

A Multivariable Regression Analysis of the Vertical Distribution of TCO_2 in the Eastern Pacific

SAM BEN-YAAKOV

Department of Geology, University of California, Los Angeles 90024

Multiple regression analysis was performed on a number of variables along two vertical profiles in the eastern Pacific. It is shown that a simple linear mixing model is consistent with the linear relation that has been found between total dissolved inorganic CO_2 , dissolved oxygen, total alkalinity, and temperature along the profiles. The regression coefficient for oxygen was found to be in the range 0.374 ± 0.014 (95% confidence limits) which is close to the value predicted by Redfield, Ketchum, and Richards (1963). The study suggests that two other processes, besides carbonate reaction, may be responsible for alkalinity variations in the oceans.

Oceanic relations between total inorganic CO_2 (TCO_2) and dissolved oxygen (O_2) are frequently studied through an intermediate variable called apparent oxygen utilization (AOU). This parameter was suggested by Redfield *et al.* [1963] as a coarse measure to the in situ oxygen variations. The basic assumption is that any given sea-water sample (or its parents, if it is a mixture) was at one time at the ocean surface and was in equilibrium with the partial pressure of oxygen in the atmosphere. Under this assumption and by using the solubility coefficient of oxygen at the potential temperature and salinity one can calculate the preformed concentration of dissolved oxygen in the sample when it was at the surface (O_2'). The apparent oxygen utilization is obtained by the relation

$$\text{AOU} = \text{O}_2' - \text{O}_2 \quad (1)$$

where O_2 is the measured concentration of dissolved oxygen in the sample. This procedure is clearly only a rough approximation of real oceanic processes since it implicitly assumes that surface sea water is in equilibrium with atmospheric O_2 and that the solubility of O_2 in sea water is linear with temperature and salinity. Neither of these assumptions is accurate [Redfield, 1948; Weiss, 1970; Culberson and Pytkowicz, 1970]. Nevertheless, the concept of AOU was found to be useful in various studies. For example, Redfield *et al.* [1963] showed that

the variation in phosphorous, nitrogen, carbon, and oxygen in the ocean are consistent with a proposed model which predicts that the variation in concentration of these elements should be in the proportion 1:16:106:276. The TCO_2 - O_2 relationships in the ocean can also be studied through the use of the AOU parameter. The procedure in this case [Park, 1965] consists of normalizing TCO_2 to constant chlorinity and correcting it for total alkalinity variations through the assumption that alkalinity changes are only due to solution or precipitation of carbonates. However, previous TCO_2 and AOU correlations were found to be linear only in restricted oceanic areas or over short segments of vertical profile, which may indicate perhaps the inaccuracies in the AOU estimations.

The procedure of using AOU as a measure of the in situ variation of dissolved oxygen was criticized by Craig [1969] who did not find a linear relation of TCO_2 production to O_2 consumption over the Pacific deep and intermediate waters. He concluded that the procedure of estimating the in situ oxygen variations by the auxiliary parameter AOU is inconsistent with a mixing model which assumes that diffusive effects are important. Culberson and Pytkowicz [1970] attempted to overcome this problem by introducing another correction that normalized the measured TCO_2 to the preformed value. They showed that this additional modification improves the linear correlation between TCO_2 and O_2 and that the TCO_2/O_2 variation ratio

obtained by this procedure is consistent with the model of *Redfield et al.* [1963].

The different procedures used during previous studies to correlate $\text{TCO}_2\text{-O}_2$ variations in the oceans were in essence, an attempt to reduce a multivariable problem to a two-variable regression. Such a procedure can be successful only if the mutual relation between the variables are known a priori. For example, when using these procedures, one generally assumes that alkalinity variations are only due to solution or precipitation of carbonates and the measured TCO_2 is usually 'corrected' by the relation

$$\Delta \text{CO}_2 = 0.5\Delta TA \quad (2)$$

[*Culberson and Pytkowicz*, 1970] where ΔTA is the change of total alkalinity from the performed value. However, additional contributions from organic alkalinities [*Gripenberg*, 1960] or silicates [*Garrels*, 1965] may be responsible, in part, for the variations of alkalinity in the oceans. It is clear, therefore, that any attempt to reduce the multivariable problem to a two-variable relation ($\text{TCO}_2\text{-O}_2$) may result in some inaccuracies due to the fact that most of the intervariable dependencies are not as yet fully understood.

The present study is an attempt to use a different approach for correlating $\text{TCO}_2\text{-O}_2$ variations in the ocean. I shall use here a multivariable regression analysis to explore possible characteristic dependencies between TCO_2 and O_2 variations in the Pacific. The results of this analysis will later be compared to different mixing models that may explain the dependencies that have been obtained from the study.

It has been shown by *Wyrki* [1962] and *Munk* [1966] that the vertical distribution of T , O_2 , and S in the oceans is not inconsistent with a simple mixing model involving a constant upward vertical velocity and eddy diffusivity. Namely, temperature, salinity, and O_2 concentration at any given point along a vertical profile can be predicted by the solution of the continuity equation which contains both advective and diffusive parameters. The particular solution at any given point depends on the boundary conditions of the profile. *Craig and Weiss* [1968] have shown that oceanic argon data are also consistent with this model. If a mixing process can explain the distribution of

variables along a vertical oceanic profile, one should be able to explore some of the intervariable relations by regression analysis on the variables along the profile. This was the method used in the present study.

METHOD

Total dissolved inorganic CO_2 in a given sea water sample is, in general, a function of dissolved O_2 , total alkalinity, temperature, and salinity of the same sample. Oxygen is used up during oxidation of organic debris which produces CO_2 . Alkalinity changes may be, in part, a result of carbonate solution or precipitation both of which will alter the TCO_2 in the water. Temperature and salinity may be important since they control the solubility of gases in sea water and may, therefore, serve as a parameter for estimating the initial concentration of the gases in the parent waters. Also, temperature and salinity may serve as an indicator for the extent of the mixing between two water masses.

