Re-tuning the Walker-Kasting global carbon cycle box model using a parameter sensitivity analysis

Simon Yang*[1], David A. Carozza[1], Lawrence A. Mysak[1]

[1]Department of Atmospheric and Oceanic Sciences, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 2K6

ABSTRACT

Introduction: The Walker-Kasting global carbon cycle box model is a simple representation of the Earth system used to study climatic events. This model has a high number of parameters whose sensitivity must be tested in order to better understand which of them dominate the behaviour of the model. In this study, we perform a parameter sensitivity analysis. Moreover, we use these results to re-tune the model to preindustrial conditions using a quantitative criterion. We then compare our results to those determined by Walker-Kasting. **Methods**: We achieved the parameter sensitivity analysis by calculating, for each parameter, an index that measures the impact of a change in the initial parameter value on the equilibrium solutions. The most sensitive parameters were determined and then tuned in the model by comparing the model equilibrium solutions to a set of 32 experimental values. **Results**: We found that nine of the tuning parameters were sensitive to a change to their initial value. Furthermore, we discovered that 5 of these parameter values were identical to those determined by Walker-Kasting, thus affirming their work. **Discussion**: A sensitivity analysis is interesting to perform because it allows the users of a model to more fully comprehend the way in which the model reacts to changes in its parameters. Sensitivity analysis is fundamental in the tuning of a model (for example, to a particular period in the Earth's history) since it allows researchers to consider only the most important parameters.

KEYWORDS

Walker-Kasting, Box model, Parameter sensitivity analysis

*Corresponding author:
simon.yang@mail.mcgill.ca

Received: 11 January 2010
Revised: 17 March 2010

INTRODUCTION

A box model is a simplified version of a complex system such that the components of the system are reduced to linked boxes (or reservoirs). Models of this type are employed to simulate processes in the climate system in a rudimentary way, and are used to determine whether model output satisfactorily describes an observed phenomenon. Because they are simple in structure and computationally efficient, box models are ideal for analyzing climate processes that occur on long time scales. In particular, a box model of the Earth system depicts the components of the Earth system (atmosphere, ocean, terrestrial biosphere, etc.) as boxes that are linked by exchanges of mass, energy or both.

In general, box models possess a large number of parameters that must be tuned with respect to a set of experimental data. This is the case because box models are not inherently based on

physical laws, but on parameterizations of those laws. Researchers perform such a tuning procedure by determining the set of model parameter values that generates the best equilibrium results (steady state solutions) in terms of a specific criterion, such as minimizing the root mean squared error between the model and experimental results.

An important element in the development of a box model is therefore the determination of which parameters have the greatest impact on the model equilibrium results. Such a procedure must be undertaken since it is computationally unfeasible to run a model, even a computationally simple one, for all values of the parameter space. Therefore, it is interesting to know which parameters are the most sensitive since it shows which parameters need to be considered with greater accuracy and which parameters can be ignored in the tuning process.

In this study, we perform the first sensitivity analysis for the Walker-Kasting (*1*) box model (denoted WK92) and thus determine the most influential model parameters. These results allow us to tune the model to preindustrial conditions using a quantitative criterion. Furthermore, we compare our tuning results to those of the qualitative approach used by WK92, and discuss applications of our sensitivity analysis.

METHODS

MODEL

The preindustrial global carbon cycle is the biogeochemical cycle that comprises both: i) the carbon stored within different reservoirs of the Earth system and ii) the exchange of carbon between these components. The WK92 box model is a simple representation of the preindustrial global carbon cycle, which consists of eight model reservoirs: atmosphere, terrestrial biomass, cold surface ocean, warm surface ocean, thermocline, Deep Atlantic Ocean, Deep Indian Ocean and Deep Pacific Ocean (Fig. 1).

