Tellus (1978), *30*, 350–357

The atmospheric carbon dioxide problem

By ERIK ERIKSSON, *Division of Hydrology, Uppsala University, P.O. Box 554, S-75122 Uppsala, Sweden*

(Manuscript received June 7, 1977; in final form January 11, 1978)

ABSTRACT

A linear systems analysis of the Mauna Loa carbon dioxide data is presented using estimated fossil carbon dioxide production as input and observed yearly means of carbon dioxide in the atmosphere as output. The analysis is done on discrete time basis with 1 year as a time step. Because of the near exponential increase in fossil fuel production, not much information is possible to extract from the data concerning system parameters except that at present, 50–60% of the atmosphere "excess" carbon dioxide is transferred to the earth's surface yearly (the oceans and land surfaces). However, the results indicate that the pre-industrial equilibrium level of atmospheric carbon dioxide may have increased to a new "equilibrium" level about 10 per mille higher. This can, however, be due to the slow exchange rate between deep sea water and the more superficial parts, the parameter of which is not possible to deduce from the available set of input–output data.

1. Introduction

The background of this problem is familiar to most readers, hence, it can be summarized very briefly. Carbon dioxide is being released in ever increasing amounts to the atmosphere mainly through combustion of fossil fuels. This causes a measurable increase in the partial pressure of atmospheric carbon dioxide. Since carbon dioxide has absorption bands in the infrared region, changing atmospheric concentrations may require changes in the earth's radiation balance mechanism presumably through increased surface temperature. The actual effect on the surface temperature is not known so that speculation is free to paint any picture of future climatic changes due to release of fossil carbon dioxide. The problem is further complicated, however, by the lack of knowledge about the mechanisms which account for the portion of emitted carbon dioxide being absorbed in the oceans or incorporated into the terrestrial biosphere.

Although some information is available on the oceanic properties pertaining to the problem, uncertainties are still considerable. As to the terrestrial biosphere, practically nothing is known about its dynamic behaviour.

There are several thorough reviews of this problem (e.g. Bolin, 1970, 1975) and a short version is found in a recent SCOPE report (1976).

Attempts have also been made to develop mathematical models which account for the atmospheric increase given the fossil output and known (and assumed) properties of oceans and the terrestrial biosphere. Such models have been presented by Machta (1973) and by Bacastow and Keeling (1973) and others, admirably reviewed by Bolin (1975). In these models natural radiocarbon information is used to derive some parameter values. Such parameters may be useful provided their extraction from data is based on physically correct models of the dynamics of the system.

The purpose of the present paper is to investigate what kind of information can be derived from observed data on atmospheric concentrations of carbon dioxide and data on fossil fuel release. As a base the observation results presented in the SCOPE report referred to earlier will be used (p. 92). The fossil output data are also taken from the same report, p. 74. The approach here is therefore different from before, insofar that no auxillary information is used. This may appear wasteful but can still be of some guidance for future work.

2. The data

These are listed in Table 1 as average concentrations of carbon dioxide at Mauna Loa, Hawaii, taken as differences from the 1958 average level 316 ppm and fossil carbon dioxide output expressed in 10^9 tons per year. The data are complete up to and including 1974.

Table 1. *Yearly averages of carbon dioxide concentrations in Mauna Loa and fossil carbon dioxide production*

Concentrations given as differences in ppm from the 1958 average, 316 ppm. Yearly fossil carbon dioxide output in 10^9 metric ton units

Year	CO_2 ppm	Prod of CO_2 10^9 tons
1958	0.00	8.5
1959	0.61	9.0
1960	1.50	9.7
1961	2.18	9.6
1962	3.04	10.1
1963	3.50	10.7
1964	4.14	11.3
1965	4.72	11.8
1966	5.40	12.3
1967	6.18	12.6
1968	7.05	13.4
1969	8.14	14.3
1970	9.45	15.2
1971	10.49	16.0
1972	11.64	16.5
1973	13.83	17.6
1974	14.32	18.0

3. Analytic method

Considering fossil carbon dioxide release as an input to the atmosphere, some very simple box models can be derived for a systems analysis approach. In Fig. 1 the simplest possible system is shown in which the box represents the atmosphere receiving an input $x(t)$. Out of the box there is a flux, $u(t)$ going into the earth's surface with no return flow. The oceans and the terrestrial biosphere act as an infinitely large sink. For the atmospheric box with mass $M(t)$ we impose the obvious condition