One would expect, of course, that the functional relationship between TCO_2 and the various parameters will be nonlinear. Such a relationship between TCO_2 , dissolved oxygen (O_2), total alkalinity (TA), temperature (T), salinity (S), and depth (D) can be formally expressed by

$$\text{TCO}_2 = F(\text{O}_2, TA, T, S, D) \quad (3)$$

and the total differential of TCO_2 is related to the differentials of the various variables by

$$d(\text{TCO}_2) = \frac{\partial F}{\partial \text{O}_2} d\text{O}_2 + \frac{\partial F}{\partial TA} dTA + \dots \quad (4)$$

The concentration of TCO_2 along a vertical profile is fairly constant and total variations are no more than 4%-5% of the mean value. Hence, one may be justified in attempting to use the conventional approximation

$$\Delta \text{TCO}_2 \cong \frac{\partial F}{\partial \text{O}_2} \Delta \text{O}_2 + \frac{\partial F}{\partial TA} \Delta TA + \dots \quad (5)$$

If the partial derivatives are linear along the profile, one can convert the last equation to the form

$$\text{TCO}_2 = A + A_0 \cdot \text{O}_2 + A_A \cdot TA + A_T \cdot T + A_S \cdot S + A_D \cdot D \quad (6)$$

$$A = \text{constant}$$

The present study is an attempt to test the linearized model expressed by the last equation. This result was achieved by using the least square method to fit equation 6 to measured data along a vertical profile. The fitting is an extension of the well-known two-dimensional regression analysis to n -dimensional space. The mathematical tools for this procedure are well-established [Draper and Smith, 1966] and are in essence a least square fit of a plane (in n dimensional space) to the given data. Similar to the two-dimensional case, one obtains an estimate of the regression coefficients which are of interest here. The arithmetic is fairly involved and calls for mechanized calculations.

The fitting was accomplished by two computer programs prepared by the biomedical research group at the University of California, in Los Angeles [Dixon, 1970]. Two programs were used, a step-wise regression program (BMD02R) and a case-combination regression program (BMD03R). The first program computes a sequence of multiple linear regression equations in a step-wise manner. At each step, one variable is added to the regression equations so that the contribution of each variable to the regression can be evaluated. The second program performs multiple regression analysis on the data within each selection of subsamples from a given population. This program is convenient for checking statistical trends within a given subsample of a population (e.g., a section of a vertical oceanic profile).

RESULTS

Multiple regression analysis was performed on two profiles in the eastern Pacific by Culbertson and Pytkowicz [1970]. One oceanographic station was taken in the tropical eastern Pacific (station 70; 4° 00.0'S, 82° 00.0'W) and the other in the eastern North Pacific (station 127; 38° 00.2'N, 124° 45.0'W). Both stations were established during *Yaloc 69*, a cruise to the eastern tropical Pacific on the *Yaquina* (Department of Oceanography, Oregon State University). Dissolved oxygen, pH , alkalinity, as well as temperature and salinity were measured along a profile at the two stations. The data consists of a set of 29 points to a maximum depth of 3754 meters (station 70) and a set of 28 points to a maximum depth of 3928 meters (station 127). It is reported that the

precision (2 sigmas) for pH measurements was ± 0.006 pH units, and for total alkalinity, ± 0.008 meq/kg. Oxygen was determined by the method of Strickland and Parsons [1965] who gave an estimated precision (2 sigmas) of ± 3 μg -atm/l. Total dissolved inorganic CO₂ was calculated from total alkalinity and pH by using Lyman's data for the apparent dissociation constants of carbonic and boric acids in sea water. By means of the step-wise regression analysis program, BMD02R [Dixon, 1970], I have calculated linear correlation coefficients and a multivariable regression for the given data.

Table 1 is a linear correlation matrix [Draper and Smith, 1966] for the data of station 70. This two-variable analysis suggests that dissolved CO₂ is highly correlated with temperature ($R = -0.983$) and to a lesser extent, with salinity ($R = -0.843$) and total alkalinity ($R = -0.825$). It is of interest that linear correlation between TCO₂ and O₂ is fairly poor ($R = -0.357$), which again substantiates the need for a multiple regression analysis. If the upper layer points are removed and the correlation matrix recomputed, one finds a somewhat different picture. Table 2 is a recomputed correlation matrix for station 70 for the water column below 300 meters. The correlation between CO₂ and O₂ is now more pronounced and the linear correlation coefficient has changed from a low negative value ($R = -0.357$) to a high positive value, 0.856. Although the procedure of removing the upper layer seems to improve the linear correlation between TCO₂ and O₂, the polarity of the linear regression seems to be incorrect since one would expect TCO₂ to increase with a decrease of O₂. It seems, therefore, that linear correlation of the variables taken two-at-a-time, without considering the other interdependent variables, is of little use.

The next step of the statistical analysis, was to compute a multivariable linear regression when TCO₂ was considered a dependent variable while T , S , O_2 , and D , and TA were considered independent variables. This was a step-wise computation in which variables were added one-at-a-time. The variable added in each step was the one that made the greatest reduction in the error sum of squares [Dixon, 1970; Draper and Smith, 1966]. Table 3 summarizes this computation. The first variable to enter

TABLE 1. Correlation Matrix for Station 70, Eastern Tropical Pacific

	<i>D</i>	<i>T</i>	<i>S</i>	O ₂	<i>TA</i>	TCO ₂
<i>D</i>	1	-0.844	-0.637	0.229	0.948	0.772
<i>T</i>	-0.844	1	0.872	0.194	-0.989	-0.983
<i>S</i>	-0.637	0.872	1	0.164	-0.723	-0.843
O ₂	0.229	0.194	0.164	1	0.223	-0.357
<i>TA</i>	0.948	-0.989	-0.723	0.223	1	-0.825
TCO ₂	0.772	-0.983	-0.843	-0.357	-0.825	1

Parameters defined in text.

regression was *T*, since it produced the highest linear correlation coefficient with TCO₂. The next variable to be added was O₂, which was calculated to be the variable that would make the highest contribution to the regression when compared to the other remaining variables. It is of interest that one could not have deduced it from the correlation matrix of Table 1, since TCO₂-O₂ showed a low level of correlation. Total alkalinity was the following variable to be added. This variable, however, made only a small contribution to the regression as the increase in multiple correlation coefficient squared (*MRSQ*) was only 0.0034. The contributions of *S* and *D* were even smaller and it seems that the addition of these variables is of little significance. The regression computation was repeated for station 127 (northeastern Pacific) and the results, regarding the contribution of the variables to the regression, were identical.

At each step of the regression, the regression coefficients (*A*'s in equation 3) were recalculated for the variables in the regression. The values of the coefficients obtained in each step is summarized in Table 4. This table also demonstrates the fact that the contribution of the last two variables to enter regression, and especially the contribution of *D*, is insignificant. This same

general behavior was found during the stepwise regression computation for station 127.

