

Institute of Nuclear Sciences INS-R--363

**A Review of
Analyser Calibration Procedures and
Calibration Gas Mixtures
used in the New Zealand
CO₂ Monitoring Program**

M. R. Manning and K.P.Pohl

July 1987

Institute of Nuclear Sciences
D.S.I.R.
Lower Hutt, New Zealand

ABSTRACT

The New Zealand atmospheric CO₂ program involves regular calibrations of gas analysers using sets of calibration gases whose concentrations are assigned by the World Meteorological Organization Central Calibrating Laboratory (CCL). This report describes the calibration procedures used and presents a summary of our calibrations in the period from 1983–1987, involving two gas analysers, 45 calibration gases, and 140 independent calibrations.

The analysis presented here separates the repeatability of the analyser measurements from an apparent consistency in the concentrations assigned by the CCL. This is believed to be the most extensive assessment undertaken so far of gas analyser calibration procedures and to provide the best independent assessment of the CCL system.

Three conclusions are drawn from this work. Firstly that the long term repeatability of our gas analyser (0.04 ppmv) is slightly better than the apparent consistency of individual calibration gases (0.05 ppmv) currently obtained from the CCL. Secondly that the component of network wide imprecision in global CO₂ monitoring directly attributable to the CCL is about 0.06 ppmv. Thirdly that there exist a small number of "rogue" calibration gases whose apparent concentration on different analysers can vary by up to 0.5 ppmv relative to the majority of calibration gases, and overall calibration strategies should be designed to detect and reject such gases.

KEYWORDS

Atmospheric Composition
Baring Head
Calibrating
Carbon Dioxide
Gas Analysis
Least Squares Analysis
Infra-red analysis

Contents

1	Introduction	1
2	A model for NDIR Analysers	1
3	General Calibration Strategy	3
4	Calibration in the presence of drift	5
5	Least squares fitting of calibration data	8
6	Analysis of calibration gas consistency	10
7	Analysis of gases used with the URAS-2T analyser	11
8	Analysis of gases used with the Siemens analyser	16
9	Comparison of the results for the two analysers	19
10	Conclusions	20
11	References	22
	Appendix A. 1983–1985 URAS-2T Data	
	Appendix B. 1986–1987 Siemens Data	

1 Introduction

This report summarizes the procedures used to calibrate the non-dispersive infra-red (NDIR) gas analysers used in the New Zealand CO₂ monitoring program, and presents results of nearly all such calibrations for the January 1983 to April 1987 period.

This extensive set of calibration data involving two analysers, 45 calibration gases including both CO₂-in-nitrogen and CO₂-in-air types, and 140 independent analyser calibrations, allows us to separate the repeatability of analyser measurements from the apparent consistency of concentrations assigned to the calibration gases, by the World Meteorological Organization (WMO) Central Calibrating Laboratory (CCL).

We believe this constitutes the most extensive independent assessment of the CCL undertaken so far. In addition our estimates of long term repeatability of NDIR measurements are relevant to considerations of the network wide imprecision in current global CO₂ monitoring. In particular we are able to estimate the component of this imprecision that is due to the CCL itself, and to the inherent repeatability of NDIR analysers operated as in the New Zealand program.

It should be stressed that careful estimates of network imprecision are essential if atmospheric CO₂ data from different laboratories are to contribute to further understanding of the carbon cycle. The problems in using data from different laboratories are well known. For example Fraser et al (1983) estimated an imprecision of 1 ppmv for data taken from 4 organizations. Whereas in order to use atmospheric CO₂ concentrations in computer models to determine regional CO₂ sources and sinks to an accuracy of 1GtC, it is necessary to achieve an imprecision of 0.1 ppmv (Tans, private communication 1987).

The theoretical estimate of network imprecision arrived at in this work is not a substitute for more direct estimates obtained by intercomparison of measurements of the same gas mixtures. However, our analysis suggests that with sufficient use of calibration gases a network imprecision of 0.1 ppmv could be attained if inter-laboratory biases were reduced to the same level as the repeatability within a single laboratory.

This work is a significant extension of an earlier report (Manning and Pohl, 1986a) particularly as data from a new NDIR analyser show better repeatability than those reported previously. A preliminary draft of the present work was presented to a WMO/NBS meeting of experts on CO₂ measurements at the National Bureau of Standards (NBS), Gaithersburg, USA, in June 1987. The authors would like to acknowledge many useful comments received from participants at that meeting, and the support of the NBS which enabled one of us (MRM) to attend.

2 A model for NDIR Analysers

NDIR gas analysers of the type considered produce a voltage signal derived from the difference in the intensity of an infra-red beam transmitted through a gas in a sample cell and another gas in a reference (zero) cell. The same infra-red beam is passed through both cells and the transmitted part is measured by a detector selective to frequencies absorbed by CO₂. The gas in the reference cell is kept at a stable CO₂ concentration and the pressures in both cells are equalized so the resulting voltage output from the analyser can be related to the CO₂ concentration in the sample cell.

A simple approximation for the intensity of the infra-red beam transmitted through a fixed length cell containing gas at a fixed pressure is given by

$$I_{\text{transmitted}} = I_{\text{incident}} e^{-2\mu C}$$

where C is the gas CO_2 concentration and 2μ is an optical depth depending among other things on the cell length, gas pressure, and frequency spectrum of the infra-red source. (The constant factor 2 is introduced here for convenience later.) A more detailed model of infra-red absorption is necessary if one is to determine the amount of absorption from the specifications of the analyser (see e.g. Griffith et al, 1982). However, the simple approximation above, known as Beer's law, is sufficient for our purposes where parameters such as μ are determined by calibrating the analyser with gases of known concentration.

Using the formula for transmission intensity given above, the voltage output from the analyser can be written as

$$V(C) = V_R + f(e^{-2\mu C_R} - e^{-2\mu C}) \quad (1)$$

where C_R is the CO_2 concentration in the reference cell, C the concentration in the sample cell, and V_R and f contain signal offset and gain parameters.