Fig. 1. The single box model of atmospheric carbon dioxide.

$$\frac{dM(t)}{dt} = x(t) - u(t) \tag{1}$$

Assuming a linear, first-order response of the atmospheric box we can state that

$$u(t) = k \cdot M(t) \tag{2}$$

so we can also write

$$\frac{du}{dt} = k(x(t) - u(t)) \tag{3}$$

It is, however, convenient to transform this to an expression valid for discrete times. Denoting the corresponding data by $x_1, x_2, \ldots, x_t$ and $u_1, u_2, \ldots, u_t$ respectively, we approximate our expression as follows. u_t is the average output during year t and x_t can be said to be the average input during the same year. The difference $u_t - u_{t-1}$ is then an estimate of the average rate of change of output in the time interval $t - 1$ to t, i.e. expresses the time derivative. As to $x(t)$ and $u(t)$ we should consequently approximate these by $(x_t + x_{t-1})/2$ and $(u_t + u_{t-1})/2$ respectively. Hence, eq. (3) will read

$$u_t - u_{t-1} = \frac{k}{2}(x_t + x_{t-1} - u_t - u_{t-1}) \tag{4}$$

which can be rearranged to

$$u_t = \frac{2-k}{2+k} u_{t-1} + \frac{k}{2+k}(x_t + x_{t-1}) \tag{5}$$

Writing

$$\alpha = \frac{2-k}{2+k} \tag{6}$$

we can write

$$u_t = \alpha u_{t-1} + \frac{1-\alpha}{2}(x_t + x_{t-1}) \tag{7}$$

Because of the first-order linearity condition imposed the atmospheric mean concentration which we call y_t is simply related to u_t by, say,

$$y_t = \beta u_t \tag{8}$$

which gives

$$y_t = \alpha y_{t-1} + \beta \frac{1-\alpha}{2}(x_t + x_{t-1}) \tag{9}$$

Noting now that the z-transform can be defined by

$$X(z) = \sum_{t=0}^{\infty} x_t z^{-t} \qquad x_t = 0 \text{ for } t = -1, -2, \ldots \quad (10)$$

where z is bounded such that the sum converges, we find that

$$z^{-k}X(z) = \sum_{t=0}^{\infty} x_{t-k} z^{-t} \quad (11)$$

Hence, the z-transform of eq. (7) will be

$$U(z) = \alpha z^{-1}U(z) + \frac{1-\alpha}{2}(1 + z^{-1})X(z)$$

or

$$U(z) = H(z) \cdot X(z) \quad (12)$$

where the transfer function $H(z)$ is

$$H(z) = \frac{(1-\alpha)(1+z^{-1})}{2(1-\alpha z^{-1})} \quad (13)$$

for a first-order linear system.

We can also use a somewhat simpler approximation, i.e. by writing

$$u_t - u_{t-1} = kx_t - ku_t \quad (14)$$

leading to

$$u_t = \frac{1}{1+k}u_{t-1} + \frac{k}{1+k}x_t \quad (15)$$

If $\alpha' = (1/1 + k)$ the z-transform becomes

$$U(z) = \frac{1-\alpha'}{1-\alpha' z^{-1}}X(z) \quad (16)$$

i.e. with a transfer function

$$H'(z) = \frac{1-\alpha'}{1-\alpha' z^{-1}} \quad (17)$$

We can apply this to a more complex system of two linear boxes connected as in Fig. 2. Here we consider the possibility of a flux back from the earth's surface, linearly related to the total mass in that box. Using the transfer functions we can formulate the conditions simply by

$$U(z) = H_a(z)[X(z) + V(z)] \quad (18)$$

$$V(z) = H_s(z).U(z) \quad (19)$$

since the transfer functions operate on the inputs. Eliminating $V(z)$ gives

$$U(z) = \frac{H_a(z)}{1 - H_a(z)H_s(z)}X(z) \quad (20)$$

a somewhat more complicated transfer function. We now introduce indices a and s to α in eq. (13) and rewrite eq. (20) as