The standard error of estimation (*s*) for the regression coefficient (Table 4) is generally increasing as more variables are entering the regression. However, the ratio between (*s*) and the calculated coefficient becomes smaller with each step. In general, the standard error of estimation is 10% to 30% of the estimated value which indicates a high level of confidence since the number of degrees of freedom is above 24. The standard error of estimation for *A*₀ and *A*₄, which are of interest here, is only about 5% of the coefficients at steps 4 and 5. This indicates that a statistically meaningful dependence has been established between TCO₂, O₂, and *TA*. The same general behavior was observed during the regression computation for the two stations with and without surface data points.

Figures 1 and 2 are plots of the regression residual versus depth and TCO₂, respectively. The range of the residuals is 0.009 mmole/kg, namely only 0.5% of the mean value of TCO₂ at station 70 (2.274 mmole/kg). It should be remembered that the precision of *TA* determination (2 sigmas) was reported to be 0.008 meq/kg, which is approximately 0.5% of the mean value of this variable at station 70 (2.376

TABLE 2. Linear Correlation Matrix for Intermediate and Deep Waters (the water column below 300 meters) for Station 70, Eastern Tropical Pacific

	<i>D</i>	<i>T</i>	<i>S</i>	O ₂	<i>TA</i>	TCO ₂
<i>D</i>	1	-0.799	-0.114	0.944	0.894	0.715
<i>T</i>	-0.799	1	0.394	-0.924	-0.952	-0.978
<i>S</i>	-0.114	0.394	1	-0.039	-0.099	-0.443
O ₂	0.944	-0.924	-0.039	1	0.981	0.856
<i>TA</i>	0.894	-0.952	-0.099	0.981	1	0.921
TCO ₂	0.715	-0.978	-0.443	0.856	0.921	1

TABLE 3. Summary of Step-wise Multivariable Regression Computation on the Data of Station 70

Step No.	Variable Entered	<i>MR</i>	<i>MRSQ</i>	Increase in <i>MRSQ</i>	<i>N</i>
1	<i>T</i>	0.9834	0.9671	0.9671	1
2	O ₂	0.9978	0.9956	0.0286	2
3	<i>TA</i>	0.9995	0.9990	0.0034	3
4	<i>S</i>	0.9997	0.9994	0.0004	4
5	<i>D</i>	0.9998	0.9995	0.0001	5

MR is multiple correlation coefficient.

MRSQ is *MR* squared.

N is number of variables in regression.

TCO₂ was assumed to be the dependent variable.

meq/kg). Since TCO₂ was calculated from *TA* and *pH*, one would expect that the precision of TCO₂ determination will be the sum of the uncertainties of *TA* and *pH* determination. Hence, the distribution of the residual is no larger (and in effect, much smaller) than one would expect it to be if the scatter was only a result of the precision of the measurement. The plots of the residuals versus depth (Figure 1) and TCO₂ (Figure 2) do not show any apparent trends with respect to any of these variables. Similar scatter was obtained when the residuals were plotted against the other variables. This again indicates that the regression is adequately (statistically speaking) representing the data along the profile of station 70.

The range of the regression residuals for station 127 was found to be somewhat larger than that for station 70. The range here is 0.015 mmole/kg, which is approximately 0.8% of the mean of TCO₂ at station 127. However, examination of the plot of residual versus depth (Figure 3) indicates that the larger range is mainly due to three data points: 49 meters, 75 meters, and 197 meters. If these data points are removed, one obtains a better fit and the range of residuals drops to approximately the same value that was calculated for the regression of station 70.

I have combined the data of stations 70 and 127 (57 data points) and recalculated the regression coefficients. A summary of the step-wise multiple regression on the combined data is given in Table 5. The multiple correlation coefficient (*MR*) reached the value of 0.9987 when the three variables *TA*, O₂, and *T* were included in the regression, whereas the addition of *S* and *D* did not improve this parameter.

Also, the inclusion of these parameters in the regression changed only slightly the computed values for *A_o* and *A_A* as can be seen in Table 6. This table compares the coefficients *A_o* and *A_A* that have been obtained during the various phases of the computations. It should be noted that the oxygen regression coefficient *A_o* remained in the range -0.3 to -0.4 when the regression was calculated for stations 70, 127, and for the combined data of the two stations. The total alkalinity regression coefficient *A_A* showed larger variations for the two stations. Possible interpretation for these results will be discussed below.

The last regression computation to be considered here is a step-wise regression on the upper and lower water layer of stations 70 and 127. In this calculation, I have removed all the data points of the intermediate water columns (200-3000 meters) and repeated the calculation on the remaining data points of the profiles. Total dissolved CO₂ was again assumed to be the dependent variable while *T*, *S*, O₂, and *TA* were considered the independent variables. The regression computation was done in the step-wise manner (program BMDO2R) when at each step an additional variable was added to the regression and the regression coefficients were recalculated. By means of the regression coefficients and the data of the intermediate water column, I have calculated the predicted TCO₂ concentration for the region that had not been included in the regression computation. The residuals (measured minus predicted TCO₂) were calculated and plotted versus depth for station 70 and 127 (Figures 4 and 5, respectively). Each plot depicts the residuals for four regressions. In the first regression only tem-

TABLE 4. Regression Coefficient at Each Step of a Step-wise Regression Computation for Station 70

Step No.	A_T		A_O		A_A		A_S		A_D	
	Calculated	s	Calculated	s	Calculated	s	Calculated	s	Calculated	s
1	-0.1546	0.0006								
2	-0.1493	0.0002	-0.1846	0.0142						
3	-0.00965	0.0006	-0.3366	0.0181	0.6701	0.0755				
4	-0.0067	0.0008	-0.3974	0.0198	0.9382	0.0843	-0.0408	0.0095		
5	-0.0060	0.0008	-0.4112	0.0193	0.9310	0.0781	-0.0495	0.0096	0*	0*

Asterisk * indicates less than 10^{-5} . s is standard error of estimate of A .

perature was entered. In the second, both temperature and oxygen were entered. In the third, T , O_2 , and total alkalinity were included while in the fourth one, T , O_2 , TA , and salinity entered the regression. Figures 4 and 5 suggest that a high degree of certainty in the prediction can be obtained if the four independent variables are included in the regression. The prediction with only T included is poor while the prediction with T and O_2 included is fair, since the largest deviation of the predicted concentration from the observed one is only approximately 1%. In station 127, we observe a special case in which the prediction of the regression which includes only T and O_2 , is better than a more elaborate regression in which T , O_2 , and TA are included. The addition of S , however, seems to improve the prediction.

