Tellus (1977), *29*, 435–444

Comparability of CO_2 measurements

By WALTER BISCHOF, *Department of Meteorology, University of Stockholm, International Meteorological Institute in Stockholm,*[1] *Arrhenius Laboratory, S-106 91 Stockholm, Sweden*

(Manuscript received June 15; in final form November 10, 1976)

ABSTRACT

In measuring the atmospheric CO_2 concentration, corrections for the carrier gas effect are required at present for each individual gas analyzer. It is shown that standard gases of CO_2/air composition offer less complication in the data comparison than CO_2/N_2 standards which are most commonly being used for calibration. CO_2/air standards which are used in Stockholm and related to the Scripps manometric calibration scale have remained stable over more than 10 years.

Aircraft data from the upper troposphere and the lower stratosphere obtained in the Stockholm project are in fair agreement with data from the Mauna Loa and the South Pole stations. An accelerating increase over the period 1963 to 1975, i.e. about 0.5 ppm/yr at the beginning and about 1.3 ppm/yr at the end of the period, has been observed.

1. Introduction

The measurement of atmospheric CO_2 has been given high priority and is an essential component of the WMO network of background air pollution stations. Suggestions of how to organize such a monitoring program have been made by the author (Bischof, 1974) in Chapter 3.3 of the WMO Operations Manual No. 299. For further discussion of possible errors and how to maintain an accurate monitoring program it is appropriate to cite some paragraphs from the manual.

> "The infrared gas analysis is a *relative* measurement and requires calibration with well-defined standard gases. Most standards are composed of mixtures of CO_2 in air, or CO_2 in N_2. However, the results are most sensitive to the kind of carrier gas (i.e. N_2 or air) being used for CO_2. Therefore, the analyzer as well as the standard gases must be calibrated against an *absolute system.* Such a system exists at the Scripps Institute and has been used as a basic calibration station for most of the CO_2 projects (Keeling, 1958). The determination of primary standards at Scripps is related to CO_2/N_2 mixtures.

Some complications exist at present because of the fact that differing types of analyzers and standard gases are used in different projects. It becomes clear from experiments made recently that this could lead to significant errors in the data interpretation (Bischof, 1973; Pearman & Garratt, 1974). The results clearly demonstrate how sensitive the IR gas analysis is with regard to the carrier gas composition of standards (i.e. N_2 or air) and the type of analyzer being used. Thus differing analyses are indicated when the N_2/O_2 ratio of a carrier gas is different in spite of the very same CO_2 concentration of the standard. It is likely that this is due to absorption line broadening involved in the IR gas analyses method.

This influence does not appear if the gas composition of sample and standard is identical. Thus the intercalibration of standards having the same carrier gas, e.g. N_2 or air, is not affected. However, in analyzing air samples, the result will differ by several ppm when changing from CO_2/air to CO_2/N_2 standards. In this case, the indicated deviation is also dependent on the

[1] Contrib. No. 345. This work has been sponsored by the Swedish Natural Science Research Council under contract No. G 0223-060.

type of analyzer. Pressure (or collision) broadening influences are reflected in the differing results obtained when using CO_2/N_2 standards at different ambient pressures. Although gas analyzers of different types are based on the same principle of operation, some characteristic differences exist in the design, e.g. the absorption band selection, the dimensions of the measuring cell and the filling of the receiver cell . . .

The conclusions are that comparable results can be obtained only if standards and analyzers are calibrated using an absolute concentration determination. At present, an intercomparison of calibration and analysis methods among the different projects is of fundamental importance. It is likely that the complications would become less if the same kind of standards were used (and probably also the same type of analyzer). However, since experiments still being carried out may result in improved techniques and better comparability between systems, no suggestions for standardization of gas standards and analyzers are made here."

Investigations made in the meantime at Scripps have shown that important adjustments are required, which are due to errors by drift, nonlinearity and carrier gas effect in the calibration system. Hence, three different calibration scales exist at present at Scripps, namely the 1959 "Index", the "Manometric" and the "new adjusted" scale. Data reported by Scripps as well as by our group have so far been related to the "Manometric" scale which has not been adjusted. In investigating the comparability of CO_2 measurements, data reported in this paper are related to this scale if not otherwise indicated.

2. Intercomparison of standards

An intercomparison of standards has been carried out recently by Pearman (1977). A number of gas tanks containing CO_2/N_2 as well as CO_2/air mixtures were analyzed at different stations. All stations used CO_2/N_2 standards except Stockholm where CO_2/air standards were employed.