$$U(z)\left[1 - \frac{(1-\alpha_a)(1-\alpha_s)(1+z^{-1})^2}{4(1-\alpha_a z^{-1})(1-\alpha_s z^{-1})}\right] = \frac{(1-\alpha_a)(1+z^{-1})}{2(1-\alpha_a z^{-1})}X(z) \quad (21)$$

using the more elaborate approximation. We can see from this that the inversion of eq. (20) will read

$$u_t = \lambda_1 u_{t-1} + \lambda_2 u_{t-2} + \mu_0 x_t + \mu_1 x_{t-1} + \mu_2 x_{t-2} \quad (22)$$

the general form for a second-order linear system. In this the parameters λ and μ will be functions of α_a and α_s. However, looking at eq. (20) we note that for $\alpha_s = 1$ we obtain eq. (12), the transform of a first-order linear system. Since

$$\alpha_s = \frac{2-k_s}{2+k_s} \quad (23)$$

α_s would approach 1 for very small values of k_s, i.e. for long turnover times of the surface reservoir. In such a case the return flux v_t will be very small compared to u_t.

We are thus led to consider equations of the form

$$y_t = \sum_{k=1}^{n} a_k y_{t-k} + \sum_{k=0}^{n} b_k x_{t-k} + c \quad (24)$$

as appropriate expressions for an n-reservoir system. Such an expression includes all possible variations in the arrangement of fluxes between the n reservoirs as well as the effect of various ways of approximating the continuous functions. For every such arrangement there will be a specific set of a_k's and b_k's. In theory it should, therefore, be possible to select the proper model structure from a study of the empirical set of a_k's and b_k's because of various restrictions placed on the parameters in the model.

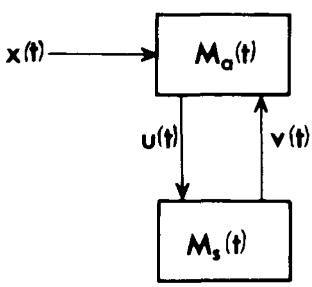

Fig. 2. A possible two-box model of carbon dioxide flux. $M_a(t)$ amount in the atmosphere; $M_s(t)$ amount in the "surface" reservoir.

Since the system is linear we can consider only the perturbation on the atmosphere—earth's surface system caused by fossil fuel combustion, imposed on a presumed stationary state where x_t had been zero for a very long time. The constant c has to be added to take care of some mean properties in x_t and y_t.

Evaluation of the parameters in eq. (24) can, of course, be done in the usual way by using least-square fit to the data in Table 1. The procedure is then one of a multiple regression and we can apply statistical criteria to describe the success of the procedure. However, there will be a limit to the number of parameters to be estimated, which depends on the significance requirements set up.

4. Results

The analysis was done with the aid of a computer programme for multiple regression analysis which computes coefficients stepwise with increasing number of "independent" variables. One run considered a first-order process, i.e. the first case considered.

1. The simplest case has only curiosity value and consists of the fitted equation (16 pairs of variables)

$$y_t = 1.0542\, y_{t-1} + 0.6006 \qquad R^2 = 0.9929$$

(R^2 being the "multiple" regression coefficient squared) which tells that the increase in atmospheric carbon dioxide fits nicely an exponential. It appears as if originating from an initial concentration $y_o = -(0.6006/0{\cdot}052) = -11{\cdot}6$. Since our zero level is at 316 ppm y_0 should be 304.4 ppm.

2. Next, the source term x_t is added which gives

$$y_t = 0.5291\, y_{t-1} + 0.7269\, x_t - 5.852 \qquad R^2 = 0.9964$$

Both coefficients are highly significant (different from zero). Here, a y_0-value is derived by setting $x_t = 0$, $y_t = y_{t-1} = y_0$. We obtain $y_0 = -(5.852/1.05291) = -12.4$, i.e. at 303.6 ppm. As to k, the transfer rate we obtain this from $(1/1 + k) = 0.5291$, i.e. $k = 0.89$. This seems very high. Using the "better" approximation we have $(2 - k)/(2 + k) = 0.5291$ and $k = 0.616$, more reasonable. According to this about 62% of the atmospheric "excess" of carbon dioxide is taken up by the earth's surface yearly.