DISCUSSION

The regression analyses described above suggest that the TCO_2 along a profile can be accurately represented by a linear multiple regression equation which includes the variables: temperature, dissolved oxygen, total alkalinity, and salinity. The contribution of salinity was found to be less pronounced as the improvement in the multiple correlation coefficient (MR) due to this variable was found to be small. The extremely high degree of fitness that has been obtained may suggest that the concentrations of TCO_2 along an oceanic profile is indeed a linear function of temperature, salinity, oxygen, and total alkalinity. An attempt will be made now to present a mixing model that may explain the linear relationship discussed above. Three models will be considered: a two-source model, a three-source model, and a multiple-source model. It will be shown that these three models may result in a linear function between TCO_2 and the other variables.

The models to be discussed here follow the treatments of *Wyrtki* [1962], *Munk* [1966], and *Craig* [1969] in that I assume that the concentration of traces along an oceanic vertical profile can be explained by a vertical mixing process. However, the treatment presented here is not concerned with the mechanism of mixing, but rather with its end result, namely, the concentration of the tracers in the mixture. To put it more bluntly, the end result of a mixing process described by the eddy diffusion concept,

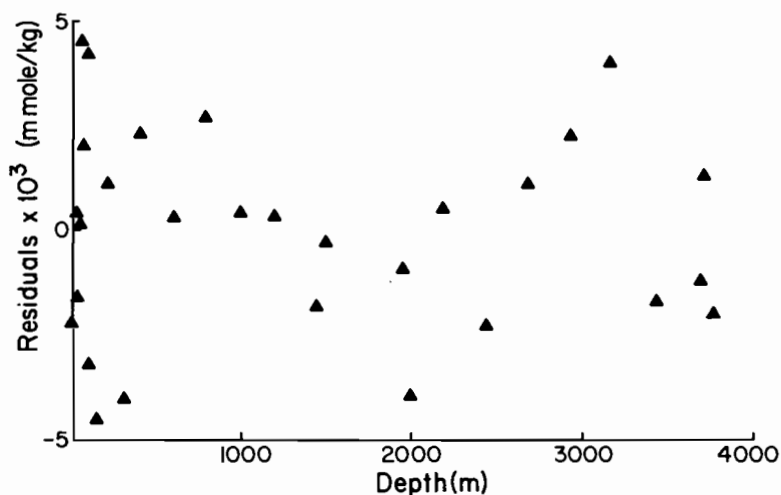


Fig. 1. Multiple regression residuals versus depth for station 70.

or the simple-minded process of mixing two buckets of waters, is the same. The present treatment is clearly incomplete because it is only concerned with the final concentrations, it does not describe the driving forces of the mixing, nor does it put any time scale on the process. It appears, however, that this simplified approach may be useful in studying the interrelations between the variables since one can obtain an estimate of the partial derivatives of equation 4. These partial derivatives (or constants in equation 6) express the functional relation between TCO₂ and each of the param-

eters (O₂, T, . . .), given that all the other variables are constant. It should be emphasized that an agreement between the suggested models and the regression analysis does not necessarily imply that the models accurately represent real oceanic processes. The only conclusion that can be drawn is that the models are not inconsistent with observed data [Munk, 1966].

Although the models to be presented here do not depend on any specific mixing mechanism, they are not in conflict with the generally accepted eddy-diffusion mixing model. This is

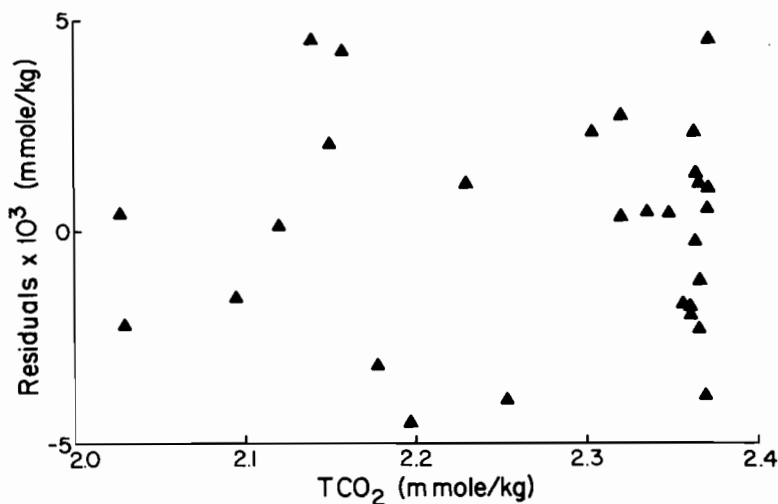


Fig. 2. Multiple regression residuals versus TCO₂ for station 70.

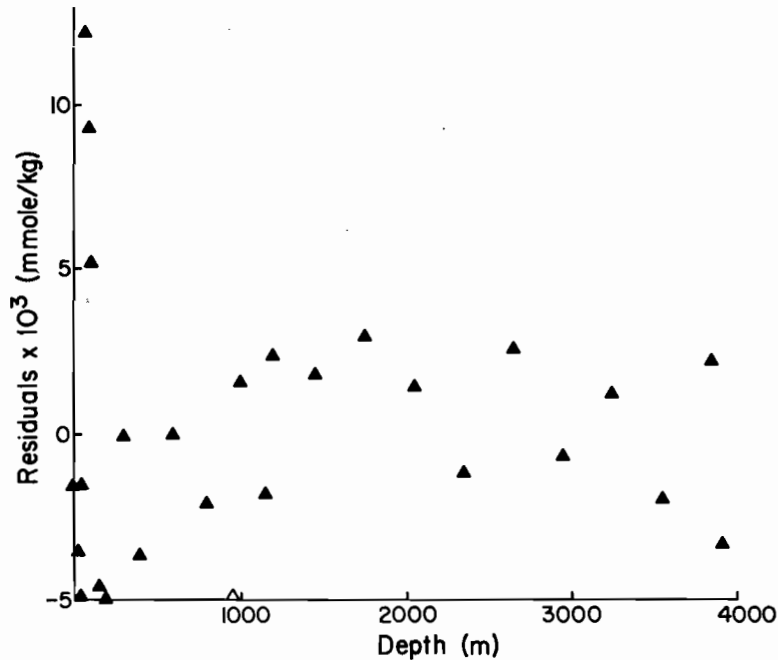


Fig. 3. Multiple regression residuals versus depth for station 127.

clearly demonstrated in the appendix to this paper in which the regression equation is derived directly from the eddy-diffusion mixing model.

Model A: Two sources. Consider two water types M_1 and M_2 having temperatures T_1 and T_2 and concentrations of some stable conservative property (e.g., argon) C_1 and C_2 , respectively. The two water types can be represented as two points on a CT plane (Figure 6). The line between the two points is the locus of water masses that will result when the two sources are mixed in different proportions. I am neglecting

here second order effects and assuming that the temperature T and concentration C of a mixture is a linear combination of T and C of the sources. Pressure effects are also neglected and it is assumed that the mixing is a simple mixing (laminar or turbulent) with no diffusive parameters. Namely, mixing is not a function of concentration gradients. The CT diagram considered here is, of course, an extension of the classical ST diagram originally suggested by *Helland-Hansen* [1916].