The model is governed by 32 ordinary differential equations that represent thermohaline (temperature and salinity) fluxes and mixing fluxes of carbon and nutrients between the reservoirs, as well as biogeochemical processes such as photosynthesis, respiration and the oceanic biological pump, which are biological processes that transport carbon from the ocean surface to the deep ocean. For each reservoir, the model calculations include: atmospheric CO_2 concentration, lysocline depth (the ocean depth below which the rate of dissolution of calcium carbonate increases dramatically), average surface air temperature and δ13C for each reservoir. Table 1 lists the prognostic variables. Note that δ13C is the ratio of the rarer ^{13}C isotope of carbon to the more common ^{12}C isotope, relative to a generally recognized standard ratio of

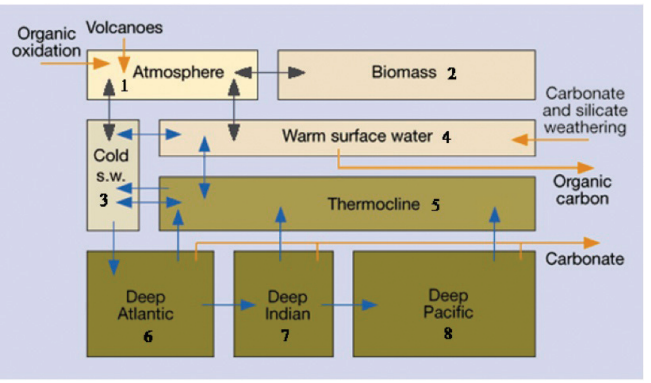

Fig. 1: Schematic of the WK92 model reservoirs (adapted from Dickens, 1999). Exchanges of carbon internal to the model are represented by blue and black arrows. Blue single- and double-headed arrows between the ocean reservoirs represent thermohaline and mixing fluxes, respectively. Black arrows between the atmosphere and biomass represent respiration and photosynthesis, whereas those between the atmosphere and ocean reservoirs represent diusive exchanges of carbon. Exchange fluxes external to the model are represented by orange arrows. Figure taken from Carozza (2009)

Table 1: Walker and Kasting (1992) model prognostic variables. PAL = preindustrial atmospheric level; 1 PAL = 280 ppmv. In the WK92 model, there are six equations for each of P, , and A in the ocean reservoirs and one equation for each of TS and [CO2]. In addition, there is one equation for MBio, three equations for R in the deep ocean reservoirs and eight equations for 13C. Note that this adds up to 32 equations.

Variable	Name	Chemical Representation	Unit
P	Dissolved phosphate	HPO_4^{2-}	mmol m^{-3}
Σ	Total dissolved carbon	$HCO_3^- + CO_3^{2-} + H_2CO_3$	mol m^{-3}
A	Alkalinity	$HCO_3^- + 2CO_3^{2-}$	mol m^{-3}
$[CO_2]$	Carbon dioxide	$pCO_{2(g)}$ or $pCO_{2(aq)}$	PAL
T_S	Surface temperature	-	K
M_{Bio}	Biomass	-	10^{18} mol C
R	Pelagic carbonates	-	10^{18} mol C
$\delta^{13}C$	$\delta^{13}C$	-	‰

the two. It is an important variable because a change indicates a flux of carbon into or out of the system in question. A list of the model parameters can be found in Table 2.

SENSITIVITY ANALYSIS

To calculate the sensitivity of the WK92 model parameters, we implemented the method described by Nordhaus (*2*). First, we determined an appropriate range for the model parameters under consideration. Each range was defined by a minimum and a maximum value (extreme values), and was determined based on a literature review of the specific parameter. We then compared the equilibrium results of the simulations with the extreme parameter values to the results with the initial parameter value by means of the following sensitivity index:

$$I_i = \sqrt{\frac{1}{n}\sum_t \left(\frac{X_i^M(t) - X_i^*}{X_i^*}\right)^2}, \quad (1)$$

where $X_i^M(t)$ is the time series of the i^{th} model variable for the extreme parameter values, X_i^* is the time series of the i^{th} model variable for the initial parameter value, and n is the number of points in the time series. Note that the time series of the model variable i that is under investigation, $X_i^M(t)$, evolves in time until it reaches equilibrium and is thus a function of time. However, the time series of the model variable under the initial parameter value, X_i^*, is a constant function, since it begins at the equilibrium value. The sensitivity index therefore does not only measure the difference between the equilibrium values for the initial and extreme parameter values, but also the distance between the two time series.

The deviation is then normalized by dividing the difference $X_i^M(t) - X_i^*$ by X_i^*, so that the quantity in parentheses in Equation 1 is dimensionless. This normalization allows us to compare sensitivity indices for every parameter. The I values were then averaged to obtain an extreme sensitivity index I for each parameter. To take both the maximum and minimum parameter results into account, the two sensitivity indices are averaged. The initial, minimum and maximum values for each model parameter, the sensitivity indices and the averaged sensitivity index I are given in Table 3. The nine most sensitive parameters are presented in Figure 2.