The results show good agreement between the CO_2/N_2 gases at all stations using these standards. For CO_2/air gases, however, deviations up to about 7 ppm were obtained, due to the N_2 carrier effect.

In a world-wide monitoring program we are generally interested in the analysis of air samples. Since all projects are comparing their standards with the Scripps calibration system, a comparison of their analyses of air samples and those obtained at Scripps is therefore of some interest. Such a comparison is shown in Table 1 using Pearman's Table 2 and by converting those values into differences relative to the Scripps results. Thus in comparing analyses of air samples as made at different stations, corrections corresponding to the numbers in Table 1 are required. It is seen that corrections for Stockholm standards are negligible.

However, the establishment of standards at Scripps is based on using CO_2/N_2 standards, therefore when analyzing individual air samples an additional correction is required for the carrier gas

Table 1. *Deviation of analysis results for three air standards analyzed at different stations and related to Scripps analysis data. (Calculated from Pearman, 1975, Table 2, station numbers are given in the same order.) Station 4 is Scripps Institution, the values given in this column are analysis results. Station 9 is Stockholm*

	Station									
Tank	1	2	3	4	5	6	7	8	9	10
DA 10	−0.65	−0.52	−0.60	316.61	4.88	0.03	5.96	(−3.28)	0.04	6.73
DA 8	−0.85	−0.97	−0.73	326.19	4.91	0.35	5.97	0.55	0.02	(8.72)
DA 3	−0.54	−0.25	−0.41	327.13	4.96	0.30	5.87	0.50	−0.17	6.98
Mean	−0.68	−0.58	−0.58		4.91	0.22	5.93	0.53	−0.04	6.85

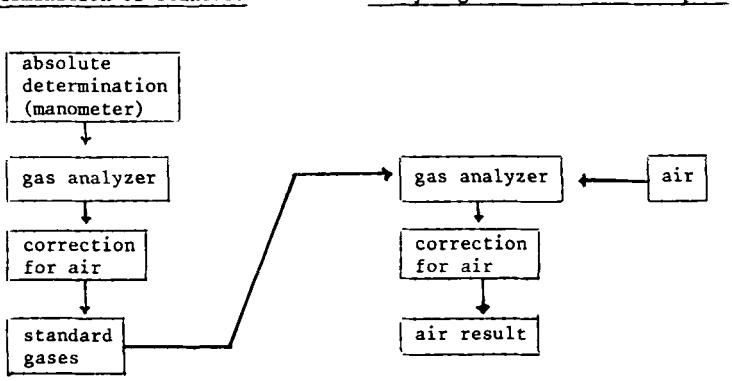

Fig. 1. Block diagram for N_2 standards.

effect on the analyzer used for standard determination. Thus, as shown in the block diagram, Fig. 1, two corrections are necessary for stations using N_2 standards (one for each individual analysis and one in determining the standards). This is not the case when using CO_2/air standards. Thus if air standards were used also at Scripps and at all stations no correction would in principle be necessary.

3. Carrier gas effect

The carrier gas effect for three different analyzers is demonstrated in Fig. 2 (Bischof, 1975). The deviation in the analysis results obtained from CO_2/N_2 mixtures (i.e. at $R = 0$) relative to CO_2/air mixtures (i.e. at $R = 0.21$) is about −6, −2, and +7 ppm, respectively.

Nearly identical results were obtained using any one of the three analyzers if CO_2/air mixtures were used. However, the data presented from this experiment probably demonstrate the carrier gas effect rather qualitatively and do not permit a conclusion with regard to the carrier gas error (see for example Pearman, 1977). In fact, even less errors can be obtained from all three analyzers if *air standards* are used instead of experimental mixtures, see Table 2.

4. Calibration

4.1. *Analyzers*

The IR–gas analyzers in principle show a non-linear relationship between analyzer response and gas concentration. Within a limited range of operation, a linear relationship can, however, be used. In practice, analyzers are designed for zero-

Table 2. *Calibration results (ppm) obtained from different analyzers. In the first column the standard mean is calculated from 1963–74 intercalibration of our standards (for details, see further section 4)*

Standard No.	Standard mean	1968-03-20		1976-03-19		
		URAS 1	UNOR 2	UNOR 2	UNOR 5	UNOR 20
3	313.51	313.52	313.58	313.55	313.53	313.52
5	321.97	321.98	321.97	322.00	321.98	322.00
7	318.57	318.58	318.58	318.44	318.53	318.57
8	313.98	313.97	313.81	—	—	—
10	321.40	321.39	321.39	321.47	321.42	321.37
11	332.06	—	—	332.06	332.07	332.07