The concentration C of a given mixture is a function of the proportions of the two sources

TABLE 5. Summary of Step-wise Multivariable Regression on the Combined Data of Stations 70 and 127

Step No.	Variable Entered	MR	$MRSQ$	Increase in $MRSQ$	N
1	TA	0.8690	0.7552	0.7552	1
2	O_2	0.9549	0.9119	0.1567	2
3	T	0.9987	0.9974	0.0854	3
4	S	0.9987	0.9974	0	4
5	D	0.9987	0.9974	0	5

TCO_2 was assumed to be the dependent variable.

MR = multiple correlation coefficient.

$MRSQ$ = MR squared.

N = number of variables in regression.

TABLE 6. Oxygen and Alkalinity Regression Coefficients for Stations 70, 127, and the Combined Data of the Two Stations

Data	Parameters in Regression	A _o		A _A		A _T	
		Computed value	s	Computed value	s	Computed value	s
Station 70	T, O, TA	-0.33659	0.01808	0.67015	0.07357	-0.00965	0.00059
	T, O, TA, S, D	-0.41120	0.01937	0.93096	0.07808	-0.00602	0.00082
Station 127	T, O, TA	-0.34729	-0.01351	0.18573	0.12106	-0.02288	0.00318
	T, O, TA, S, D	-0.33011	-0.03343	0.33834	0.11909	-0.01473	0.00442
Station 70 + 127	T, O, TA	-0.37386	0.00717	0.69919	0.01846	-0.00960	0.00023
	T, O, TA, S, D	-0.36911	0.02166	0.64632	0.06739	-0.00996	0.00091

s is standard error of estimate of A.

in the mixture. Since a linear mixing process is assumed, one can express the concentration C of a given mixture as a linear function of T, the temperature of the same mixture

$$C = \left(\frac{C_2 - C_1}{T_2 - T_1} \right) T + C_1 \quad (7)$$

or, in general

$$C = K_T T + C_0 \quad (8)$$

The stable conservative variable T is used here as a measure of the degree of mixing. Any other stable conservative parameter could have

been used since we are assuming simple linear mixing. Another stable conservative variable that can be used is, of course, salinity. However, salinity variations along a vertical profile in the ocean might be small as compared to temperature variations. Hence, temperature can be considered a more sensitive measure for mixing.

Since we are assuming a linear advective mixing, we can use the principle of superposition for calculating the concentration of a nonconservative variable. We will distinguish here between the preformed concentration C' and the real (measured) concentration C, which is different

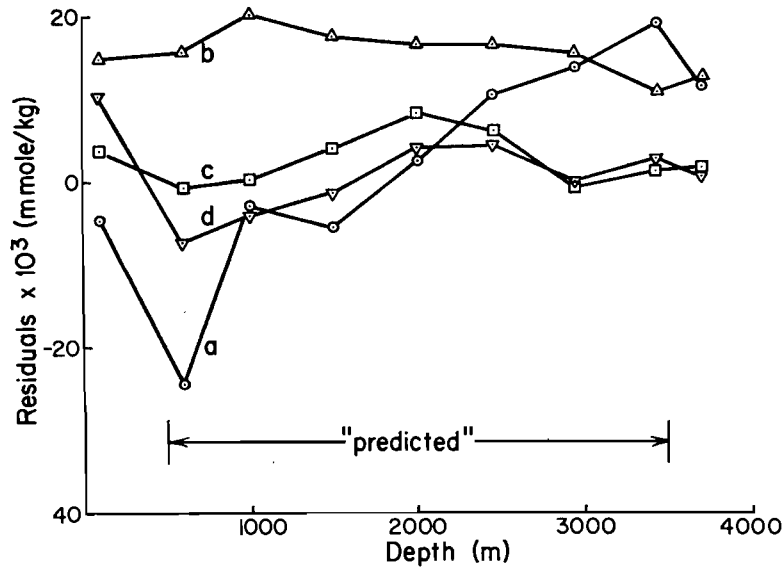


Fig. 4. Residuals of TCO₂ prediction in intermediate waters when only surface and bottom data points were used in regression calculation for station 70. The plots are for regressions in which the independent variables are (a) T; (b) T, O₂; (c) T, O₂, TA; (d) T, O₂, TA, S.

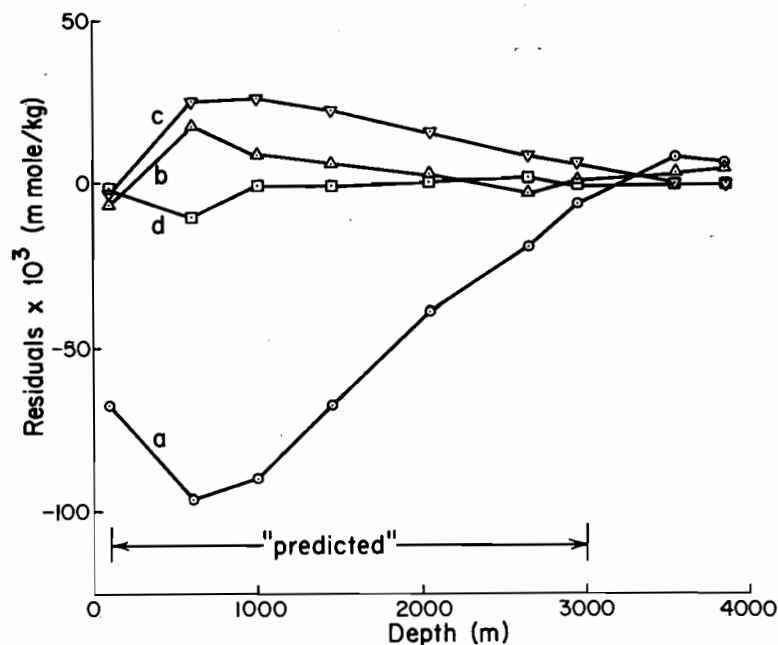


Fig. 5. Residuals of TCO_2 prediction in intermediate waters when only surface and bottom data points were used in regression calculation for station 127. The plots are for regression in which the independent variables are (a) T ; (b) T, O_2 ; (c) T, O_2, TA ; (d) $T, \text{O}_2, \text{TA}, S$.

from C' due to production or consumption of this variable (ΔC).

$$\Delta C = C - C' \quad (9)$$

I shall further assume that ΔC is linearly proportional to ΔG

$$\Delta C = A_G \cdot \Delta G \quad (10)$$

Where G is another nonconservative variable and A_G is the exchange constant. Of course, ΔC

may be a function of more than one variable, in which case, equation 10 should be extended to include the additional dependencies.