Table 2: Model parameters and their description

Parameter	Description
f_{PT}	Fraction of particulate matter dissolved in the thermocline
R_{CP}	Ratio of particulate matter organic carbon to phosphorus
R_{CO}	Ratio of particulate matter $CaCO_3$ to organic carbon
z_1	Overall lysocline depth
z_2	Rate of change of lysocline with respect to $[CO_3^{2-}]$
C_{bio}	Initial biomass
a_T	Total surface area of surface ocean reservoirs
uv	Upwelling velocity in m/y
vmv	Vertical mixing velocity in m/y
swmt	Surface water mixing time in years
tcmt	Thermocline mixing time in years
α	Planetary albedo
f_{PD}	P_4 / P_8
P_R	Concentration of phosphate in river water
τ_{OA}	CO_2 dissolution time in years
M_{atm}	CO_2 mass in the atmosphere
τ_{Bio}	Forest growth and decay time in years
sfcarb	Reactive pelagic carbonate per square meter
a_3 / a_T	Fraction of cold surface ocean area to ocean area
C_{surf}	Earth's surface heat capacity
P_4	Phosphate in the warm surface ocean
P_8	Phosphate in the Deep Pacific

Table 3: Initial, minimum, and maximum parameter values and sensitivity index 1. The subscripts are the reservoir numbers (Fig. 1)

Parameter	Initial value	Test value	Sensitivity index I	Average
f_{PT}	0.925	0.950 0.900	0.165 0.092	0.128
R_{CP}	120.000	140.000 100.000	0.072 0.077	0.074
R_{CO}	0.090	0.200 0.030	0.382 0.670	0.526
z_1	5.800	6.300 5.300	0.033 0.029	0.031
z_2	50.000	60.000 40.000	0.020 0.032	0.026
C_{bio}	0.200	0.400 0.100	0.048 0.186	0.117
uv	1.150	1.380 0.920	0.032 0.046	0.028
vmv	10.500	12.600 8.400	0.034 0.043	0.039
swmt	50.000	60.000 40.000	0.011 0.015	0.013
tcmt	250.000	300.000 200.000	0.008 0.009	0.009
α	0.300	0.308 0.292	0.020 0.018	0.019
f_{PD}	0.040	0.048 0.032	0.019 0.019	0.019
P_R	0.001	0.001 0.001	0.024 0.014	0.019
τ_{OA}	10.000	12.000 8.000	0.001 0.001	0.001
M_{atm}	0.050	0.060 0.040	0.001 0.001	0.001
τ_{Bio}	50.000	60.000 40.000	0.000 0.000	0.000
sfcarb	0.0004	0.0008 0.0002	0.0000 0.0000	0.000
a_3/a_T	0.360	0.432 0.288	0.044 0.055	0.037
C_{surf}	50.300	60.360 40.240	0.000 0.000	0.000
careafr	0.003	0.005 0.001	0.000 0.000	0.000

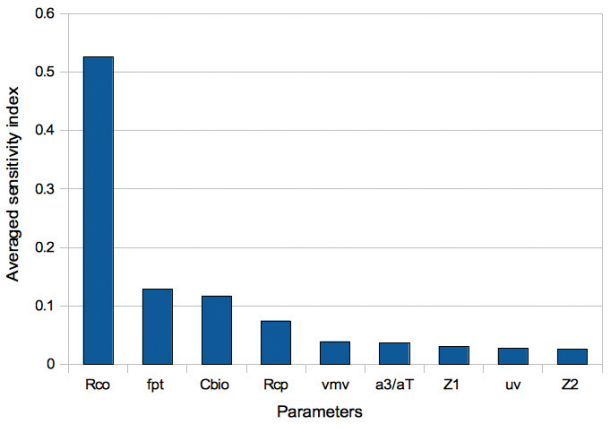

Fig. 2: Averaged sensitivity index I (see table 3) for the nine most sensitive model parameters, presented in decreasing order.

Table 4: Selected proxy values and their sources. The subscripts of the variables refer to a specific reservoir (see Fig.1) i.e., P_3 refers to phosphate in the cold surface water reservoir.