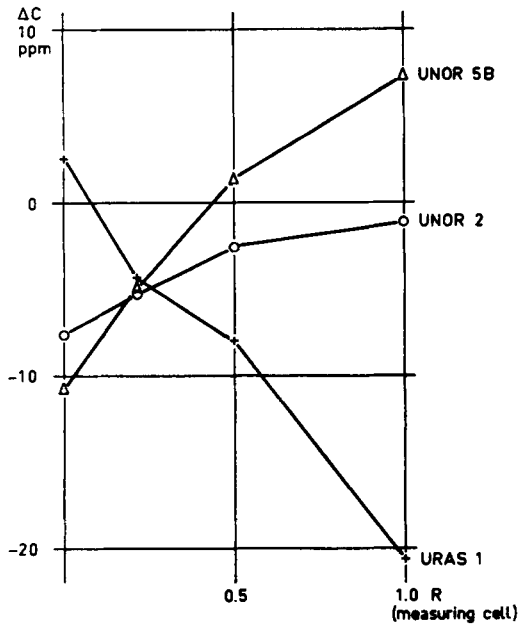

Fig. 2. The carrier gas effect shown for three different analyzers, from Bischof (1975). (*R* is the ratio $O_2/(N_2 + O_2)$.)

Table 3. *Intercalibration of four 1-l standards (CO_2/air) at Scripps and in Stockholm*

Standard No.	Predicted Scripps results	Calculated Stockholm results
1-1	305.81	305.78
1-2	333.49	333.58
1-3	329.00	328.85
1-4	316.75	316.84

point suppression and are equipped for appropriate amplification to full range response. Because of some differences in their design, different characteristics can be observed. Thus, different analysis results can be obtained if the calculation is based on the assumption of linear response and using two reference points. It should be noted, however, drift or scatter occur in the interpolated standards if the analyzer characteristics change by time.

In order to obtain comparable results from different analyzers, an attempt was made to describe the characteristics of each analyzer using a non-linear function. In view of the accuracy of

Table 4. *Calibration of the Stockholm analyzer (UNOR 2) at Scripps Institution*

Scripps standard No.	Scripps index	UNOR 2 result
4293	316.31	316.34
4289	312.03	312.02
2399	322.31	322.29
4288	328.23	328.24
18220	340.43	340.43

0.3 ppm obtained for our standards from calibrations at Scripps, a parabolic function was found to be adequate. Hence, concentration and analyzer response are used as variables in a regression program, and concentration for reference points interpolated.

Table 3 shows the result of a comparison between calibrations at Scripps and values computed as described above for four 1-l CO_2/air temporary standards. An almost linear relationship is obtained for this analyzer (UNOR 2). This was further confirmed by the results obtained from a direct calibration of this analyzer, where five Scripps basic standards (CO_2/N_2) with well-defined concentrations were used in a usual run of ten comparisons, see Table 4. The good agreement between measured and calculated concentrations indicates that the characteristics of our UNOR 2 and the Scripps APC analyzer are similar, at least within this operation range (the carrier effect does not affect this comparison).

These findings permit us to re-check our URAS 1 which was found to have non-linear characteristic using three primary standards calibrated at Scripps in 1964. Fig. 3 shows the deviation from linearity as expressed by the difference between linear and parabolic computation, which thus demonstrates the error which would occur if linearity were assumed. This error can of course be reduced by the choice of a smaller span, providing that such standards are available. However, as will be shown and further discussed in the next paragraph, this rather extreme non-linearity changes to become almost linear if calibration data from the 1969–74 period are used. About the same linearity exists also in our UNOR 5 B which has been in operation since 1972, as well as in our UNOR 2/200 which has a 200 ppm range and has been used since 1975.

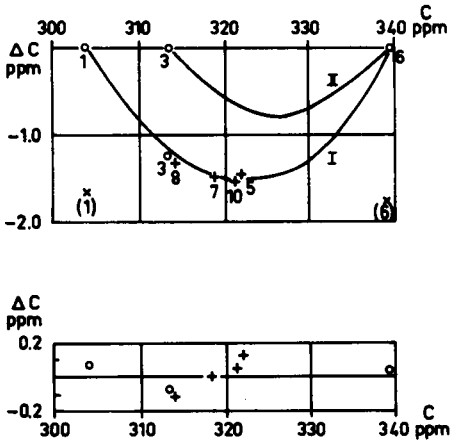

Fig. 3. Calibration results, URAS 1, with predicted standards Nos. 1, 3 and 6, according to Scripps calibrations in 1964. (a) Deviation from linearity if standards No. 1 and No. 6 (I), and standards No. 3 and No. 6 (II) are chosen as low and high span standards respectively. The curves are calculated by parabolic regression, symbols represent interpolated values from the computation. (b) Deviation between calculated and interpolated values. × marks the "shift" between the 1964 and 1969–74 data.