3. This case includes two x_t terms and reads

$$y_t = 0.6020\, y_{t-1} + 0.8144\, x_t - 0.1960\, x_{t-1} - 4.976 \qquad R^2 = 0.9965$$

The fit is thus about as good as in the previous case. The y_0 becomes $(-4.976/0.3980) = -12.5$, a carbon dioxide concentration of 303.5 ppm, nearly the same as before. For k, the transfer rate we use $(2 - k)/(2 + k) = 0.602$, $k = 0.491$, hence 49% per year transfer of the atmospheric excess. This case then looks quite good. However, the weighting of the x_t and x_{t-1} is not equal as eq. (5) demands. This is, of course, a question of timing observations with the input. We find, however, from the analysis that the coefficient -0.1960 is not statistically significant, whereas the other coefficients are, using conventional significance criteria.

Turning now to second-order linear models, the computation of transfer rates becomes very complicated, depending ultimately on the model structure used. As to the cases considered we obtain

4. $y_t = 0.4544\, y_{t-1} + 0.1102\, y_{t-2} + 0.6878\, x_t - 5.444 \qquad R^2 = 0.9959.$

5. $y_t = 0.5277\, y_{t-1} + 0.1061\, y_{t-2} + 0.7742\, x_t - 0.1899\, x_{t-1} - 4.610 \qquad R^2 = 0.9960.$

6. $y_t = 0.6413\, y_{t-1} - 0.3641\, y_{t-2} + 0.7497\, x_t - 0.5732\, x_{t-1} + 0.9209\, x_{t-2} - 8.791 \qquad R^2 = 0.9975.$

Firstly, considering the fit we notice no improvement comparing cases 4 and 5 to case 3. Only the last case increases R^2 a little. Secondly, in all the cases only the first autoregression coefficient is significant, i.e. the coefficient for y_{t-1}. The coefficient for y_{t-2} is not significantly different from zero. For that reason alone we may question the applicability of the second-order model. It is

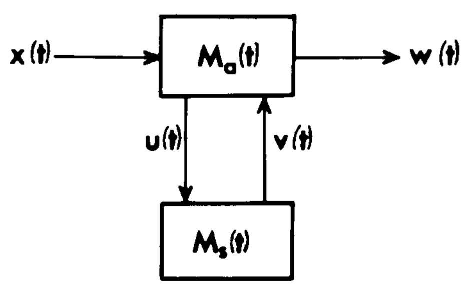

Fig. 3. Another possible two-box model of atmospheric carbon dioxide flux. $w(t)$ is, say, direct loss to deep water or uptake by the terrestrial biosphere. In that case $M_s(t)$ may represent the sea surface reservoir.