Following the above procedure, we can distinguish between TCO_2 , the measured quantity, and TCO_2' , the preformed concentration that would have been found if the dissolved inorganic CO_2 would have been a conservative quantity. In reality, CO_2 is not a conservative variable, since oxidation of organic debris, respiration, and solution or precipitation of carbonates will change the concentration of CO_2 in a given water mass. Since oxidation of organic matter (both biological and nonbiological) as well as carbonate reactions are probably the most important processes we can express ΔTCO_2 (the difference between the observed and preformed CO_2) as

$$\Delta \text{TCO}_2 = A_O \cdot \Delta \text{O}_2 + A_A \cdot \Delta \text{TA} \quad (11)$$

where A_O is the exchange constant between CO_2 and O_2 and A_A is the exchange constant between CO_2 and titration alkalinity. The preformed concentrations of O_2 and TA can be expressed by an equation similar to (8), and by introducing the relation of equation 9, one

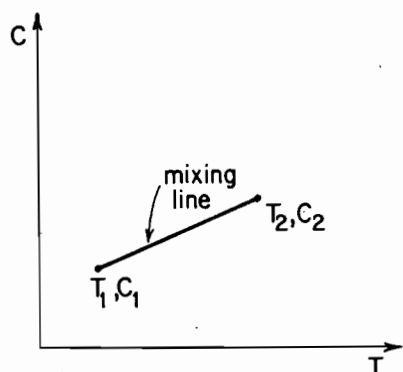


Fig. 6. CT diagram for a linear, two-source, mixing model.

obtains

$$\begin{aligned} \text{TCO}_2 = & K_{\text{CO}_2} \cdot T + \text{TCO}_{20} + A_0 \{ O_2 \\ & - (K_O \cdot T + O_{20}) \} + A_A \{ TA \\ & - (K_A T + TA_0) \} \end{aligned} \quad (12)$$

which can be reduced to the form

$$\text{TCO}_2 = A + A_T \cdot T + A_O \cdot O_2 + A_A \cdot TA \quad (13)$$

Hence, the simple two-source model suggests that TCO₂ of a given water mass, which is a mixture of two water types, should be a linear function of T, O₂, and TA.

Model B: Three sources. In this case, one has to consider two possibilities regarding the concentrations C, of the three sources. One case is that all the three water types are on the same straight line in the CT diagram (Figure 6). In this event, one can follow the analysis of model A, which will result in an equation similar to equation 13. In the general case, however, the three water types would not fall on the same line. In such a case, one can introduce an additional stable conservative parameter, such as salinity, to force a linear relationship for a mixture of the three sources.

Consider three water types M₁, M₂, and M₃, with temperature T₁, T₂, T₃, salinities S₁, S₂, S₃, and concentrations of some quantity C, C₁, C₂, C₃. The three water types can be represented by three points, C₁T₁S₁, C₂T₂S₂, and C₃T₃S₃ in a three-dimensional space, CTS (Figure 7). The three points define a plane which can be mathematically expressed as

$$C = K_{CT} \cdot T + K_{CS} \cdot S + C_0 \quad (14)$$

The last equation defines a plane which is the locus of all possible mixtures of the three sources. This equation can be used, therefore, to calculate the concentration of any mixture provided that the temperature and salinity of the mixture are known.

Equation 14 is similar to equation 8 except that it includes salinity S as an additional variable. Therefore, by following the analysis presented above one will obtain here also a linear relation for a nonconservative quantity C. Hence, the concentration of TCO₂ in a water mass that is a mixture of three primary sources can be presented by

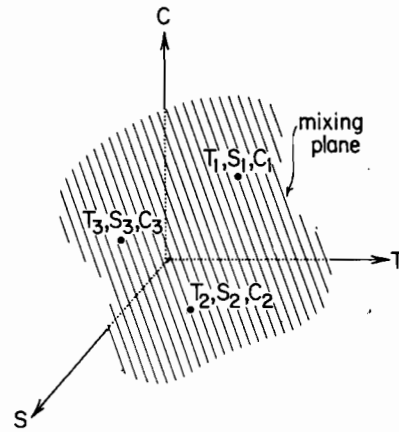


Fig. 7. CTS space for a linear, three-source, mixing model.

$$\begin{aligned} \text{TCO}_2 = & A + A_O \cdot O_2 + A_A \cdot TA \\ & + A_S \cdot S + A_T \cdot T \end{aligned} \quad (15)$$

Model C: Multiple sources. Here again one has to distinguish between two cases regarding the locus of the points which are representing the primary water types. If the points fall on the same plane in the CTS space, one can express the concentration of any mixture by a linear equation similar to equation 15. If the sources do not define a single plane, one will have to introduce more conservative variables until a linear relationship is established. This has the effect of adding more dimensions to the space and the final expression for C (equations 14 or 15) will have to include more stable conservative variables.

Addition of more independent conservative variables is not required if all the sources fall on the same plane in the CTS space. Or, in our particular case, if each of the preformed quantities: O₂', TA', and TCO₂' define a unique plane in the corresponding spaces, one will be able to express TCO₂ of a water mass by an equation similar to equation 15. In the ocean, this requirement is partially fulfilled since the solubility of gases are, to a first approximation, linear with temperature and salinity and surface sea water is approximately in equilibrium with the atmosphere. Hence, to a first approximation, the preformed quantities, O₂', TCO₂', TA' define a unique plane in their CTS spaces. One would expect, therefore, to find in the oceans a linear relation similar to equation 15.

The multiple regression analysis given above suggests that the vertical distribution of TCO_2 along a vertical profile in the eastern Pacific can be expressed by a linear expression similar to equation 13 which includes T , O_2 , and TA as independent variables. This implies that the TCO_2 concentration along a vertical profile can be explained by a simple, two-source mixing model (model A). However, a linear dependence similar to equation 13 should hold also when temperature is replaced by another stable conservative tracer. One would expect, therefore, that a good fit would be obtained when T is replaced by S in the regression. This, however, has not been observed during the present analyses. Calculations with S , O_2 , and TA as independent variables produced, in general, regressions with relatively poor statistics when compared to the regressions in which T was incorporated. This is not surprising as the S - T diagram of both stations 70 and 127 are nonlinear. The diagram of station 70 has two break points (at approximately 6°C and 16°C) while the one for station 127 has some nonlinearity at the upper layer.