4.2. *Standard gas composition*

In the Scripps project, N_2 standards are used "because it was not known how stable gas mixtures containing oxygen could be, and it was believed that an inert gas such as N_2 appeared more likely to maintain integrity than oxygen containing gas mixtures" (according to Dr Keeling at the WMO expert meeting 1975).

In Stockholm, air standards have been used, because it was felt that calibration gases should have the same composition as the medium to be analyzed, provided that such standards can be kept stable enough over a longer time. The experience of others (see for example Egle & Schenk, 1951) and long-range experiments carried out in our laboratory in co-operation with the gas manufacturer (AGA Stockholm) have shown that air standards indeed remain stable if tanks and gases are treated in the way as explained in the WMO Manual No. 299. Concentration drift occurs in the N_2 standards as well as in CO_2 standards, if tanks and standards are not dried adequately.

4.3. *Intercomparison of standards*

A set of ten standards stored in 50-l gas tanks have been available since 1963 in our laboratory and some of them are still in operation while others have been re-filled after several years of use and recalibrated. Some of these (Nos. 1, 3, 6 and 11), "selected primary standards", were shipped and calibrated at Scripps during 1964 and 1971. Additional intercalibrations were made occasionally by means of 1-l gas tanks being filled at our institute and shipped to Scripps. Intercomparisons of our ten standards are made periodically in our laboratory. Three different analyzers have been in operation during this period and one of them was brought to Scripps for calibration in 1970 (see Section 4.1).

Thus, we are able to study the behavior of our standards and analyzers over a long period of time, thanks to the cooperation with the Scripps Institute. As will be seen below, the results, which in fact are too numerous to be reported here in detail, show that our CO_2/air standards have remained stable with a few exceptions.

Fig. 4 shows the results of the calibrations of primary standards Nos. 1, 3, 6 and 11 (*) and the 1-l (+) standards at Scripps. The circles indicate values obtained for intermediate standards using the non-linear regression program as described in Section 4.1. Gaps exist in 1965–67 and 1972–73 because of the lack of intercalibrations.

Even though more numerous intercalibrations would have been desirable the results show (Fig. 4): Four standards (Nos. 3, 5, 7 and 10) which are still in operation after 12 years have remained unchanged in relation to Scripps standards. In particular, the constancy of No. 3 is controlled by calibrations at Scripps in 1964 and 1971. One standard (No. 8) remained unchanged over its entire lifetime of 7 years. One standard (No. 11), i.e. the refilled tank of No. 8, remained unchanged over more than 3 years and is still in operation. These six standards obviously cover the total range of atmospheric CO_2 variation in our project. On the other hand two gases (Nos. 2 and 9) have clearly drifted and have therefore not been used as standards.

The behavior of three standards, Nos. 1, 4 and 6, is different. The first two seem to have changed between 1964–68 and the same is apparent for No. 6 from the values extrapolated with the aid of the other standards using the calibration curve discussed above. It is remarkable, however, that this shift occurred only in our low and high standards whereas the others remained stable with an accuracy of 0.2 ppm. In examining the response

