Abstract

In this dissertation, the aim is to investigate the empirical relationship between the partial pressure of CO_2 (pCO_2) and other ocean variables in the Southern Ocean, by using a small percentage of the available data.

CO_2 is one of the main greenhouse gases that contributes to global warming and climate change. The concentration of anthropogenic CO_2 in the atmosphere, however, would have been much higher if some of it was not absorbed by oceanic and terrestrial sinks. The oceans absorb and release CO_2 from and to the atmosphere. Large regions in the Southern Ocean are expected to be a CO_2 sink. However, the measurements of CO_2 concentrations in the ocean are sparse in the Southern Ocean, and accurate values for the sinks and sources cannot be determined. In addition, it is difficult to develop accurate oceanic and ocean-atmosphere models of the Southern Ocean with the sparse observations of CO_2 concentrations in this part of the ocean.

In this dissertation classical techniques are investigated to determine the empirical relationship between pCO_2 and other oceanic variables using *in situ* measurements. Additionally, sampling techniques are investigated in order to make a judicious selection of a small percentage of the total available data points in order to develop an accurate empirical relationship.

Data from the SANAE49 cruise stretching between Antarctica and Cape Town are used in this dissertation. The complete data set contains 6103 data points. The maximum pCO_2 value in this stretch is 436.0 μatm, the minimum is 251.2 μatm and the mean is 360.2 μatm. An empirical relationship is investigated between pCO_2 and the variables Temperature (T), chlorophyll-a concentration (Chl), Mixed Layer Depth (MLD) and latitude (Lat). The methods are repeated with latitude included and excluded as variable respectively. D-optimal sampling is used to select a small percentage of the available data for determining the empirical relationship. Least squares optimization is used as one method to determine the empirical relationship. For 200 D-optimally sampled points, the pCO_2 prediction with the fourth order equation yields a Root Mean Square (RMS) error of 15.39 μatm (on the estimation of pCO_2) with latitude excluded as variable and a RMS error of 8.797 μatm with latitude included as variable. Radial basis function (RBF) interpolation is another method that is used to determine the empirical relationship between the variables. The RBF interpolation with 200 D-optimally sampled points yields a RMS error of 9.617 μatm with latitude excluded as variable and a RMS error of 6.716 μatm with latitude included as variable. Optimal scaling is applied to the variables in the RBF interpolation, yielding a RMS error of 9.012 μatm with latitude excluded as variable and a RMS error of 4.065 μatm with latitude included as variable for 200 D-optimally

Investigating the empirical relationship between oceanic properties observable by satellite and the oceanic pCO_2

M.van der Walt
20238703

Dissertation submitted in fulfilment of the requirements for the degree *Master of Science in Applied Mathematics* at the Potchefstroom Campus of the North-West University

Supervisor: Prof G. Drevin
Co-supervisor: Dr S. Kok
Assistant supervisor: Prof J. Spoelstra

November 2011

sampled points.

Keywords

Southern Ocean; Global warming; Genetic algorithm; Radial basis function interpolation; Curve fitting

Opsomming

Die doel van hierdie verhandeling is om die empiriese verwantskap tussen parsiële druk van CO_2 (pCO_2) en ander oseaaneienskappe te ondersoek in die Suidelike Oseaan deur slegs 'n klein persentasie van die beskikbare data te gebruik.

CO_2 is een van die primêre kweekhuisgasse wat bydra tot aardverwarming en klimaatsveranderinge. Die konsentrasie van antropogeniese CO_2 in die atmosfeer sou egter baie hoër gewees het as 'n gedeelte daarvan nie deur die oseane en landelike gebiede opgeneem was nie. Die oseane neem CO_2 op vanaf die atmosfeer en stel ook CO_2 vry in die atmosfeer. Daar word beweer dat groot areas in die Suidelike Oseaan 'n CO_2 sink is, menende dat dit meer CO_2 opneem vanuit die atmosfeer as wat dit vrylaat. Die lesings van CO_2 in die Suidelike Oseaan is egter yl, en dus kan die mate waartoe die oseaan 'n CO_2 sink of bron is, nie akkuraat bepaal word nie. Benewens is dit moeilik om akkurate oseaan en oseaan-atmosfeer modelle te ontwikkel vir die Suidelike Oseaan met die yl gesaaide metings van CO_2.