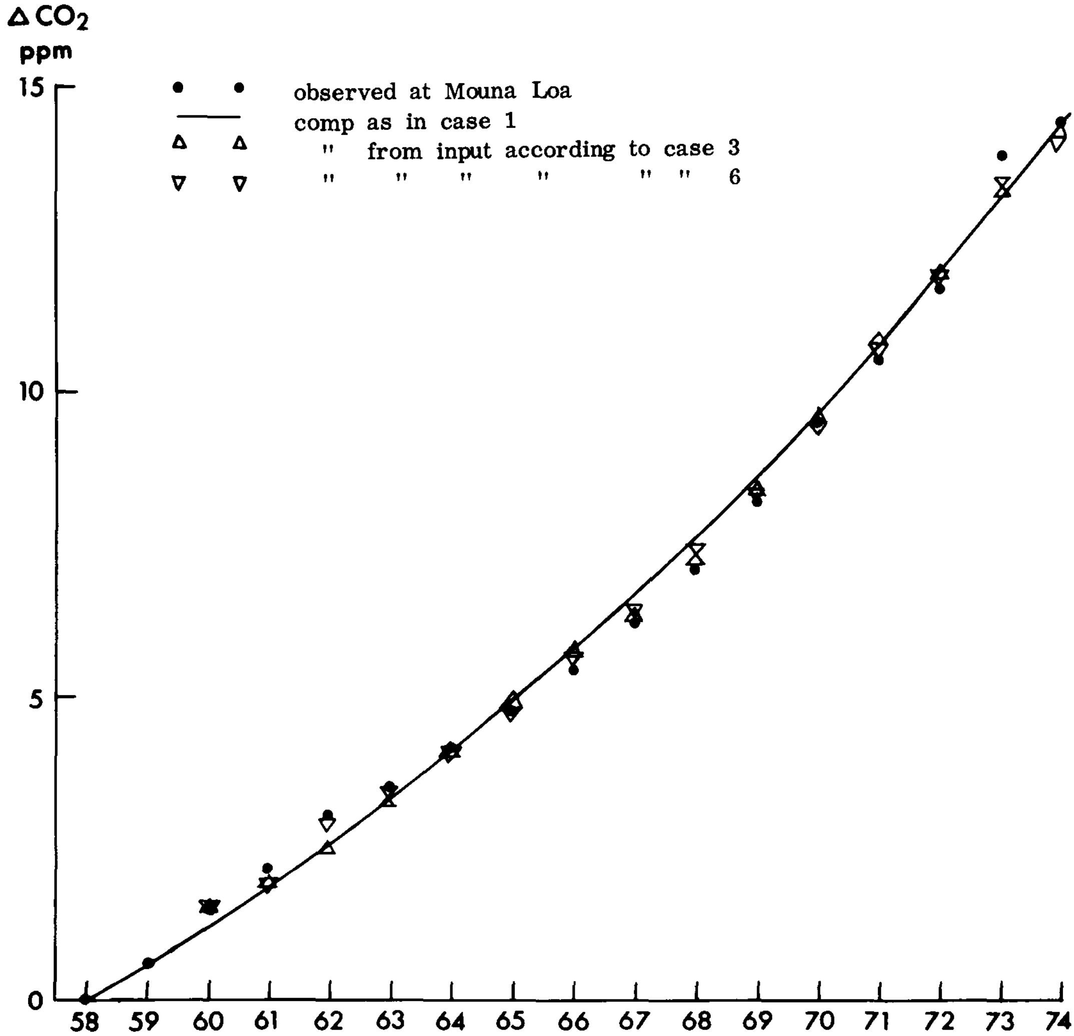

Fig. 4. Yearly averages of atmospheric carbon dioxide.

interesting to note that for case 6 the most reasonable model structure seems to be that one pictured in Fig. 3 with w_t dominating over u_t. However, the data hardly warrant a detailed computation so far. Maybe in future, when more data are available it can be justified.

As to y_0 in the three cases we obtain it as before by putting x_t, x_{t-1} and x_{t-2} equal to zero and assume $y_t = y_{t-1} = y_{t-2} = y_0$. This gives

case 4 $\quad y_0 = 304.0$
case 5 $\quad y_0 = 303.4$
case 6 $\quad y_0 = 303.8$

hence, about the same as before.

The various equations can be used to compute concentrations y_t from inputs x_t starting the series where $y_{t-1} = 0$ or for a second-order model, where $y_{t-2} = 0$. In this computation all succeeding y_{t-1} and y_{t-2} used are those computed earlier. Three sets of computed values are shown in Fig. 4 together with actual observations. One set is computed from case 2, i.e.

$$y_t = 0.5291\, y_{t-1} + 0.7269\, x_t - 5.852$$

the second from the 6 equation,

$$y_t = 0.6413\, y_{t-1} - 0.3641\, y_{t-2} + 0.7497\, x_t - 0.5732\, x_{t-1} + 0.9209\, x_{t-2} - 8.781$$

The third represented by the line is the simple exponential increase obtained in case 1. As Fig. 4 demonstrates, the fit to observed data is remarkably good. The difference between the sets of computed values, is not very pronounced except for the exponential case.

The result is to some extent disappointing. The past records do not contain much information on the system. One reason for this, perhaps the most important, is the "deterministic" nature of the input function x_t, being almost exponential. This effect can be demonstrated by the following example. Consider eq. (24) with x_{t-k} being a geometric progression given, say, by

$$x_{t-k} = \beta x_{t-k-1} \tag{25}$$

This corresponds to an exponential increase in continuous time ($\beta > 1$). Obviously we can also write

$$x_{t-k} = \beta^{-k} x_t \tag{26}$$

so that eq. (24) could be written

$$y_t = \sum_{k=1}^{n} a_k y_{t-k} + x_t \sum_{k=0}^{n} b_k \beta^{-k} \tag{27}$$

Now we assume that a solution to this is also a geometric progression such that

$$y_{t-k} = \alpha^{-k} y_t \tag{28}$$

With this relation eq. (27) becomes

$$y_t = y_t \sum_{k=1}^{n} a_k \alpha^{-k} + x_t \sum_{k=0}^{n} b_k \beta^{-k} \tag{29}$$

i.e.