It is convenient to refer concentrations to some standard concentration C_0 near the mid-point of the region of interest. (e.g. $C_0 = 340$ ppmv) and to work with relative concentrations

$$X = C - C_0 \quad (2)$$

in terms of which

$$V(C) = V_R + f e^{-2\mu C_0} (e^{-2\mu X_R} - e^{-2\mu X}).$$

In practice the μX terms are small (for the analysers described in this report typically ≤ 0.02) so this can be approximated adequately using the first three terms of an expansion of the exponentials as

$$V(C) \approx V_R + S [X(1 - \mu X) - X_R(1 - \mu X_R)] \quad (3)$$

$$\approx V_0 + SX(1 - \mu X). \quad (4)$$

The "sensitivity" S is defined as

$$S = 2\mu f e^{-2\mu C_0},$$

and V_0 is the analyser output at $X = 0$ ($C = C_0$).

From now on equation 4 is used as a model of the NDIR analyser response to a sample gas concentration C . It will also be convenient to use a "curvature corrected relative concentration"

$$G = X(1 - \mu X) \quad (5)$$

in terms of which the analyser response is linear. Note that the analyser response model used relates non-linearity directly to the optical depth.

The absorption of infra-red radiation by CO_2 molecules is modified, through a collision broadening effect, by the presence of certain types of other gas molecules in particular oxygen. The effect on the NDIR analyser is to change the effective optical depth. That is μ depends on the other major constituents of the carrier gas for the CO_2 .

If the same carrier gas (i.e. "natural" air) was always used, the optical depth parameter and hence both S and μ could be treated as constant. However in the NZ CO_2 program both CO_2 -in-nitrogen and CO_2 -in-air gases have been used and it is necessary to allow for different analyser response characteristics for the two types of gas. From equation 1 it can be seen that a change

$$\mu \rightarrow \mu + \delta\mu$$

corresponding to a change in carrier gas has exactly the same effect as scaling the corresponding CO₂ concentration by

$$C \rightarrow \left(1 + \frac{\delta\mu}{\mu}\right)C.$$

It is more convenient in practice to use this fact and work with concentrations adjusted by some factor depending on the nature of the carrier gas. In this way the same analyser response parameters S and μ can be used independent of the carrier gas.

3 General Calibration Strategy

The calibration strategy adopted so far in the NZ CO₂ monitoring program has been to use only CO₂-in-nitrogen gases for determining the analyser parameters S and μ (note that V_0 will not be used below). This is because of past concern about the consistency of CO₂-in-air gases.

A “carrier gas correction” is determined separately by measuring CO₂-in-air gases after a CO₂-in-nitrogen calibration scale has been defined. This carrier gas correction is allowed to be a general linear expression in concentration so that

$$C_A = C_N + (k_1 C_A + k_2) \quad (6)$$

where C_A is the concentration of the CO₂-in-air gas; C_N is the apparent concentration determined by the analyser output using the CO₂-in-nitrogen calibration scale; and the (positive) correction term in parentheses is the carrier gas correction. Note that k_2 is found to be significantly different from zero.

Before considering the determination of the various calibration parameters introduced above it is important to realize that the model used so far for NDIR analysers is an idealized one. A major complication in practice is that the analyser response to a fixed gas concentration is not steady in time. Rather it is subject to short term and long term drift.

The factors which cause the analyser output to drift are many and varied and difficult to quantify. They include:

- variations in water vapour concentrations in sample and reference cell (e.g. due to changes in cryogenic drying efficiency, or to back diffusion of water vapour from the gas cell outlet area to the cell itself);
- absorption of the infra-red beam by other trace gases and variability in the concentration of such trace gases;
- drift in electronic gain and voltage offset parameters (particularly through temperature dependence, and power supply fluctuations);
- variation in the infra-red source strength and its output frequency spectrum.
- variation in pressure and temperature of the two cells.

Our NDIR model suggests that effects which act on one cell more than the other (e.g. back diffusion of water vapour which will be greater for the reference cell as it is operated at lower flow rates) will cause drift in V_0 . However there are effects that potentially act on both cells equally (e.g. fluctuations in ambient pressure or infra-red source strength) and these will produce drift in the sensitivity S .

In order to take drift in the analyser response into account it is necessary to determine the calibration parameters repeatedly by fitting the model response given by equations 4 and 6 to gas mixtures of known CO₂ concentration. Ideally such calibrations should be done before and after measurement of a gas of unknown CO₂ concentration.

Our experience with NDIR analysers over many years suggests that both V_0 and S are subject to significant drift over periods of an hour or so. In particular the change in predicted atmospheric CO₂ concentration due to drift in these parameters is often of the order of 0.1 ppmv/hour for the Hartmann and Braun URAS-1 and URAS-2T analysers used in the NZ program. The Siemens Ultramat 3 analyser now in use has much lower drift rates, but drift in V_0 is still allowed for in calibration procedures.

It has been found that changes in the curvature parameter μ are of much less significance at least for the URAS-2T and Siemens analysers. The curvature term contributes only of the order of 0.3 ppmv to calculated concentrations and it is difficult to separate drift in this term from random noise in the analyser output. Thus average values of μ are determined using many separate calibrations. However as μ is liable to change whenever major changes, such as adjustment of flow rates or of infra-red beam stops, are made to the analyser, values for μ are not averaged across such changes.

The carrier gas correction parameters k_1 and k_2 in equation 6 are also stable over long periods, and average values for these are used with the same approach to averaging as used for μ . An exception to this generally observed stability in carrier gas effect occurred in the 1981-1982 period for our URAS-2T analyser when the carrier gas correction drifted significantly and beyond our control despite replacement of most of the critical optical and electronic components of the analyser. This anomalous period has been documented separately (Manning and Pohl, 1986b).

Particular attention has been paid to correcting for analyser drift in the NZ program. This is required as the baseline data to be measured varies by little more than 1 ppmv per year and determination of inter-annual changes in the shape of the annual cycle require stability to better than 0.1 ppmv over a few years. On the shorter time scale, baseline data taken during an extensive southerly wind event may vary by less than 0.1 ppmv over 48 hours. In order to determine whether there are systematic trends during this time, stability to considerably better than 0.1 ppmv is required over a few days.