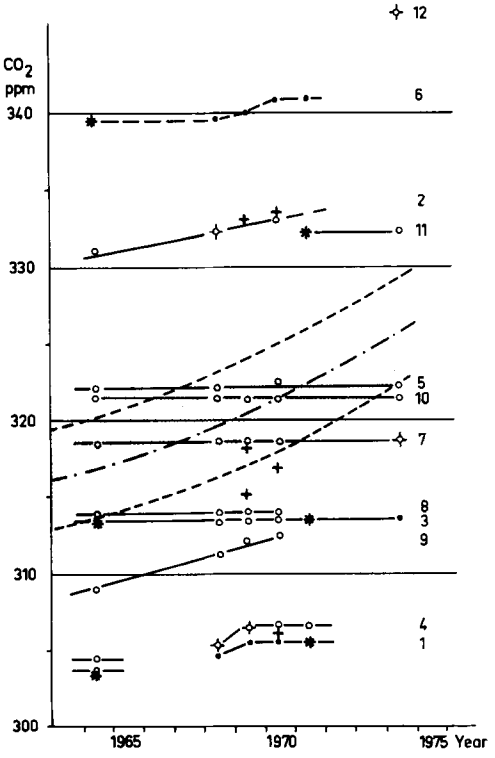

Fig. 4. Calibration results according to Scripps calibration numbers.
* Primary standards, calibrated at Scripps.
+ 1-l standards, calibrated at Scripps.
o Annual means, calculated by interpolation.
• Annual means, calculated by extrapolation.
The number assigned to each individual standard is given as well as the range of actual air sample analyses as it has changed during this time (range between dashed lines).

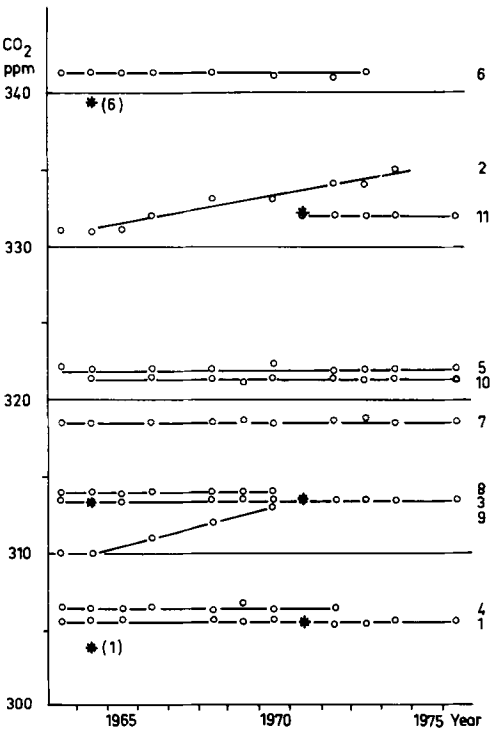

Fig. 5. Calibration results obtained from predicted standard concentration numbers as calculated from the 1969–74 averaged means. (Symbols as for Fig. 4.)

characteristics of our analyzers using Scripps standards it was, however, found that the shift noted corresponds to a remarkable change. Calibrations made after 1969 yield an almost linear response function whereas calibrations made earlier correspond to a quasi-parabolic one as shown in Fig. 3. It should be noted that the deviation from linearity of the URAS 1, as given before 1969, was far too large according to the manufacturer's calibration data. Although it is not claimed that the analyzer characteristic should necessarily be linear, we do not understand why such a drastic change to a quasi-linear one should occur simultaneously with the shift.

If we use the response characteristics of the analyzers as obtained from the calibrations after 1969 and deduce the values that would apply for the standards Nos. 1, 4 and 6 in 1964 (see Fig. 5), we find that all three are changed throughout this period. It should further be noted that the other standards are practically unchanged which is not surprising since they are all in the range 313 to 323 ppm and thus close to the minimum in the parabolic calibration curve (see Fig. 3). In particular we note that the standard deviation for each individual standard averaged for the whole period does not exceed 0.17 ppm (standards Nos. 2 and 9 excluded) and conclude that a practically linear response curve gives a better fit as indicated by a higher correlation coefficient and a lower standard error of estimate. Thus in reality only two of our standards (Nos. 2 and 9) may have drifted and the comparatively large number of standards is a good guarantee that the calibration system employed has been stable. Although this finding does not explain the discontinuity, it is obvious that our standards

within the region 310 to 330 ppm are not affected, nor are our air sample analyses, being situated within this range.

5. Data comparison

5.1. *Accuracy of measurements*

Our sampling and analysis technique has been described elsewhere (Bischof, 1970). Improvements have been made in course of time. Particularly the paper strip recorder was replaced by a Digital voltmeter in 1970, which provides higher accuracy in the data evaluation. Our analyzer has also been connected to a computer system which stores the data for further calculations. Hereby, a reproducibility of the measurement of about 0.03 ppm has been obtained. The accuracy of the analyses is of course lower, i.e. about 0.05 ppm for standard gases and about 0.1 ppm for flask samples. Generally speaking, the accuracy has improved by about a factor of two since 1970.