Variables	units	Range	selected value	source
P_3	$10^{-3}molm^{-3}$	0.7 - 1.0	0.85	*Walker and Kasting* (1992) fig.7 p.163
P_4	$10^{-3}molm^{-3}$	0.1 - 0.2	0.15	*Walker and Kasting* (1992) fig.7 p.163
P_5	$10^{-3}molm^{-3}$	0.2 - 2.5	1.5	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
P_6	$10^{-3}molm^{-3}$	1.0 - 2.5	1.75	*Walker and Kasting* (1992) fig.7 p.163
P_7	$10^{-3}molm^{-3}$	2.3 - 2.5	2.4	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
P_8	$10^{-3}molm^{-3}$	2.0 - 3.25	2.625	*Walker and Kasting* (1992) fig.7 p.163
Σ_3	$10^{-3}molm^{-3}$	2.14 - 2.23	2.185	*Broecker and Peng* (1982), fig-2-9 p.70
Σ_4	$10^{-3}molm^{-3}$	1.90 - 2.02	1.96	*Broecker and Peng* (1982), fig-2-9 p.70
Σ_5	$10^{-3}molm^{-3}$	2.0 - 2.3	2.15	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
Σ_6	$10^{-3}molm^{-3}$	2.16 - 2.20	2.18	*Broecker and Peng* (1982), fig-2-9 p.70
Σ_7	$10^{-3}molm^{-3}$	2.34 - 2.40	2.37	*Broecker and Peng* (1982), fig-2-9 p.70
Σ_8	$10^{-3}molm^{-3}$	2.33 - 2.38	2.355	*Broecker and Peng* (1982), fig-2-9 p.70
A_3	$10^{-3}molm^{-3}$	2.36 - 2.40	2.38	*Broecker and Peng* (1982), fig-2-9 p.70
A_4	$10^{-3}molm^{-3}$	2.275 - 2.34	2.31	*Broecker and Peng* (1982), fig-2-9 p.70
A_5	$10^{-3}molm^{-3}$	2.3 - 2.35	2.325	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
A_6	$10^{-3}molm^{-3}$	2.325 - 2.355	2.34	*Broecker and Peng* (1982), fig-2-9 p.70
A_7	$10^{-3}molm^{-3}$	2.375 - 2.39	2.382	*Broecker and Peng* (1982), fig-2-9 p.70
A_8	$10^{-3}molm^{-3}$	2.44 - 2.475	2.458	*Broecker and Peng* (1982), fig-2-9 p.70
M_{Bio}	mol C	-	0.2×10^{18}	*Walker and Kasting* (1992)
T_S	°C	-	15.0	*Walker and Kasting* (1992)
pCO_2	ppmv	-	280	*Walker and Kasting* (1992)
$z_{iy,6}$	km	-	4.0	*Broecker and Peng* (1982), fig-2-14 p.77
$z_{iy,7}$	km	-	4.0	*Broecker and Peng* (1982), fig-2-15 p.77
$z_{iy,8}$	km	3.5 - 4.0	3.75	*Broecker and Peng* (1982), fig-2-16 p.77
$\delta^{13}C_1$	‰	-	-7	*Ruddiman* (2001), fig.11-10 p.242
$\delta^{13}C_2$	‰	-	-22.0	*Panchuk et al.* (2008)
$\delta^{13}C_3$	‰	1.0 - 2.0	1.5	*Broecker and Peng* (1982), fig-6-13 p.310
$\delta^{13}C_4$	‰	1.5 - 2.4	1.95	*Broecker and Peng* (1982), fig.6-12 p.309
$\delta^{13}C_5$	‰	-	1.2	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
$\delta^{13}C_6$	‰	0.5 - 1.5	1.0	*Broecker and Peng* (1982), fig.6-12 p.309
$\delta^{13}C_7$	‰	-	0.4	*WOCE* (1990-1998) Atlas volume 4 : Indian Ocean
$\delta^{13}C_8$	‰	-0.5 - 0.8	0.65	*Broecker and Peng* (1982), fig.6-12 p.309

Table 5: Parameters that are to be tuned and their test values. The values in bold are the values that were selected during the tuning, i.e., together, these values are the combination of parameters that gave the equilibrium solutions that were closest to the experimentally derived values.