In hierdie verhandeling word klassieke tegnieke ondersoek, om die empiriese verwantskap tussen pCO_2 en ander oseaaneienskappe te bepaal deur *in situ* metings te gebruik. Verder word die seleksie van 'n klein persentasie punte (van die volledige data stel) ondersoek. Die doel is om oordeelkundig 'n seleksie van punte te maak, sodanig dat die fout op die benadering van pCO_2 vanuit die empiriese verwantskap so klein as moontlik is.

Die data wat gebruik word in hierdie verhandeling, is verkry deur lesings wat geneem is gedurende die SANAE49 skeepsvaart tussen Antarktika en Kaapstad. Die volledige datastel bevat 6103 punte. Die maksimum pCO_2 waarde gedurende die vaart is 436.0 μatm, die minimum is 251.2 μatm en die gemiddeld is 360.2 μatm. 'n Empiriese verwantskap word ondersoek tussen pCO_2 en Temperatuur (T), chlorofil-a konsentrasie (Chl), Gemengde Laag Diepte (MLD) en breedtegraad (Lat). Die metodes word herhaal met en sonder breedtegraad onderskeidelik. D-optimale seleksie word gebruik om 'n klein persentasie punte van die volledige datastel te selekteer waarmee die empiriese verwantskap bepaal word. Kleinste-kwadrate-optimalisering is een metode wat gebruik word om die empiriese verwantskap te ondersoek. Vir 200 punte wat met D-optimale seleksie gekies is, lewer die 4de orde vergelyking 'n wortel-gemiddelde-kwadraat-fout (RMS) van 15.39 μatm (op die benadering van pCO_2) sonder breedtegraad ingesluit as veranderlike en 'n RMS fout van 8.797 μatm met breedtegraad ingesluit as veranderlike. Radiale basisfunksie-interpolasie (RBF) is nog 'n metode wat gebruik is om die empiriese verwantskap tussen die veranderlikes te bepaal. Die RBF interpolasie, wat gedoen is met 200 punte gekies deur D-optimale seleksie, lewer 'n RMS fout van 9.617 μatm sonder breedtegraad as veranderlike en 'n RMS fout van 6.716 μatm met breedtegraad ingesluit as veran-

derlike. Wanneer die veranderlikes optimaal skaleer word met die RBF interpolasie, met 200 punte wat met D-optimale seleksie gekies is, lewer dit 'n RMS fout van 9.012 μatm sonder breedtegraad as veranderlike en 'n RMS fout van 4.064 μatm met breedtegraad ingesluit as veranderlike.

Sleutelwoorde

Suidelike Oseaan; Aardverwarming; Genetiese algoritme; Radiale basisfunksie-interpolasie; Krommepassing

Contents

Definitions

Advection: The transfer of a property of the atmosphere, such as heat, cold or humidity, by the horizontal movement of an air mass. The rate of change of an atmospheric property caused by the horizontal movement of air [18].

Anthropogenic: Anthropogenic refers to something that is caused by humans or human activity [18].

Buffer capacity: Buffer capacity is referred to as the relative ability of a buffer solution to resist pH change when an acid or base is added to the solution [17].

Carbon budget: An amount of carbon dioxide that a country, company or organization has agreed upon to be the largest amount that it will produce in a particular period of time [3].

Carbon cycle: The combined processes of photosynthesis, decomposition and respiration by which carbon compounds are transferred between the major carbon reservoirs, including the atmosphere, ocean and living organisms [18].

Carbon dioxide flux: (Or CO_2 flux) Refers to the CO_2 transfer between the ocean and the atmosphere. It is the flow of CO_2 particles through a given surface [18].

Carbon sink: A carbon sink is a terrestrial or oceanic area that absorbs carbon dioxide released by the burning of fossil fuels [16].

Entrainment: To carry (suspended particles, for example) along in a current [18].

Fugacity: The thermodynamic property of a gas that is related to its partial pressure. It is an indication of the tendency of the gas to escape or expand [18].

Kyoto Protocol: The Kyoto Protocol is an international agreement linked to the United Nations Framework Convention on Climate Change. In short, the Kyoto protocol sets targets for 37 industrialized countries and the European community for reducing greenhouse gas emissions. This protocol

commits the countries to reduce the emissions by 2012. It is generally seen as an important first step towards a fully global emission reduction regime that will contribute to stabilizing greenhouse gas emissions [77].

Nautical mile: The nautical mile is a unit of length used in marine navigation [8]. One nautical mile = 1.852 km [8].

Ocean deep layer: The deep layer of the ocean is the bottom part of the ocean, and makes up most of the ocean's volume (90% of the ocean) [60], [14]. The density of the water increases as the depth of the ocean increases in the deep layer [60]. The colder, denser waters sink down to the deep waters of the ocean where the deep ocean layer acts as a "store" for these waters [14].