$$y_t = \frac{\sum_{k=0}^{n} b_k \beta^{-k}}{1 - \sum_{k=1}^{n} a_k \alpha^{-k}} \cdot x_t \tag{30}$$

Since x_t is a geometric progression y_t will also show a geometric progression, hence relation (28) is valid. We can thus write

$$y_t = A x_t + \text{constant} \tag{31}$$

where A is a constant, when the input term is a geometric progression. A will of course depend on the number of boxes and their internal arrangement. However, eq. (31) permits estimate of two parameters only one of which is a measure of means. Only one parameter describing the system can be obtained irrespective of how many boxes are employed in the model. The answer is trivial. For an exponential input the output will be an exponential function irrespective of the order of the model, as long as it is linear. This is perhaps one reason for the "success" of various linear models proposed in the past. And it explains why it is difficult to extract any reliable information on the system from the data. Nor is the observed increase in atmospheric carbon dioxide of much help in validating models. Another reason for the poor success may be inherent in the Mauna Loa data. We may raise the question: Do the yearly averages from Mauna Loa really represent *global* yearly averages? This means that we do not require that the Mauna Loa data be exactly global averages but the analysis requires that the Mauna Loa data should be proportional to the global averages. If they are not, a "noise" is introduced into the calculations which increases the uncertainty of the parameter estimates. Since most of the carbon dioxide is released in the northern hemisphere there must be a large-scale eddy flux into the southern hemisphere. The large-scale eddy pattern over the tropics is in no way stationary from year to year. Hence, the effectiveness of the transport across the equator may fluctuate from year to year or at least may show some climatic fluctuations. Hence, also the Mauna Loa yearly means may contain climatic fluctuation patterns.

In Fig. 5 differences between observed and computed for cases 3 and 6 are shown and it is seen that the differences display a considerable persistence resembling to some extent short-term climatic fluctuations. At first sight one would perhaps interpret this as a failure of the model. Although the models are primitive they would hardly respond in this way for a steadily increasing output of carbon dioxide. Hence, climatic fluctuations are more likely.

A final note on the y_0's, discounting case 1 which is not of much interest. In the other cases the standard deviation of this estimate is close to 4 ppm. Hence, the computed y_0's could be somewhere between 300 and 310 ppm. The value normally quoted for the pre-industrial state is 290 ppm, thus well below y_0.

The computed y_0's can in general not be taken to represent initial atmospheric carbon dioxide concentrations. The exception to this is case 1 provided the output had been a simple geometric progression at all previous times. Since this is doubtful, an extrapolation back in time is hardly permitted. In