Parameter	WK92 value	test values
R_{CO}	0.09	0.03, 0.045, 0.06, 0.075, **0.09**, 0.105, 0.12, 0.135
f_{PT}	0.925	0.9, 0.91, 0.92, **0.925**, 0.93, 0.94
C_{bio}	0.2	0.10, 0.15, 0.20, 0.25, **0.30**, 0.35, 0.40
R_{CP}	120.0	100.0, 110.0, 120.0, **130.0**, 140.0
vmv	10.5	9.45, **10.5**, 11.55
a_3/a_T	0.362033	**0.2896**, 0.362033, 0.4344396
z_1	5.8	5.3, **5.8**, 6.3
uv	1.15	0.92, **1.15**, 1.38
z_2	50.0	40.0, 50.0, **60.0**

TUNING

GENERAL METHOD

In the original work of WK92, the set of model parameters was qualitatively determined by arbitrarily selecting a set of parameters and then adjusting those parameters until a reasonable match to the experimental results was found. As an application to the sensitivity analysis described in the previous section, we have re-tuned the WK92 box model using the most sensitive model parameters.

The goal of tuning is to determine the parameter values that give equilibrium solutions that best reproduce experimentally derived values (Table 4). To achieve this, we must first find the experimentally determined values. Based on the nine most sensitive parameters (Fig. 2), values between the minimum and maximum parameter value were chosen to be part of the tuning process (Table 5). The number of values chosen for each parameter, which ranged from three to eight, depended on the sensitivity of the parameter. Otherwise put, the more sensitive parameters required greater accuracy, and thus more values were used for their calculation.

Equilibrium runs involving every combination of these chosen parameter values were then implemented. This procedure can represent a tremendous number of runs and is why only the most sensitive parameters were used in the tuning. In our study, we performed 408 240 simulations (each running for a period of 2 million model years with a time-step of 1000 model years) to determine the best set of nine tuning parameters. Every set of equilibrium variables was compared to the set of experimental values using the following cost function:

$$\sqrt{\sum_{i=1}^{32}\left(\frac{Xe_i - Xp_i}{Xp_i}\right)^2} \quad (2)$$

where Xe_i is the i^{th} member of the 32 equilibrium solutions, and Xp_i is the i^{th} member of the 32 experimentally derived values. Again, the function is normalized (the difference $Xe_i - Xp_i$ is divided by Xp_i) so as to make the quantity in parentheses in (Eq. 2) dimensionless. A cost, or normalized cumulated root mean squared error, was calculated for every set of equilibrium solutions. Finally, the set of equilibrium solutions that had the smallest cost value was selected as the optimal set of parameter values and equilibrium solutions (Table 5).

EXPERIMENTAL DATA

We reviewed the literature to find experimental preindustrial values of the 32 dependent variables. Important sources for experimental values were *Tracers in the Sea* by Broecker and Peng and *Ocean Biogeochemical Dynamics* by Sarmiento and Gruber (*3*, *4*). These texts provided us with the concentration of phosphate, alkalinity, the lysocline depth and $\delta^{13}C$ values for several of the model boxes. Moreover, the World Ocean Circulation Experiment (WOCE) website provided vertical cross section maps of the concentration of alkalinity, phosphate, $\delta^{13}C$ and total dissolved carbon for the Deep Indian Ocean and the thermocline (*5*). Although the WOCE data represents the present day, we assume that they are a reasonable representation of these two reservoirs in the preindustrial era. The $\delta^{13}C$ values for terrestrial biomass and the atmosphere were taken from Ruddiman (*6*). We often found plausible ranges of values rather than specific results, and therefore selected the value to tune by taking the

average of the maximum and minimum for a given range. Note that measurements in the thermocline vary tremendously as it is a non-homogeneous layer. The experimentally derived data for this reservoir were therefore estimates chosen from the middle of the thermocline layer. The 32 experimental values of the model variables and the associated references are listed in Table 4.

RESULTS

Figure 2 presents the averaged sensitivity index of the nine most sensitive parameters. The most sensitive parameter is R_{CO}. When tested from its initial value to its minimum value, we found a sensitivity index of 0.67 (Table 3). This implies that the 32 model variables change by an average factor of 0.67 when R_{CO} is changed from its initial value to its minimum. We find that other parameters, such as z_2 or α, are less sensitive and therefore less important since they change the model variables by a smaller factor. Some parameters, such as C_{surf}, have no influence on the model equilibrium solutions (Table 3).

Table 4 presents the 32 model prognostic variables, the experimental range of these variables and the selected value applied in the tuning procedure. The WK92 parameter values, the values examined in the tuning procedure, and the values that were selected from the tuning (bold) are given in Table 5.