The apparent discrepancy between our results, which suggest that eastern Pacific is not inconsistent with a linear mixing model type A, and the nonlinearity of the S - T diagram, can be resolved by considering the preformed concentrations of TCO_2 , O_2 , and TA in the water sources. The preformed values are defined as the concentration that would be measured if the variables were stable and conservative. The preformed concentrations are, therefore, a function of the solubility of the gases in sea water which is strongly dependent on temperature and to a lesser degree on salinity. Hence, the variable T in equation 13 can serve not only as a measure of the mixing that has taken place, but also reflects the temperature dependence of the preformed concentrations. This point can be made more lucid by reconsidering the C - T diagram of Figure 6. The fact that we have obtained a good model A fit suggests that the preformed concentrations of TCO_2 , O_2 , and TA , along the profiles, lie on a straight line on their respective C - T diagrams. This picture is consistent with a two-source mixing model (type A) although the oceanic profiles may involve more than two physical sources as suggested by the S - T diagrams. Clearly, the fact that the pre-

formed concentrations of the primary sources seem to lie on a straight line, on their respective C - T diagrams, does not affect the estimation A_0 and A_A which reflect the TCO_2 - O_2 and TCO_2 - TA exchange, respectively.

The exchange constant A_0 was found to be bounded by the values -0.33 to -0.42 during all the regression analyses. This may indicate that the ratio between oxygen consumption to TCO_2 production is fairly constant in the eastern Pacific. The parameter A_0 is essentially the ratio $\Delta\text{TCO}_2/\Delta\text{O}_2$ (mole/gram-atom) when all the other variables (TA , S , T) are constant. This coefficient expresses the real TCO_2 - O_2 relation and is independent of processes such as the solubility of the gases in sea water. The best estimate (lowest s) of A_0 was obtained when the regression computation was carried out on the data of the two stations (Table 6). The standard error of estimate s was only 0.00727 while the number of degrees of freedom was 53 (57 data points minus 4 variables in regression). Hence, one may conclude, with a 95% confidence level, that A_0 is in the range -0.374 ± 0.014 . It is of interest that the value obtained here is very close to the value predicted by the model of Redfield *et al.* [1963]. They have predicted the ratio to be $-106/276$ or in decimal units, -0.384 . This prediction is based on the composition of average plankton analyses given by Fleming [1940].

The TCO_2 - TA exchange coefficient A_A was found to be more variable than A_0 . Also, the standard error of estimate (s), was found to be larger (Table 4). This may indicate that carbonate solution or precipitation are not the only important processes which cause TA variations in the oceans. If only carbonate reaction was responsible for TA variation, A_A would have been less variant and its values would have converged around 0.5. This was not observed during the present study. The alkalinity coefficient A_A was consistently calculated to be lower than 0.5 at station 127, and higher than 0.5 at station 70. This may indicate that at least two additional processes are responsible for TA variations; one has the effect of increasing TA while the other has the effect of decreasing it.

Total dissolved inorganic CO_2 was considered in the present regression analysis to be the dependent variable whereas dissolved oxygen

was taken as an independent variable. This choice, although arbitrary from a mathematical point of view, was considered to best express real oceanic processes. Concentration of TCO₂ at any given sea-water sample is a function of the concentrations of the 'preformed' TCO₂, variation due to respiration, nonbiological oxidation of organic matter and possible variations due to carbonate reactions. The concentration of O₂, although linearly related to TCO₂ through equation 1, is not a function of TCO₂. The O₂ concentration depends on the rate of formation of the 'preformed' O₂ by solution of atmospheric oxygen and photosynthesis, and removal of O₂ by respiration and oxidation of organic debris. The latter, which is probably the most important process, is independent of TCO₂ at least to a first approximation.

The present analysis is different from previous related studies in three respects: (1) no attempt was made to assume a priori relations between variables [Culberson and Pytkowicz, 1970], (2) no assumption was made regarding the rate of oxygen consumption [Wyrski, 1962], and (3) a simple linear mixing model is proposed for explaining the vertical distribution of TCO₂ and O₂ along a profile. The fact that the vertical distribution of these parameters can be explained without resorting to eddy diffusion, does not necessarily imply that diffusion processes are unimportant. The present study is only an attempt to look at the mixing problem from a different angle.

The small TCO₂ residual ranges (e.g., Figures 1 and 3) obtained during the present regression analyses may suggest that the various parameters were measured with high precision. Our results indicate that the precision estimates given by Culberson and Pytkowicz [1970] are very conservative and that the real precision of measurement was, in effect, better. It should be pointed out that a systematic error is less important here since a constant shift in the data should be absorbed by the constant *A* (equation 13) while a constant proportional inaccuracy should result in a corresponding error in the corresponding regression coefficient. However, a particular regression analysis may not follow the ideal behavior due to the differences in the statistics of the variables and due to the fact that the regression equation is only an idealized approximation of the actual oceanic

process. The effect of constant error on the present calculation was investigated by perturbing the variables one at a time and recalculating the regression coefficients. The results of these calculations are given in Table 7. The calculations were carried out with an IBM 360/91 digital computer and some of the variations should be attributed to round-off error. Table 7 demonstrates that practical systematic errors (such as errors in concentrations of titration reagents, systematic volumetric and gravimetric errors, and uncertainty in thermodynamic constants) should result in relatively minor errors in the calculated regression coefficients. It appears that the estimated 95% confidence interval is wide enough to include possible contribution of practical systematic errors.

CONCLUSIONS

1. The distribution of O₂, TCO₂, and *TA* along a vertical profile in the eastern tropical Pacific and northeastern Pacific are not inconsistent with a linear mixing model type A.
2. The ratio of TCO₂ production to oxygen consumption was calculated to be in the range -0.374 ± 0.014 (95% confidence level) which is close to the predicted value of -0.384 given by Redfield *et al.* [1963].
3. The present study suggests that O₂ is consumed in situ if a vertical mixing model is pre-assumed.
4. Carbonate solutions or precipitations may not be the only processes that result in *TA* variations in the oceans.

The present study demonstrates that multiple regression analyses can be useful in studying functional relationships between variables in the oceans. This method might be useful in resolving some pending problems such as the questions whether the Redfield *et al.* [1963] model can indeed characterize the *C:N:P* ratio everywhere in the oceans and whether oxygen is consumed in situ in deep waters as suggested by the present preliminary study.