Ocean surface mixed layer: The ocean surface mixed layer is assumed to be approximately the upper 500m of the ocean [58]. This is the layer that is considered when studying climate, biological activity and pollution of the ocean [11]. Satellites, ships and aeroplanes can easily monitor the mixed layer dynamics, thus we know the most about this layer of the ocean [60].

Pack ice: Ice in the polar regions of the ocean, that consists of floating ice that joined together to cover the sea surface such that there is little or no open water in these parts [18].

Polar Frontal Zone: The Polar Frontal Zone in the Southern Hemisphere is the area where the salinity of the ocean is low, between the Antarctic Polar Front and the Subantarctic Front [15]. This zone is shown in Figure 1.

Radial: Moving or directed along a radius, or similarly developing symmetrically about a central point for a specified radius [18]. Spreading out or developing uniformly on all sides of a central point [16].

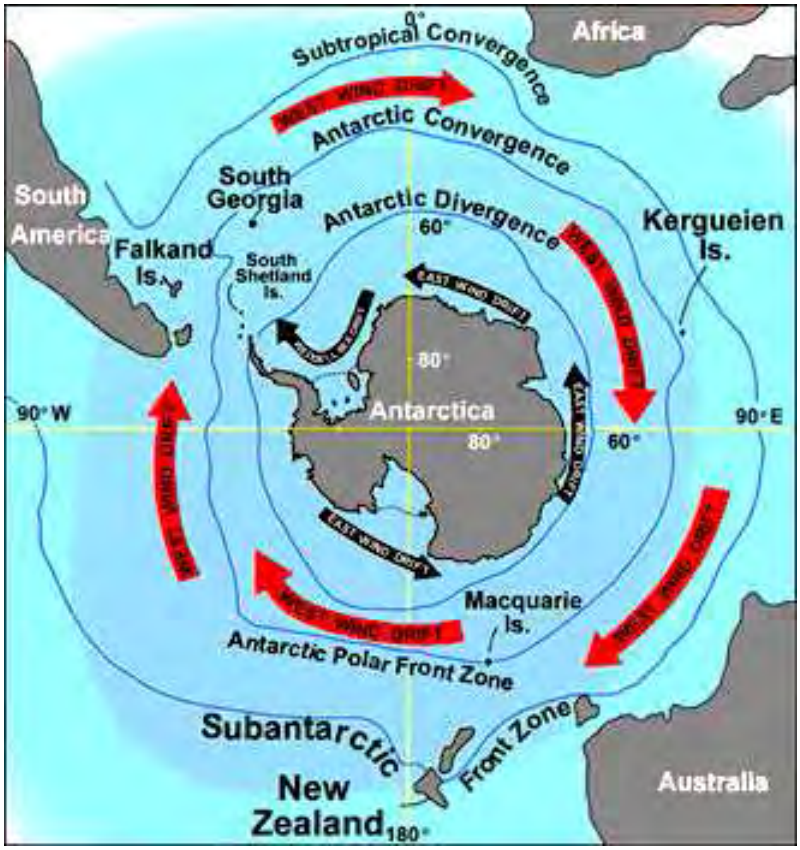

Figure 1: Position of the Polar Frontal Zone [21].

Acronyms

AATSR	Advanced Along Track Scanning Radiometer
ACC	Antarctic Circumpolar Current
AVHRR	Advanced Very High Resolution Radiometer
CDIAC	Carbon Dioxide Information Analysis Center
CFSR	Climate Forecast System Reanalysis
CLIVAR	Climate Variability
CSIR	Center for Scientific and Industrial Research
DDS	Diagnostic Data Sets
DIC	Dissolved Inorganic Carbon
DoE	Department of Energy
EEZ	Exclusive Economic Zone
GA	Genetic Algorithm
GEOSECS	Geophysical Sections Experiment
GIS	Geographic Information System
GMES	Global Monitoring for Environment and Security
GODAS	Global Ocean Data Assimilation System
HIRS	High-resolution Infra-red Radiation Sender
HPLC	High Performance Liquid Chromatography
IOCCG	International Ocean Colour Coordinating Group
IOCCP	International Ocean Carbon Coordination Project
ISIN	Integerised Sinusoidal projection
JGOFS	Joint Global Ocean Flux Study
LDEO	Latmont Doherty Earth Observatory
MERIS	Medium Resolution Imaging Spectrometer