5.2. *Comparison of air samples*

The longest record of atmospheric CO_2 measurements exists from the Mauna Loa and the South Pole stations of the Scripps project (Ekdahl & Keeling, 1973). In our project, air samples have been collected mainly on flights over the North Atlantic and over northern Sweden. Data from different periods have been published elsewhere (Bischof, 1965; Bischof & Bolin, 1966; Bolin & Bischof, 1970; Bischof, 1973). Revised data for the period 1963–74 are presented in Figs. 6 and 7 for the upper troposphere and the lower stratosphere. Crosses represent sample means. The mean seasonal variation and the trend for this period are also shown. The seasonal variation represents a best fit of actual sample means with regard to a quadratic function for seasonal trend. The seasonal trend has been approximated by a parabolic regression equation, which has been determined to

$$C_t = C_0 + 0.542t + 0.0314t^2$$

using the method of least squares. C_0 is the annual mean for 1963, i.e. 316.28 for the troposphere and 315.18 for the stratosphere, respectively, t is the time in years ($t = 0$ for 1963). The calculated seasonal variation and the trend have been extrapolated through 1975 and actual data not used for computation added to the graph.

Since Mauna Loa and South Pole data related to the 1959 manometric scale frequency appear in the literature and are being used for considerations of exchange processes and possible climate effects, a comparison with our data is of interest. In Fig. 8 the trend as given in Figs. 6 and 7 has been plotted together with the data published by Ekdahl & Keeling (1973).

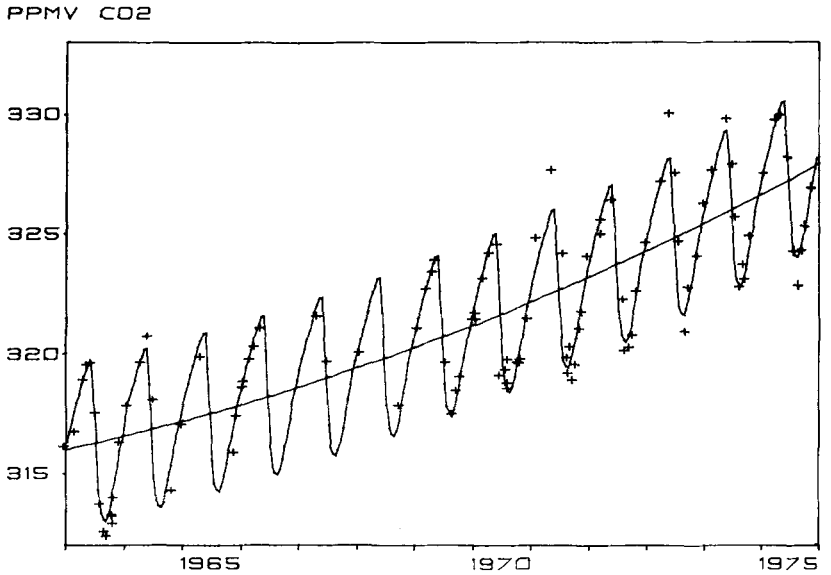

Fig. 6. CO_2 seasonal variation and increase trend, upper troposphere (Northern Hemisphere).

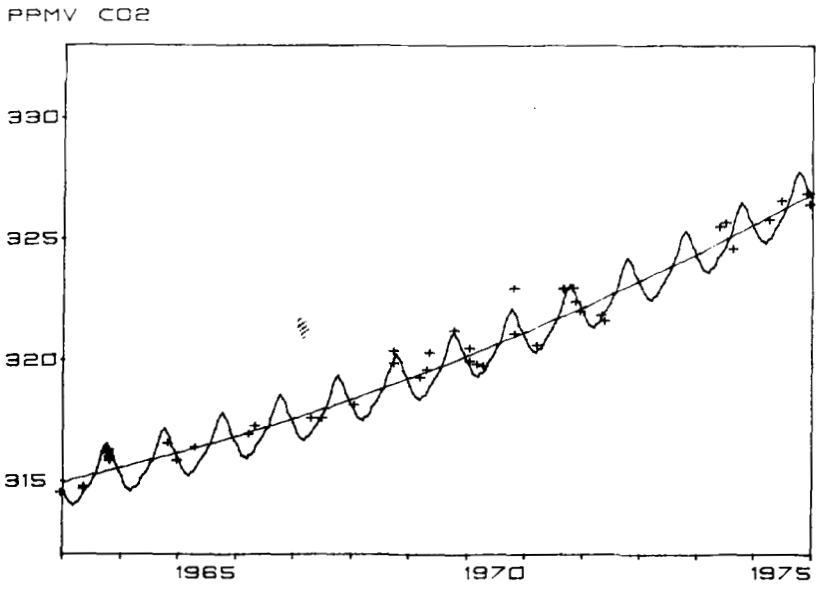

Fig. 7. CO_2 seasonal variation and increase trend, lower stratosphere (Northern Hemisphere).