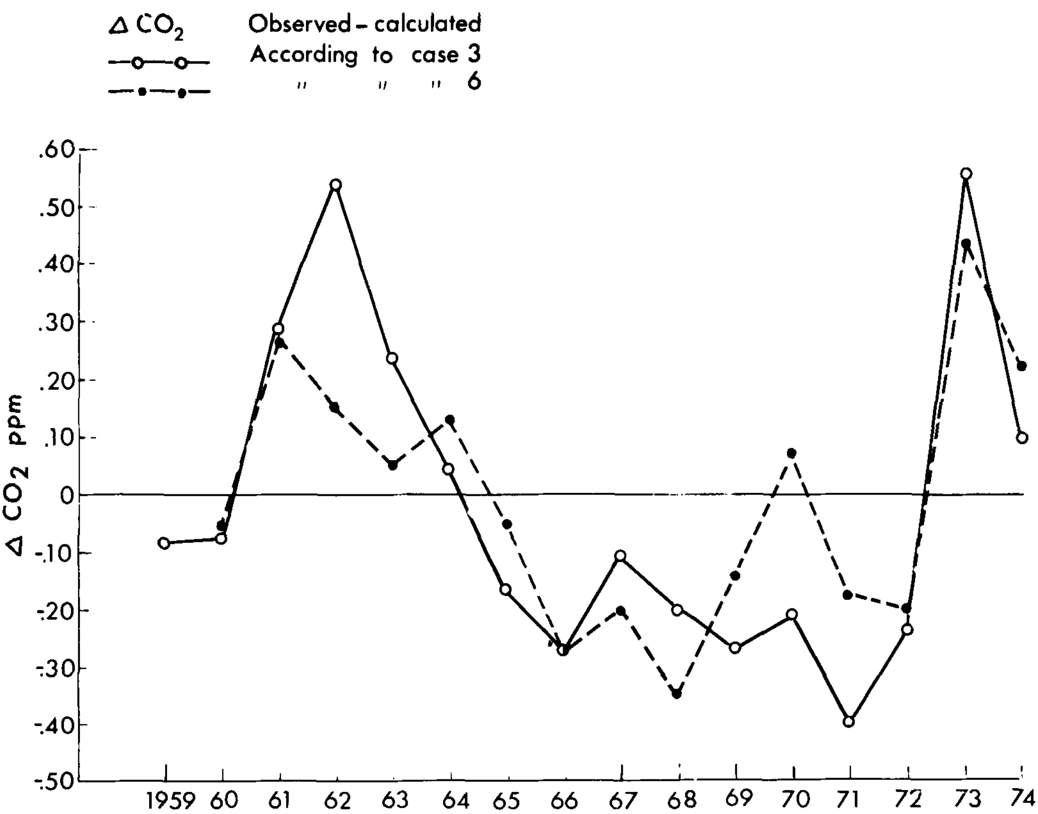

Fig. 5. Differences between observed and computed CO_2 for two different models.

all the other cases y_0 is the equilibrium concentration which would be obtained some years after the output of fossil CO_2 stopped. The actual decrease during such a phase can be simulated easily from the equations given. Hence, y_0 is more a predicted value than an extrapolation back in time. The expectancy value of y_0 in those cases is independent of the source function x_t. There is an implicit assumption made in the calculations. Somewhere in nature there is an infinite pool of CO_2 which takes care of all excess due to fossil combustion. Hence, any pre-industrial state must be the same as a post-industrial state. In reality this is of course not true. The real pool is not infinite, so whatever is taken up is likely to increase the pre-industrial equilibrium level. Unfortunately, because of the near exponential increase in output, it is not possible to obtain information from the data on this large pool, presumably represented by the deeper parts of the oceans.

With this in mind one can provisionally consider the y_0's computed as the new equilibrium level in the atmosphere caused by past inputs of "excess" CO_2 into the oceans. However, even so it cannot be taken to be a very accurate measure of this new level.

The equations derived can of course be used for predictions.

REFERENCES

Anonymous 1976. Environmental issues. A report to the International Council of Scientific Unions by the Scientific Committee on Problems of the Environment SCOPE report nr 10, Paris August 1976.

Bocastow, R. and Keeling, C. F. 1976. Atmospheric Carbon Dioxide and Radiocarbon in the Natural Carbon Cycle. Part II. In: *Carbon and the Biosphere.* AEC Symp Series 30.

Bolin, B. 1970. The Carbon Cycle. In: *The Biosphere*, San Francisco, California: W. H. Freeman & Co.
Bolin, B. 1975. A critical appraisal of models for the Carbon cycle. Appendix 8 in: *The physical basis of climate and climate modeling.* GARP publ serie No 16.
Machta, L. 1973. Prediction of CO_2 in the Atmosphere. In: *Carbon and the Biosphere.* AEC Symposium Series 30.

ПРОБЛЕМА АТМОСФЕРНОГО УГЛЕКИСЛОГО ГАЗА

Представлен анализ данных Мауна Лоа по двуокиси углерода, рассматриваемых как линейная система с оцениваемым производством двуокиси углерода из ископаемых топлив на входе и с наблюдаемыми среднегодовыми значениями концентрации двуокиси углерода в атмосфере на выходе. Анализ дан на основе дискретного времени с одним годом в качестве шага по времени. Из-за близкого к экспоненциальному роста в производстве ископаемого топлива не так уж много информации о параметрах системы можно извлечь из данных за исключением того, что в настоящее время 50–60% "избытка" атмосферного углекислого газа за год переносится к земной поверхности (океаны и поверхности суши). Однако, результаты указывают, что доиндустриальный равновесный уровень двуокиси углерода в атмосфере мог увеличиться до результирующего "равновесного" уровня примерно на 10 промилле выше. Это, однако, может быть из-за медленной скорости обмена между глубинными водами океана и его более поверхностными слоями, параметра, который невозможно получить из имеющихся входных и выходных данных.