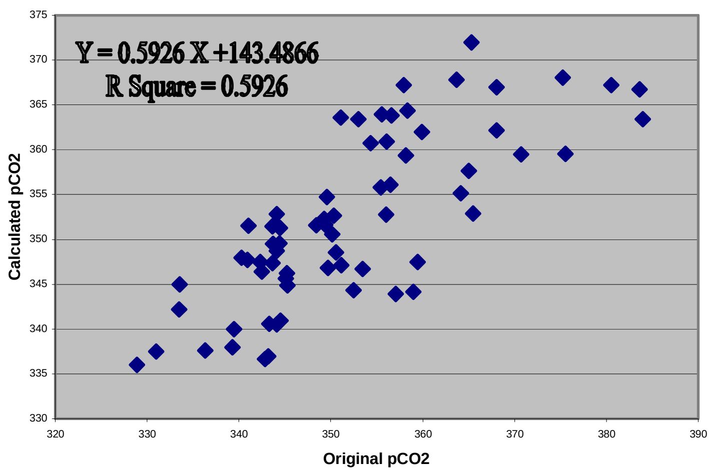

Figure 5. Comparison between Original and Calculated pCO_2. Temperature, salinity, and chlorophyll a were used to calculate pCO_2.

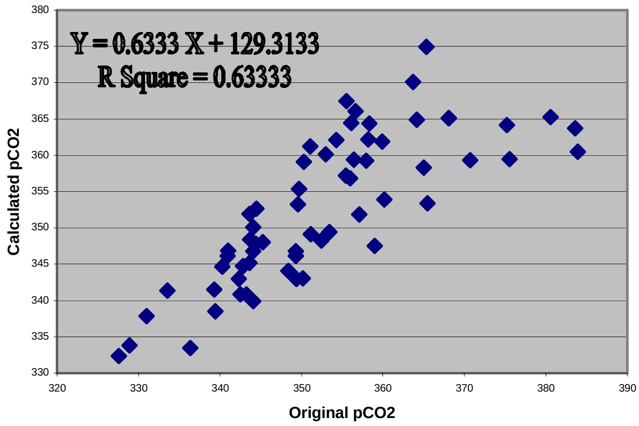

Figure 6. Comparison between Original and Calculated pCO_2. Temperature, salinity and DOC were used to calculate pCO_2.

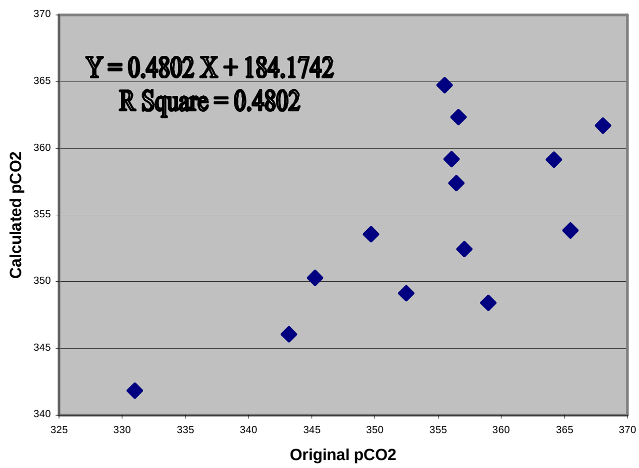

Figure 7. Comparison between Original pCO_2 and Chlorophyll *a*. Temperature, salinity, and PC were used to calculate pCO_2 .

Three multiple linear correlations are tested. In all those linear correlations, pCO_2 is the independent variable. In the first linear correlation, independent variables are temperature, salinity and chlorophyll a. The multiple linear correlations yield a R Square value of approximately 0.5925 (table 1). Although the R Square is relatively small, the P-values for all three variables are less than 2%. These P-values indicate that all three parameters are significant in multiple linear correlations.