In order to take all the above factors into consideration it is necessary to calibrate the analyser over several different time scales. The most frequent calibrations use 1 to 3 calibration gases and track variations in the S and V_0 parameters from hour to hour and day to day. These are called minor calibrations.

As an example of minor calibration strategy, the URAS-2T analyser measured one gas of known concentration every hour and 3 known gases every 24 hours. The Siemens analyser is controlled by a microprocessor which monitors the steadiness of sampled air and automatically schedules minor calibrations using 3 calibration gases every 6 hours during periods of steady atmospheric CO₂ concentration. At other times minor calibrations using 1 calibration gas are run every 2 hours and using 3 calibration gases every 30 hours.

Minor calibrations involving up to 3 calibration gases are not adequate for estimating the curvature correction parameter μ , or the carrier gas correction parameters k_1 and k_2 . These are obtained from major calibrations, using typically 6 to 20 calibration gases, which are run every few weeks. A major calibration can take from 4 to 24 hours to run.

The standard procedure with the URAS-2T was to run a major calibration approximately every month, although this was not always possible. When there was evidence of particular

problems such as during the change in the carrier gas correction in 1981–1982, major calibrations were run much more frequently. For the Siemens analyser the microprocessor schedules major calibrations of up to 11 calibration gases at preset intervals, and typically these are run 2 to 4 times per month. In the 1986–1987 period covered below major calibrations were run with higher frequency to obtain a large set of calibration data quickly.

Major calibrations are also used to monitor the calibration gases themselves. The calibration curve determined in any one major calibration run is a best fit in a least squares sense (to be detailed below) and does not give an exact fit to each of the individual calibration gases. Over a number of major calibrations the deviation of an individual calibration gas from the calibration curve can be monitored to produce both an average offset value and an analysis of its variation from one calibration to another.

Typically individual gases lie at about the same distance from the calibration curve whenever they are used. This indicates that the repeatability of the analyser measurement is better than the apparent consistency of the concentrations assigned to the calibration gases.

However the position of the calibration curve and to a lesser extent its slope and curvature vary with the set of calibration gases used to determine it. In section 6 procedures for removing this systematic variation in position of the calibration curve are given. This then allows a better assessment of the consistency of the CO₂ concentration assigned to an individual calibration gas in terms of the concentrations assigned to all the others.

4 Calibration in the presence of drift

As already mentioned the parameters V_0 and S both drift in time. Drift over the short term (time scales of up to an hour) can be eliminated by always using analyser measurements in sets of three as

gas 1	with concentration	C_1	giving volts	V_1	at time	t_1
2		C_2		V_2	t_2	
1		C_1		V_3	t_3	

The NDIR analyser model then gives

$$V_1 = V_0(t_1) + S(t_1)X_1(1 - \mu X_1) \quad (7)$$

$$V_2 = V_0(t_2) + S(t_2)X_2(1 - \mu X_2) \quad (8)$$

$$V_3 = V_0(t_3) + S(t_3)X_1(1 - \mu X_1) \quad (9)$$

and this triplet of measurements is regarded as a comparison of gas 1 and gas 2 during the time interval t_1 to t_3 . Note however that the comparison is not symmetric in gas 1 and gas 2, and gas 1 will be referred to as the “base” gas for the comparison.

In general one calibration gas is used as the base gas for a series of comparisons with several other gases run as the “comparison gas” (gas 2 above) between two measurements of the base gas which may be up to 30 minutes apart. A major calibration consists of from 30 to 200 such comparisons, involving between 6 and 20 calibration gases. One or at most two of these gases are used as the base gas throughout the calibration.

Assuming that the drift in V_0 and S is approximately linear between t_1 and t_3 the value V corresponding to the base gas at time t_2 is given by

$$V_{13} = \frac{(t_3 - t_2)V_1 + (t_2 - t_1)V_3}{(t_3 - t_1)} \quad (10)$$

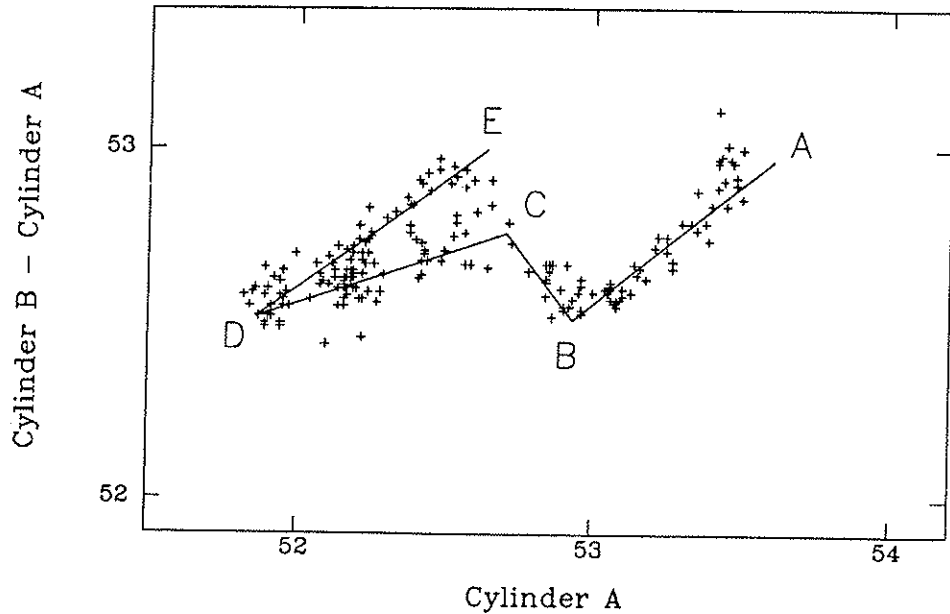


Figure 1: The difference in analyser output for two gas mixtures plotted against the output for the lower concentration mixture, showing short term sensitivity drift. Units on both axes are proportional to analyser voltage output. Data points are 6 minutes apart and the variation with time roughly follows the path shown as ABCDE.

and this value V_{13} corresponds to the same V_0 and S parameters that determine the value V_2 . Thus

$$V_2 = V_0(t_2) + S(t_2)X_2(1 - \mu X_2) \quad (11)$$

$$V_{13} = V_0(t_2) + S(t_2)X_1(1 - \mu X_1) \quad (12)$$

One way of eliminating much of the drift in the calibration scale is to eliminate V_0 directly by subtracting these equations giving

$$(V_2 - V_{13}) = S(t_2)(G_2 - G_1)$$

where the curvature corrected relative concentrations G are now used. This however still leaves a sensitivity parameter S which may be time dependent.