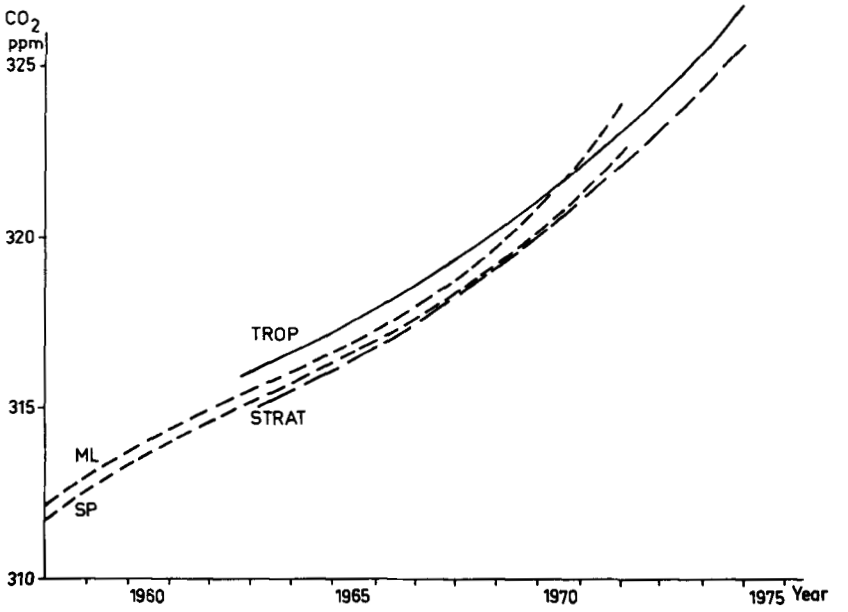

Fig. 8. The seasonal trend of atmospheric CO_2 in the upper troposphere and lower stratosphere (Northern Hemisphere) from Figs. 6 and 7, and from the Mauna Loa and the South Pole stations, according to Ekdahl & Keeling (1973).

Fig. 9 is a comparison of annual increase rates during this 15-year period using a different data set. The Stockholm data for the period 1963–68 (Bolin & Bischof, 1970) were calculated using a linear trend because of the short period of investigation. The results from three NOAA stations, as reported by Miller (1974), have also been included in the graph.

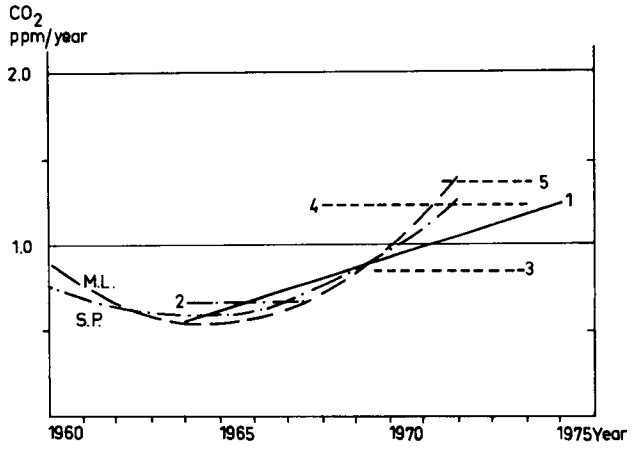

Fig. 9. CO_2 annual increase rate at different stations.
ML Mauna Loa
SP South Pole
1 aircraft, upper troposphere and lower stratosphere Northern Hemisphere
2 same as 1
3 "Charlie"
4 Nivot Ridge
5 Point Barrow

6. Summary and conclusions

By using calibrations of our standards made at Scripps, it is shown that our calibration system in general has remained stable over more than 10 years in relation to the Scripps manometric scale. Intercalibrations of our standards ensure that they have not drifted in relation to each other, with the exception of two which are not in use as standards any longer. Although a discontinuity in our high and low standards remains unexplained so far, there are reasons to believe that this shift is due to inadequate calibration at an initial stage. However, our air sample analyses are in the range of six stable standards and our results have therefore not been influenced. A comparison of our data with those obtained at Scripps (Mauna Loa and the South Pole) is therefore possible. It is found that the different data fit each other well.

In a world-wide monitoring program, as proposed by WMO, which includes N_2 standards, corrections for gas-composition and analyzer characteristic are at present needed at each station. Such corrections are practicable but presume careful control as verified in some detail here. It is evident, however, that the application of CO_2/air standards offers less complication in the processing and intercomparison of data. It is shown that such standards indeed can be kept stable enough over a sufficiently long time.