APPENDIX

DERIVATION OF MODEL A EQUATION FROM AN EDDY DIFFUSION MIXING MODEL

The continuity equation for a stable nonconservative tracer (*C*) along a vertical profile in

TABLE 7. Errors in the Regression Coefficients as a Function of a Hypothetical Systematic Error in the Variables

Variable	'Error'	ΔA_T , %	ΔA_S , %	ΔA_O , %	ΔA_A , %
T	0.1 C	0.00	-0.08	-0.01	+0.02
T	1%	+0.5	+1.3	+0.2	-0.4
S	0.1 gm/kg	+19.9	-74.1	-8.5	+16.3
S	1%	-8.3	+31.4	+3.4	+6.7
O_2	5 μ g-atm/l	-0.39	+12.1	-0.4	+0.4
O_2	1%	+0.2	-1.5	+0.9	+0.1
TA	0.02 meq/kg	+0.5	+11.7	-0.7	+1.4
TA	1%	+16.0	-47.0	-7.3	+12.9
TCO_2	0.02 mmole/kg	-4.7	+27.9	+1.6	-3.2
TCO_2	1%	-1.6	+6.7	-1.0	+1.0

the ocean can be expressed by [Wyrski, 1962; Munk, 1966; Craig, 1969]

$$\frac{\partial C}{\partial t} = K \frac{\partial^2 C}{\partial X^2} - w \frac{\partial C}{\partial X} + J \quad (1A)$$

where X is the vertical distance, K is the eddy diffusion coefficient, w is vertical advection, and J the production rate of the tracer.

For a steady-state condition

$$\frac{\partial C}{\partial t} = 0 \quad (2A)$$

and equation 1A can be rewritten in the form

$$KC'' - wC' + J = 0 \quad (3A)$$

where the prime and double-prime stand for first and second derivative with respect to X , respectively.

Equation 3A can now be used to express the vertical distribution of TCO_2 , O_2 , and TA :

$$K(TCO_2)'' - w(TCO_2)' + J_{CO_2} = 0 \quad (4A)$$

$$K(O_2)'' - w(O_2)' + J_{O_2} = 0 \quad (5A)$$

$$K(TA)'' - w(TA)' + J_{TA} = 0 \quad (6A)$$

However, if the production of TCO_2 , along the mixing distance, is controlled only by oxidation of organic matter and carbonate reaction one obtains

$$J_{CO_2} = A_O \cdot J_{O_2} + A_A \cdot J_{TA} \quad (7A)$$

where A_O and A_A are the exchange coefficients of TCO_2 - O_2 and TCO_2 - TA . Note that A_O is expected to be a negative number while A_A should be a positive number. This notation is in accor-

dance with previously defined parameters. If A_A and A_O can be assumed constant along the mixing zone, equations 4A, 5A, 6A, and 7A can be combined to obtain

$$K(\Sigma)'' - w(\Sigma)' = 0 \quad (8A)$$

where

$$\Sigma = TCO_2 - A_O \cdot O_2 - A_A \cdot TA \quad (9A)$$

It is evident from the last equation that Σ behaves as a stable conservative tracer. One can now add another stable conservative tracer, say T , and obtain the following relation for a vertical mixing between two boundaries [Munk, 1966; Craig, 1969]

$$\frac{\Sigma - \Sigma_0}{\Sigma_m - \Sigma_0} = \frac{T - T_0}{T_m - T_0} \quad (10A)$$

where the subscripts 0 and m indicate concentrations at the boundaries. By simple manipulation of the last equation one obtains

$$TCO_2 = A + A_T \cdot T + A_O \cdot O_2 + A_A \cdot TA \quad (11A)$$

where

$$A = \frac{\Sigma_0 - \Sigma_m}{T_m - T_0} \cdot T_0 + \Sigma_0 \quad (12A)$$

$$A_T = \frac{\Sigma_m - \Sigma_0}{T_m - T_0} \cdot T \quad (13A)$$

Equation 11A is identical to the mixing equation derived for model A.

Acknowledgments. I wish to thank Professors I. R. Kaplan and H. Craig for critically reading the manuscript. This study was supported by an AEC contract AT(04-3)-34 PA 178.

REFERENCES

- Craig, H., Abyssal carbon and radio carbon in the Pacific, *J. Geophys. Res.*, *74*, 5491-5506, 1969.
- Craig, H., and R. F. Weiss, Argon concentrations in the oceans: A discussion, *Earth Planet. Sci. Lett.*, *5*, 175-183, 1968.
- Culberson, C., and R. M. Pytkowicz, Oxygen-total carbon dioxide correlation in the Eastern Pacific Ocean, *J. Oceanogr. Soc. Japan*, *26*, 25-30, 1970.
- Dixon, W. J. (Ed.), *Biomedical Computer Programs*, 2nd edit., 600 pp., University of California Press, Berkeley, Los Angeles, 1970.
- Draper, N. R., and S. Smith, *Applied Regression Analysis*, 407 pp., John Wiley, New York, 1966.
- Fleming, R. H., The composition of plankton and units for reporting population and production, *Proc. Sixth Pacific Sci. Congress California, 1939*, *3*, 535-540, 1940.
- Garrels, R. M., Silica: Role in the buffering of natural waters, *Science*, *148*, 69, 1965.
- Gripenberg, S., On the alkalinity of Baltic waters, *J. Cons. Perma. Int. Explor. Mer.*, *26*, 5-20, 1960.
- Helland-Hansen, B., Nogen Hydrograhiske Metoder, *Forh. Skand. Naturf. Mote*, *16*, 357-359, 1916.
- Munk, W. H., Abyssal recipes, *Deep Sea Res.*, *13*, 707-730, 1966.
- Park, K., Total carbon dioxide in seawater, *J. Oceanogr. Soc. Japan*, *21*, 54-59, 1965.
- Redfield, A. C., The exchange of oxygen across the sea surface, *J. Mar. Res.*, *7*, 347-361, 1948.
- Redfield, A. C., B. H. Ketchum, and F. A. Richards, The influence of organisms on the composition of seawater, in *The Sea*, vol. 2, pp. 26-77, edited by M. N. Hill, Interscience, New York, 1963.
- Strickland, J. D. H., and T. R. Parsons, A manual of seawater analysis, *Bull. Fish. Res. Bd. Can.*, *125*, 1965.
- Weiss, R. F., The solubility of nitrogen, oxygen, and argon in water and seawater, *Deep Sea Res.*, *17*, 721-735, 1970.
- Wyrтки, K., The oxygen minima in relation to ocean circulation, *Deep Sea Res.*, *9*, 11-23, 1962.

(Received May 26, 1971;
revised July 23, 1971.)