DISCUSSION

SENSITIVITY

R_{CO}, the parameter that occurs in most of the model equations, is more sensitive than the other parameters. Hence, a change in its value will affect the model equilibrium solutions more than the other parameters, making it the most important tuning parameter. In addition, changes in certain parameters exhibit a sensitivity index of zero in equilibrium solutions (Table 3). This result can be explained by analyzing the model equations. Considering the C_{surf} parameter, Table 3 indicates that a change in the value of C_{surf} does not influence any of the equilibrium solutions. It does not, for example, affect T_s, the average global surface temperature. The equation relating C_{surf} and T_s is the following:

$$\frac{d}{dt}T_s = \frac{Q - F_{IR}(T_s)}{C_{surf}}, \tag{3}$$

where the terms in the numerator of the right hand side represent incoming and outgoing solar radiation, respectively. At equilibrium, $dT_s = 0$, and since C_{surf} is constant, $Q - F_{IR}(T_s)$ must be zero. Hence, changing the value of C_{surf} will not affect the equilibrium value of T_s, and accordingly, the sensitivity of C_{surf} with respect to T_s is zero. For this reason, C_{surf} is not a relevant tuning parameter. In a similar manner, the other parameters that exhibit a sensitivity index of zero are also not pertinent tuning parameters.

Note that for a simulation where carbon is being released into the atmosphere, $Q - F_{IR}(T_s)$ is not zero because the amount of outgoing solar radiation, $F_{IR}(T_s)$, is changing. In this case, the argument presented in the previous paragraph is not valid. Choosing two different values of C_{surf} for the same carbon emission scenario will indeed generate two different evolutions of T_s. Hence, although a parameter such as C_{surf} is irrelevant with respect to tuning the model, it is nevertheless important in the evolution of a simulation where the model is being forced by a release of carbon.

TUNING

We found that five out of the nine most sensitive parameter values resulting from our tuning procedure were identical to those determined by WK92 (Table 5). This is a strong affirmation of the parameters determined by WK92. Among the remaining four parameters, the difference between the original parameter values and those we calculated can be explained by the fact that our tuning procedure was more rigorous; it contained significantly more model simulations and a greater number of experimentally derived values for the dependent variables.

The goal of this study was to perform a sensitivity analysis on the parameters of the WK92 carbon cycle box model. The sensitivity analysis allowed us to understand which parameters most affected the equilibrium solutions and study how the model reacted to a change in parameter values. Furthermore, it permitted us to calculate the nine most sensitive parameters, to use this set of parameters to quantitatively tune the WK92 model and to compare and contrast our tuned parameter results to those originally determined qualitatively by WK92. In future work, these sensitivity results will be used to tune the WK92 box model to the Paleocene/Eocene boundary period (approximately 55 million years ago) so that it may be used to analyze the Paleocene Eocene thermal maximum, a period of abrupt and intense global warming (*7*, *8*).

ACKNOWLEDGMENTS

S.Y. was supported by an Undergraduate Student Research Award from the Natural Sciences and Engineering Research Council of Canada (NSERC). This work was also supported by an NSERC discovery grant awarded to L.A.M.

REFERENCES

1. J.C.G. Walker, J.F. *Kasting, Palaeogeography, Palaeoclimatology, Palaeoecology* **97** 151–189 (1992).
2. W.D. Nordhaus, *Managing the Global Commons* (The MIT Press, Cambridge, Mas- sachusetts, 1994).
3. W.S. Broecker, T.H. Peng, *Tracers in the Sea* (Eldigio Press, Lamont Doherty Geological Observatory, Palisades, 1982).
4. J.L. Sarmiento, N. Gruber, *Ocean Biogeochemical Dynamics* (Princeton University Press, Princeton, 2006).
5. WOCE, *World Ocean Circulation Experiment Atlas volume 4* (1990-1998) http://www.pord.ucsd.edu/whp atlas/indian index.html.
6. W.F. Ruddiman, *Earth's Climate, Past and Future* (W.H. Freeman and Company, New York, 2001).
7. G.R. Dickens, M.M. Castillo, J.C.G. Walker, *Geology* **25** 259–262 (1997).
8. D. Carozza, *Carbon cycle box modeling studies of the Paleocene-Eocene thermal maximum* (M.Sc. Thesis, McGill University, Montreal, Canada, 2009).