It is possible to improve the drift correction in cases where there is significant sensitivity drift by allowing for a correlation between V_0 and S . Examples of sensitivity drift are shown in Figures 1 and 2 which are typical of our general observation that where the sensitivity does drift it is quite highly correlated with drift in the baseline over time scales of the order of 10 hours.

Figure 1 shows data taken from a series of consecutive 6 minute measurements (run by mistake!) of two gases using a URAS-2T analyser over a period of about 40 hours starting on September 7th 1983. The variation with time roughly follows the path shown as ABCDE demonstrating a good correlation between S and V_0 over the 10 hour time scale. Note however that the sign of the correlation becomes negative for a short time in the region B-C.

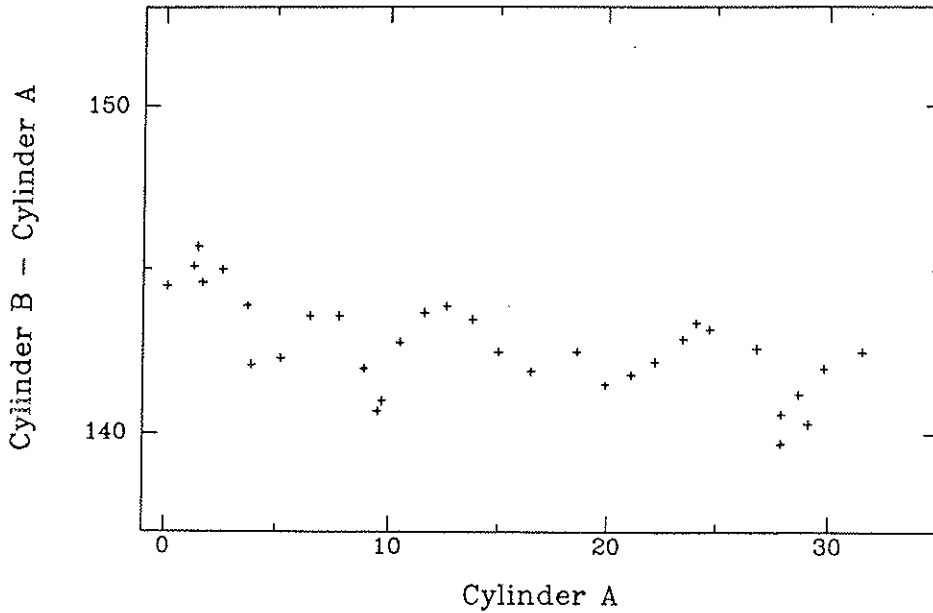


Figure 2: The difference in analyser output for two gas mixtures plotted against the output for the lower concentration mixture, showing sensitivity drift over 33 days. Units on both axes are proportional to analyser voltage output. Data points are 24 hours apart and the variation with time is strictly from left to right.

Figure 2 shows data from regular measurements made every 24 hours using a URAS-2T analyser over a 33 day period starting in July 1984. Although there is an overall correlation between S and V_0 in this data, it is clear that short and long term correlation coefficients are different even in sign. Consequently correlations between S and V_0 are not used over this time scale.

In order to exploit short term correlation between sensitivity drift and baseline drift the sensitivity is written as

$$S(t_2) = S_0 + \alpha V_0(t_2)$$

where α determines the amount of drift in S for unit drift in V_0 , and S_0 is to be taken as constant over the whole period of sensitivity drift. Substituting this form for S in the equations 12 and eliminating V_0 now leads to

$$V_2 = V_{13} + \left(\frac{S_0 + \alpha V_{13}}{1 + \alpha G_1} \right) (G_2 - G_1) \quad (13)$$

which reduces to the more familiar form if α is zero.

Note that where there is little or no drift in sensitivity the introduction of the extra parameter α cannot be justified as it has the undesirable side effect of interacting with other parameters being determined by the least squares fit to be detailed below. However for the URAS-2T analyser there were periods, particularly during major calibrations when this sensitivity drift model is clearly indicated by the data, and its use produces results that are more consistent from one major calibration to the next.

Before proceeding to the details of the least squares fitting procedure it should be noted that once the calibration parameters are known, a concentration is determined for an unknown

gas by using the same analyser model. A comparison of an unknown gas with a known one measures values for V_{13} and V_2 ; parameters S_0 , μ , and α are taken from adjacent calibrations; and the concentration X_1 assigned to the base gas is known. From these values G_1 is calculated and then G_2 using

$$G_1 = X_1(1 - \mu X_1) \quad (14)$$

$$G_2 = [V_2 - V_{13} + G_1(S_0 + \alpha V_2)] / (S_0 + \alpha V_{13}). \quad (15)$$

The relative concentration X_2 is then given as

$$X_2 = \frac{2G_2}{1 + \sqrt{1 - 4\mu}}. \quad (16)$$

Note that V_0 is not required.

5 Least squares fitting of calibration data

To determine S_0, μ, α etc of the analyser model, a least squares fit is carried out on data from a set of comparisons of gases of known concentration giving values $(X_2(i), V_2(i))$ and $(X_1(i), V_{13}(i))$. These data are taken over a period when it is reasonable to assume that the analyser parameters have remained constant.

The least squares method minimizes a sum of squares of residuals which have either equal statistical weight, or are explicitly weighted by the inverse of their expected variance. The method adopted here is to use residuals from equation 13 assuming that values of V output from the analyser have approximately the same level of instrumental noise independent of the concentration X . Strictly the statistical errors in equation 13 contain an effect due to noise in V_{13} , and there is a (positive) correlation between adjacent comparisons. Both these effects are small and are ignored here.