MLD	Mixed Layer Depth
MODIS	Moderate Resolution Imaging Spectrometer
NODC	National Oceanographic Data Centre
NCEP	National Centres for Environmental Prediction
OC-TAC	Ocean Colour Thematic Assembly Centre
OISST	Optimum Interpolation Sea Surface Temperature
pCO_2	Partial pressure of Carbon Dioxide
PiRATA	Prediction and Research Moored Array in the Tropical Atlantic
POC	Particulate Organic Carbon
POOZ	Permanent Open Ocean Zone
PSU	Practical Salinity Unit
RAMA	Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction
RBF	Radial Basis Function
RMS	Root Mean Square
SEM	Scanning Electron Microscopy
SIZ	Seasonal Ice Zone
SOCAT	Surface Ocean CO_2 Atlas
SRP	Strategic Research Project
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
SSU	Stratospheric Sounding Unit
THC	Thermohaline Circulation
TOVS	TIROS-N Operational Vertical Sounder
UCTD	Underway Conductivity Temperature and Depth
VOS	Volunteer Observing Ships
WOA05	World Ocean Atlas 2005
WOCE	World Ocean Circulation Experiment
XBT	Expendible Bathythermograph

Nomenclature

Symbol	Description
$\boldsymbol{\alpha}$	Vector of coefficients for the RBF
a_i	The i^{th} scaling coefficient of the variables
$\mathbf{A}$	The $m \times n$ matrix where m is the number of rows in the entire data set and n is the number of terms in the equation that is considered
α_j	Coefficient of the j^{th} term in the radial basis function
$\mathbf{B}$	The $k \times n$ matrix obtained by extracting k rows of matrix $\mathbf{A}$
$\boldsymbol{\beta}$	Vector of coefficients for the polynomial part of RBF interpolation
Chl	Chlorophyll-a concentration
Chl_j	j^{th} chlorophyll-a concentration value
Chl_{max}	Maximum chlorophyll-a concentration value
Chl_{min}	Minimum chlorophyll-a concentration value
e	Measurement error
$\mathbf{E}$	Error on the estimation
f	Fitness
fCO_2	Fugacity of CO_2
$\boldsymbol{\gamma}$	Vector of coefficients of the RBF interpolation $\begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}$
$\mathbf{I}$	Identity matrix
k	Number of points sampled from the data set for processing
Lat	Latitude
Lat_j	j^{th} latitude value
Lat_{max}	Maximum latitude value
Lat_{min}	Minimum latitude value

m	Number of points in the complete data set
$\mathbf{M}$	Fisher information matrix
$\mathbf{M}_{b,b}$	Matrix for the RBF part of the RBF interpolation
MLD	Mixed layer depth
MLD_A	Mixed layer depth at point A
MLD_B	Mixed layer depth at point B
MLD_C	Mixed layer depth at point C
MLD_j	j^{th} MLD value
MLD_{max}	Maximum MLD value
MLD_{min}	Minimum MLD value
n_b	Number of points selected for RBF interpolation
OF	Objective function
$\mathbf{P}$	Variance covariance matrix
pCO_2	Partial pressure of CO_2
$\mathbf{P}_b$	Matrix for the polynomial part of the RBF interpolation
P_{equ}	Pressure in the equilibrator
pH_2O	Water vapour pressure of CO_2
ϕ	Radial basis function
P_{water}	Equilibrium of vapour pressure at the temperature of equilibration
r_j	Radius defined as $\parallel \mathbf{x} - \mathbf{x}_j \parallel$
$\mathbf{R}$	Matrix for RBF interpolation
$\mathbf{S}$	Sensitivity matrix
SS	Specific sampling schedule
SST	Sea surface temperature
σ	Standard deviation
σ^2	Variance
σ_{random}	Standard deviation of randomly sampled points
$\sigma_{D-optim}$	Standard deviation of D-optimally sampled points

T	Sea surface temperature
T_{eq}	Temperature of equilibration
T_j	j^{th} Temperature value
T_{max}	Maximum temperature value
T_{min}	Minimum temperature value
u	Input signal
x_i	i^{th} x value
x_{max}	Maximum x value
x_{min}	Minimum x value
x_{scaled}	Scaled x values
XCO_2	Mole fraction of CO_2
ξ	The ratio of the x value divided by the radius r
y_{avg}	Average modelled pCO_2 value
y_i	Modelled pCO_2 values for the i^{th} point in space
$\mathbf{y}$	Actual pCO_2 values
$\tilde{\mathbf{y}}$	Modelled pCO_2 values

List of Figures

List of Tables