The Scripps calibration scale has been subject to changes, an inevitable re-calculation of our data is therefore necessary. From preliminary information (personal communication from Dr Keeling), the consequences of the new scale on our data would be as follows:

According to drift at Scripps, our standards, previously being calibrated there, should be corrected by −0.69 ppm for January 1963 and gradually by +0.06 ppm/yr for the following years.

Non-linearity of the Scripps analyzer includes corrections not exceeding −0.4 ppm for our operation range.

Due to the carrier gas error at Scripps, a correction by +3.9 ppm becomes necessary for our standards.

From our point of view, it is likely that adjustments of our data within the accuracy of analysis, i.e. 0.3 ppm, generally given for our standards at Scripps, are justified; however, it seems very unlikely that our entire calibration system has drifted.

7. Acknowledgements

We want to thank Dr C. D. Keeling and his staff at the Scripps Institute for many years' co-operation, particularly for the calibration of our standards, which enabled us to produce our data. The author is grateful to Professor B. Bolin for a critical review of the manuscript.

REFERENCES

Bischof, W. 1965. Carbon dioxide concentration in the upper troposphere and lower stratosphere, Part 1. *Tellus 17*, 365–402.

Bischof, W. 1970. Carbon dioxide measurements from aircraft. *Tellus 22*, 545–549.

Bischof, W. 1971. Carbon dioxide concentration in the upper troposphere and lower stratosphere, Part 2. *Tellus 23*, 558–561.

Bischof, W. 1973. Carbon dioxide concentration in the upper troposphere and lower stratosphere, Part 3. *Tellus 25*, 305–308.

Bischof, W. 1974. The measurements of atmospheric CO_2. Principal part of chapter 3.3 in the WMO Operations Manual No. 299.

Bischof, W. 1975. The influence of the carrier gas on the infrared gas analysis of atmospheric CO_2. *Tellus 27*, 59–61.

Bischof, W. & Bolin, B. 1966. Space and time variations of the CO_2 content of the troposphere and the lower stratosphere. *Tellus 18*, 155–159.

Bolin, B. & Bischof, W. 1970. Variation of carbon dioxide content of the atmosphere in the Northern hemisphere. *Tellus 22*, 431–442.

Egle, K. & Schenk, W. 1951. Die Anwendung des Ultrarot-Absorbtionsschreibers in der Photosyntheseforschung. *Ber. Dsch. Bot. Ges. 64*, 180–197.

Ekdahl, C. A. & Keeling, C. D. 1973. *Atmospheric carbon dioxide and radio carbon in the natural carbon cycle. Carbon and the biosphere.* Published by Technical Information Center, Office of Information Services, U.S. Atomic Energy Commission.

Keeling, C. D. 1958. The concentration and isotropic abundances of atmospheric CO_2 in rural areas. *Geochim. et Cosmochim. Acta 13*, 322–334.

Miller, J. M. (editor), 1974. Geophysical monitoring for climatic change No. 3, NOAA, Boulder, Colorado, U.S.A.

Pearman, G. I. & Garrat, J. R. 1975. Errors in atmospheric CO_2 concentration measurements made at monitoring stations. *Tellus 27*, 62–65.

Pearman, G. I. 1977. Further studies of the comparability of baseline atmospheric carbon dioxide measurements. *Tellus 29*, 171–181.

СРАВНИМОСТЬ ИЗМЕРЕНИЙ CO_2

При измерениях концентрации CO_2 в атмосфере для каждого отдельного газоанализатора в настоящее время нужны поправки на эффект газа-носителя. Показано, что стандартные газовые смеси CO_2–воздух вызывают меньше осложнений при сравнении данных, чем стандартные смеси CO_2–N_2, наиболее часто используемые для калибровки. Стандартные смеси CO_2–воздух, используемые в Стокгольме и связанные со скриппсовской манометрической калиброво чной шкалой, остаются стабильными уже более 10 лет. Самолётные данные по верхней тропосфере и нижней стратосфере, полученные в Стокгольмском проекте, находятся в хорошем согласии с данными станций Маука Лоа и [Южный полюс. Наблюдалось ускоряющееся увеличение концентрации CO_2 за период с 1963 по 1975 гг., именно. $0{,}5 \cdot 10^{-6}$ за год в начале и около $1{,}3 \cdot 10^{-6}$ за год конце периода измерений.