In doing a least squares fit to determine the analyser parameters, it is necessary to allow for "errors" in the concentrations assigned to a calibration gas. In fact as already indicated the repeatability of our analyser measurements is better than the apparent consistency of the calibrations, thus much of the deviation of a data point from the calibration curve should be treated as an "error" in the assigned concentration.

Furthermore if calibration gases are treated as having rigidly assigned concentrations the position of the calibration curve becomes less stable where two calibration gases of similar concentration "disagree" and pull the fitted curve in opposite directions. The position of the curve is liable to flip between fitting one or the other of these competing calibration gases on adjacent major calibrations.

Consequently the assigned concentration is treated as having a statistical error, and the calibration curve is estimated by minimizing a weighted sum of squared residuals from the analyser output model, plus a sum of squares of the inferred errors in the assigned concentrations. Different weighting is used for the two types of residual, giving more weight to the assigned concentration for the calibration gas than to individual analyser measurements.

Note carefully that this procedure does not constitute a re-assignment of concentrations for the calibration gases. The CCL assigned concentrations for calibration gases are not changed, and the least squares fit for each independent calibration always starts with the CCL values. A local re-assignment of calibration gas concentrations would require the passing of results from one analyser calibration to another in some way, and this is not done in the procedures described here.

In summary, analyser parameters S_0 , α and μ , together with inferred errors d_i in assigned concentrations, are determined as the values which minimize

$$Z = \sum_j \left(\frac{\xi_j}{\sigma_r} \right)^2 + \sum_i \left(\frac{d_i}{\sigma_d} \right)^2;$$

where the residual ξ_j for the j 'th comparison is

$$\xi_j = V_{2j} - V_{13j} - \left(\frac{S_0 + \alpha V_{13j}}{1 + \alpha G_{1j}} \right) (G_{2j} - G_{1j});$$

where the curvature corrected relative concentrations are given by

$$G = X(1 - \mu X);$$

and where the relative concentrations X are related to assigned concentrations C by

$$X = C - d_i - C_0.$$

The weighting factors σ_r and σ_d are error estimates for the two types of residuals. Typically we use $\sigma_r = 0.25$ ppmv and $\sigma_d = 0.1$ ppmv, although only the ratio of these is significant.

The presence of αG terms in ξ_j makes the minimization of Z a non-linear least squares problem. However the degree of non-linearity is usually quite small and a good first approximation for the parameters can be obtained by a linear least squares fit in which α , μ and all the d_i are set to zero.

Prior to 1985 all the major calibration periods were scanned manually before a least squares fit was done, and individual measurements for a gas were rejected where there was evidence that these were inconsistent with better measurements of the same gas on the same day. Typically hysteresis or "settling down" effects are seen in which one or two early measurements of a particular gas are significantly lower or higher than later measurements.

From 1985 onwards rejection of some initial measurements for a particular gas, has been largely automated with an automatic outlier rejection procedure. This is done first by calculating a simple linear least squares fit to all the calibration gas data in which μ , α and all the d_i are set to zero. Then for each calibration gas in turn, an exponentially weighted mean deviation from the fitted calibration line is calculated. This exponential weighting favours later points over early ones. Individual measurements for the calibration gas are then rejected where the corresponding deviation from the calibration line differs by more than a preset amount (typically 0.05 ppmv) from the weighted mean. Various safeguards are built into this procedure to avoid discarding too much data.

The results of processing a major calibration are estimates of the parameters S_0 , μ , α and the d_i . In order to do further analyses of the consistency of gas concentrations the average deviations between the calibration curve and the assigned concentration for each gas are stored. These average deviations are calculated by using equations 15 and 16 to calculate an average over all comparisons for the concentration of a particular calibration gas and then taking the difference between this and the CCL assigned concentration. The sense of the average deviations denoted by E below is the same as for the d_i above, viz

$$E \sim (\text{CCL assigned value} - \text{Local estimated value})$$

The analysis of average deviation values from a number of major calibrations is discussed in the next section.

6 Analysis of calibration gas consistency

This section presents procedures for the statistical analysis of deviations from the calibration curve of individual calibration gases used in a major calibration. But first it is important to be clear as to what this deviation represents.

The deviation from the calibration curve is the result of four separate factors:

1. the error in measurement made with our analyser, *big offsets*.
2. the error in the concentration assigned by the CCL
3. drift in the concentration of the calibration gas between the time measured at the CCL and on our analyser
4. an inherent difference in the response of our analyser compared to that used by the CCL for some calibration gases.

Much of the motivation for regular analyses of the deviations of calibration gases from the calibration curves is to track any drift in concentration, the third factor. For example in Manning and Pohl, 1986b (figure 1), we reported drift of more than 1 ppmv in a calibration gas, and further examples of drift can be seen in section 7 of this report. In our experience concentration drift is most likely for gases contained in steel cylinders and is only likely when cylinder pressures fall below 600 psi.

Results given below suggest that variability between analysers, the fourth factor, is significant for a few "rogue" gases, which have large deviations from calibration curves and show different effective concentrations on the two analysers used in this study. Detection and elimination of such "rogue" gases is important in an overall calibration strategy.

In order to analyse gas deviations over a number of major calibrations a procedure that is essentially a two-way analysis of variance is used. The deviations $E_{t,i}$ can be considered as a matrix of values labelled by the date t of the major calibration and by the gas i , that produced it. These deviations are separated into a time dependent factor F_t and a gas dependent factor D_i leaving a residual error term $e_{t,i}$ so that

$$E_{t,i} = F_t + D_i + e_{t,i}.$$

The factors F_t and D_i are obtained by minimizing the sum of squares

$$\sum_i (e_{t,i})^2 = \sum_i (E_{t,i} - F_t - D_i)^2$$

subject to the constraint that the gas dependent factors produce no nett shift in the concentration scale, i.e

$$\sum_i D_i = 0.$$

The D_i will be referred to as offsets.