TABLE I. Regression results using temperature, salinity, and Chlorophyll a

R-Square 0.5926	**Coefficients**	**P-Value**
Intercept	-62.0098	0.7578
X-Variable 1	-7.2258	8.8441E-14
X-Variable 2	17.2522	0.004
X-Variable 3	-91.9544	0.0112

In the second multiple linear correlation, independent variables are temperature, salinity, and DOC. This correlation yields a slightly higher R Square. The P-values for the slopes of the linear correlation are less than 5%. Again, these P-values indicate that all three parameters are significant.

TABLE II. Regression results using temperature, salinity, and DOC

R-Square 0.6333	**Coefficients**	**P-Value**
Intercept	123.9069	0.5487
X-Variable 1	-7.1275	1.32E-13
X-Variable 2	12.2405	0.0406
X-Variable 3	-0.2267	0.0017

In the third Multiple Linear correlation, the variables are temperature, salinity, and PC. This correlation yields the smallest R Square of 0.4802. Additionally, P-value for the slopes of two parameters are more than 0.35, these P-values indicate that pc and salinity are insignificant variables in this correlation.

TABLE III. Regression results using temperature, salinity, and PC

R-Square 0.4802	**Coefficients**	**P-Value**
Intercept	69.2194	0.8777
X-Variable 1	-5.2363	0.0301
X-Variable 2	11.7758	0.3534
X-Variable 3	2.6125	0.6983

IV- DISCUSSION

If data are available, the second correlation (i.e. using temperature, salinity, DOC as independent variables) would yield more accurate estimates of pCO_2. It's important to note that the third correlation may have improved considerably with more measurements of PC available. Although the first correlation did not produce the highest R Square, sea-surface chlorophyll *a* is more routinely determined from satellite, so it may be more practical to estimate pCO_2 from the first Multiple Linear correlation. Salinity is involved in all our multiple linear correlations. Satellite sensors for salinity is planned to be launched in 2008 by NASA.

V- CONCLUSION

A multiple linear correlation between pCO_2, temperature, salinity, and chlorophyll *a* was obtained from sea-surface water in Central North Pacific Ocean. All three independent variables are significant in the correlation. The correlation can be used to estimate pCO_2 from parameters that can be determined from satellite remote sensing. We used P-values to confirm that temperature, salinity and chlorophyll a were significant factors in calculating pCO_2.

ACKNOWLEDGMENTS

The 2007 Oceanography Team would like to thank Dr. Jinchun Yuan for his assistance and dedication to the team and project. Also to be acknowledged are Dr. Linda Hayden, Principal Investigator of the 2007 URE Program at Elizabeth City State University, CERSER, and CReSIS for their sponsorship, Mr. Jeff Wood, multimedia developer at ECSU for his help in building our webpage.

REFERENCES

[1] Maritorena, S., D.A. Siegel and A.R. Peterson, (2002). Optimization of a semianalytical ocean color model for global-scale applications. Applied Optics-LP, 41, 2705-2714.

[2] McClain E.P.,Pichel, W. and Walton C.C. (1985). Comparative performance of AVHRR-based multichannel sea surface temperature. Journal of Geophysical Resear, 90(C6):11587-11601.

[3] O'Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Garder, S.A. Garver, M. Kahru, and C. McClain, (1998). Ocean Color chlorophyll algorithms for SeaWiFS, Journal of Geophysical Research, 103 (C11), 24937-24953.

[4] Lohrenz, S.E.,W.J. Cai, (2006). Satellite ocean color assessment of air-sea fluxes of CO2 in a river-dominated coastal margin, Geophysical Research Letters, VOL. 33, L01601, doi:10.1029/2005GL023942, 2006.

[5] Wang, Z. H. A., and W. J. CAI (2004), Carbon dioxide degassing and inorganic carbon export from a marsh-dominated estuary (the Duplin River): A marsh CO2 pump, *Limmol. Oceanogr.*, 49, 341-354.