This separation into factors is a form of generalized least squares analysis similar to analysis of variance but special care is needed in the algorithm used, to allow for the sparsity of the data matrix E . If the data matrix has no missing values (i.e. there is a measured $E_{t,i}$ for every t and i) then the D_i and F_t are simply related to row and column means.

A FORTRAN program is used to analyse the data presented here, and this has been checked by an independent analysis (Rhoades, 1984) using the GENSTAT statistical package.

An alternative technique known as "median polish" has also been considered, but although this is generally regarded as being more resistant to outlier data we prefer the approach above which produces averages rather than medians.

When such an analysis of variance is carried out for the deviations of *CO₂-in-nitrogen* gases, the F_i terms show small shifts in the position of the calibration curve from one calibration to another and have an average value close to zero. The D_i values provide a measure of the average offset of one individual *CO₂-in-nitrogen* gas from the calibration scale determined by the entire set. They provide a better estimate of the consistency of the calibration gases than individual deviations $E_{t,i}$ as they are in effect averaged over many different calibrations usually involving different sets of calibration gases.

The residuals $e_{t,i}$ having had any systematic gas dependent and time dependent terms removed tend to show the presence of drift in individual gases, or anomalous individual measurements, more readily than do the $E_{t,i}$ values.

In the case of *CO₂-in-air* gases the deviations $E_{t,i}$ and offsets D_i are measured from the calibration curve based on *CO₂-in-nitrogen* gases. The F_i factor gives an estimate of the mean carrier gas correction term for each calibration date, and the D_i give the variation of the carrier gas correction about this mean for each individual gas as well as systematic effects associated with the gas.

Following the definition of the carrier gas correction in equation 6, estimates of $(C_A - C_N)$ for the i 'th calibration gas after pooling results of all calibrations are given by

$$C_{A,i} - C_{N,i} = D_i + \langle F \rangle$$

where $\langle F \rangle$ is the mean value of F_i for the calibrations (the mean value of D_i is constrained to be zero). Thus estimates of the carrier gas correction coefficients k_1 and k_2 are obtained by a least squares fit line

$$D_i + \langle F \rangle = k_1 C_A + k_2 + r_i \quad (17)$$

where the r_i are residuals minimized in the fit. For *CO₂-in-air* gases the residuals r_i are used to assess consistency rather than the D_i .

7 Analysis of gases used with the URAS-2T analyser

In the 1983-1985 period 28 independent calibrations were carried out for the URAS-2T analyser involving 24 *CO₂-in-nitrogen* and 15 *CO₂-in-air* calibration gases at various times. After some data selection a total of 324 deviation values $E_{t,i}$ are obtained. Details of the calibration gases are given in tables 1 and 2, and in figures 3 and 4. The values for the deviations $E_{t,i}$, together with the separation into time dependent and gas dependent factors is given in the computer output in appendix A.

From table 1 and figure 3 it can be seen that all 5 of the *CO₂-in-nitrogen* gases whose pressure dropped below 400 psi showed an upwards drift in concentration. The greatest change in concentration measured by the CCL is 2.59 ppmv for cylinder 35355. It has been shown previously (Manning and Pohl, 1986b) in cases of large drift that the phenomenon is not simply linear in time or in pressure. Rather that the concentration remains constant till some critical pressure is reached, and then the concentration rises linearly in time. The behaviour of all 5 drifting cylinders shown here is consistent with this pattern.

The data analysis procedures used in the NZ program normally assume linear drift in time for each calibration gas between CCL calibrations, although the effect is usually not

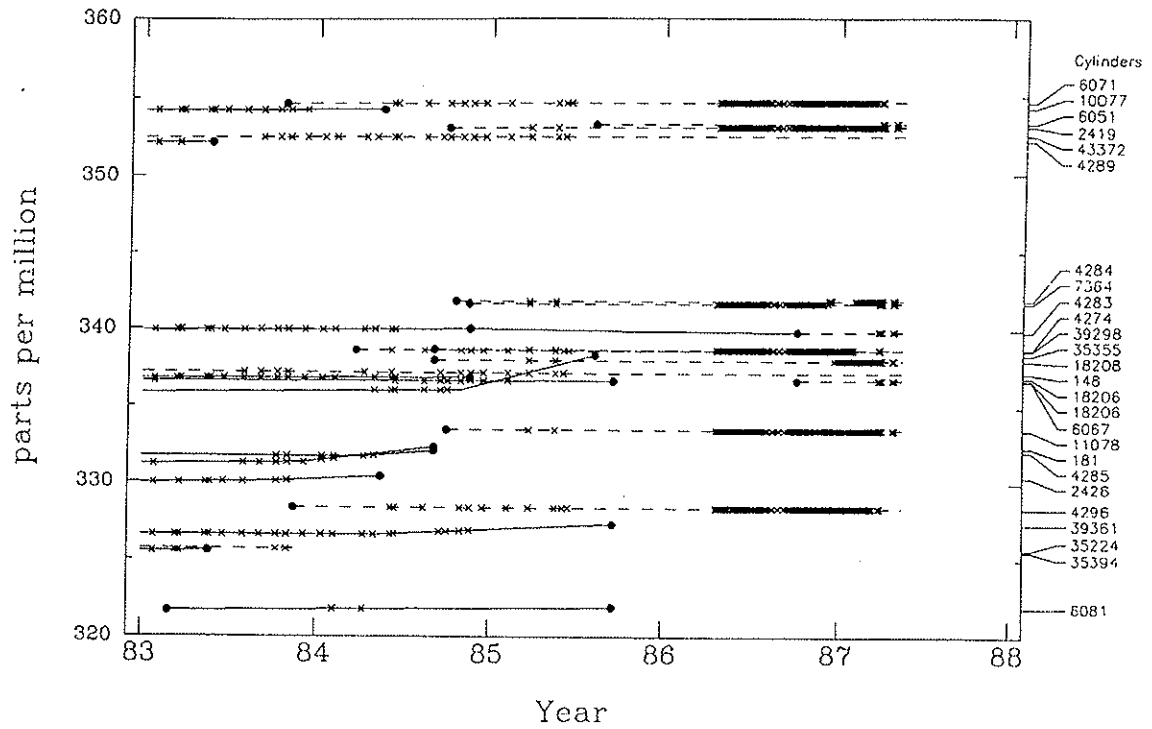


Figure 3: CO₂-in-nitrogen calibration gas concentrations and calibration times. Solid dots denote CCL calibrations, solid lines are drawn between CCL calibrations, and dotted lines are drawn where a gas has no closing CCL calibration. Crosses denote where a gas is used in a calibration of the New Zealand analyser.

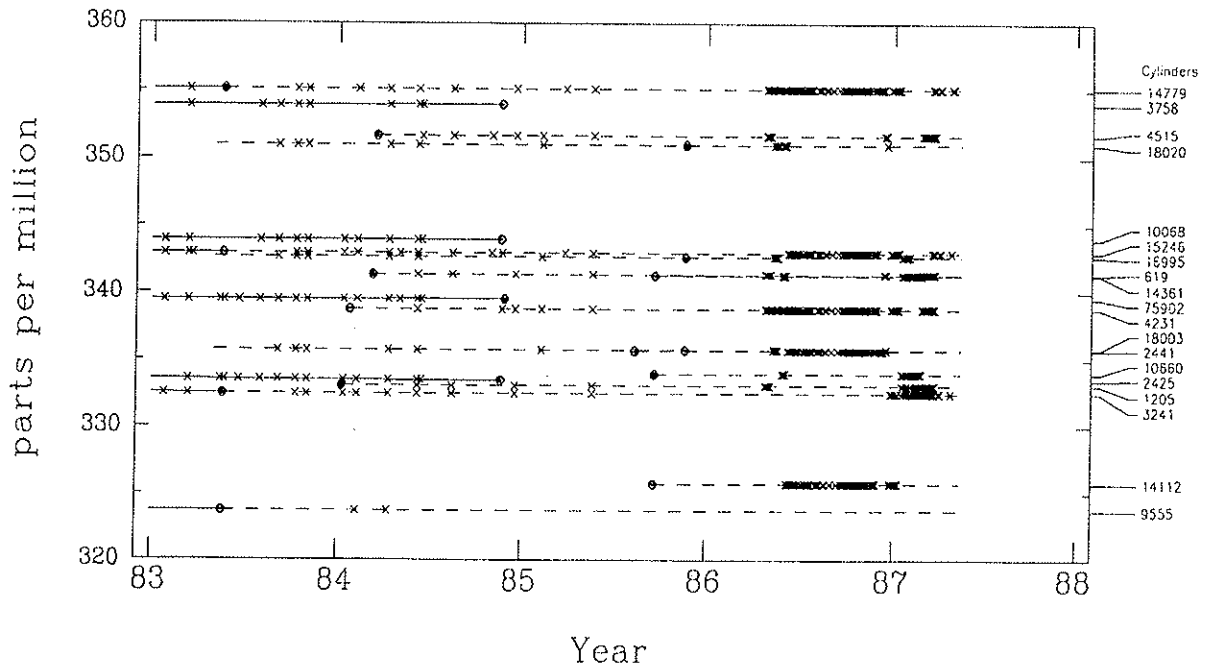


Figure 4: CO₂-in-air calibration gas concentrations and calibration times. See figure 3 for details.

Table 1: CO₂-in-nitrogen calibration gases used with the URAS-2T analyser in the 1983–1985 period. The definition of offset values and a discussion of outliers marked with *, is given in the text.

Cylinder No	First Calibration		Second Calibration		Lowest pressure (psi)	Offset D_i (ppmv)	
	Conc (ppmv)	Date	Conc (ppmv)	Date			
6081	321.69	Oct 1981	321.93	Sep 1985	1660	-0.08	
35394	325.13	May 1980	325.58	May 1983	700	+0.05	
35224	325.68	Apr 1981			1160	-0.02	
39361	326.60	Mar 1982	327.32	Sep 1985	380	-0.07	
4296	328.36	Nov 1983			1750	+0.01	
2426	329.92	Jun 1982	330.43	May 1984	380	-0.09	
181	331.13	Jun 1982	332.37	Sep 1984	380	-0.10	
4285	331.73	Jun 1982	332.10	Sep 1984	340	-0.13	
11078	333.49	Oct 1984			1580	+0.25 *	
35355	335.80	Jun 1982	338.39	Sep 1985	< 400	-0.08	
6067	336.60	Jun 1982	336.72	Sep 1985	520	+0.05	
18206	336.76	Apr 1982	336.91	Nov 1984	1790	+0.00	
148	337.22	Jun 1982			700	-0.18	
18208	338.04	Sep 1984			1910	+0.17 *	
39298	338.67	Mar 1984			1940	-0.01	
4274	338.73	Sep 1984			1950	+0.20 *	
4283	339.89	Apr 1982	340.08	Nov 1984	1220	-0.06	
7364	341.74	Nov 1984			1900	+0.09	
4284	341.91	Oct 1984			1900	+0.18 *	
4289	352.25	Aug 1980	352.18	May 1983	930	-0.08	
43372	352.51	Sep 1982			570	-0.07	
2419	353.09	Oct 1984			1910	+0.02	
10077	354.21	Mar 1981	354.26	May 1984	540	-0.07	
6071	354.64	Oct 1983			1810	+0.00	
mean and standard deviation for all offsets						0.00	± 0.11
mean and standard deviation excluding outliers						-0.04	± 0.06

significant. However, in cases where there is large drift the non-linearity must be taken into account. As described in Manning and Pohl (1986b) in more detail for cylinder 181, this is done by using the NZ calibrations to assign concentrations to the drifting gas. A point in time can be determined at which the rate of drift appears to increase markedly, and a local calibration of the gas at that time is assigned from the NZ analyser data.

Although three of the gas mixtures used are NBS standards, viz cylinders 18003 (SRM 1670), 16995 (SRM 1671) and 18020 (SRM 1672), all concentrations used are from measurements by the CCL.

Figures 5 and 6 show plots of the gas offsets D_i against concentration for the CO₂-in-nitrogen and CO₂-in-air gases respectively. In both figures an indication of the repeatability of individual measurements is given by showing as error bars the standard deviations of the $e_{i,i}$ values for each cylinder.

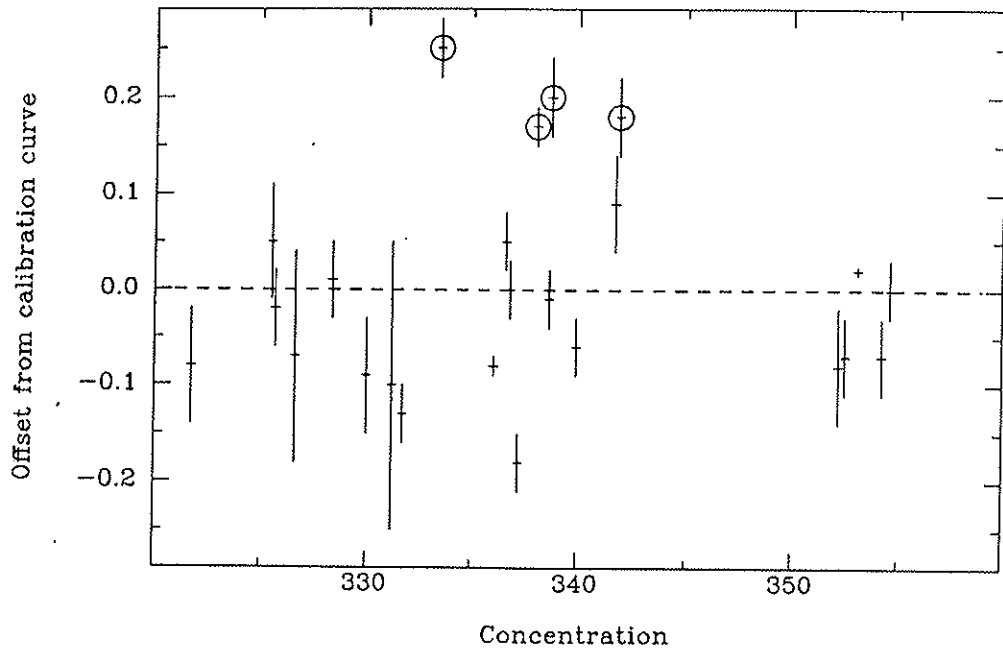


Figure 5: Calibration gas offsets for CO_2 -in-nitrogen gases measured with the URAS-2T analyser between 1983 and 1985. The outlier points discussed in the text are ringed.

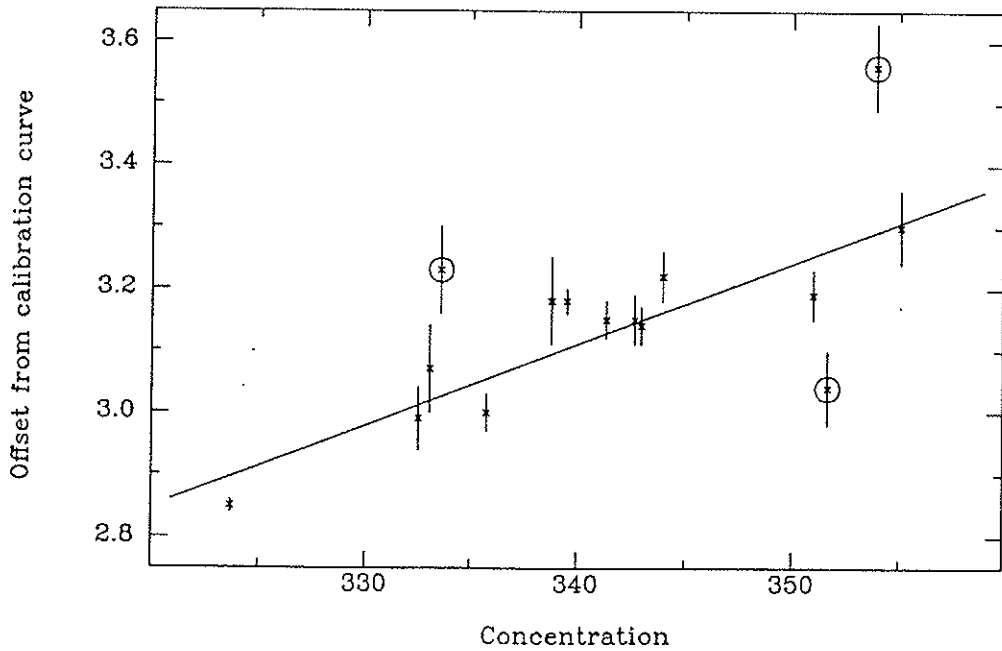


Figure 6: Calibration gas offsets for CO_2 -in-air gases measured with the URAS-2T analyser between 1983 and 1985. The outlier points discussed in the text are ringed.

Table 2: CO₂-in-air calibration gases used with the URAS-2T analyser in the 1983-1985 period. The definition of offset values and a discussion of outliers marked with *, is given in the text.

Cylinder No	First Calibration		Second Calibration		Lowest pressure (psi)	Offset $D_i + \langle F \rangle$ (ppmv)	Residual r_i (ppmv)	
	Conc (ppmv)	Date	Conc (ppmv)	Date				
9555	323.83	Jun 1981	323.71	May 1983	1770	2.85	-0.05	
3241	332.64	Jun 1981	332.50	May 1983	1750	2.99	-0.03	
1205	333.08	Jan 1984			1820	3.07	+0.04	
2425	333.62	Jul 1982	333.50	Nov 1984	2100	3.23	+0.20 *	
18003	335.76	Nov 1985			1810	3.00	-0.06	
4231	338.79	Jan 1984			1860	3.18	+0.08	
75902	339.49	Jul 1982	339.54	Nov 1984	2050	3.18	+0.07	
619	341.37	Mar 1984			1920	3.15	+0.01	
16995	342.67	Nov 1985			1840	3.15	+0.00	
15246	343.05	Jun 1981	342.96	May 1983	1750	3.14	-0.02	
10068	343.97	Jul 1982	343.96	Nov 1984	2090	3.22	+0.05	
18020	350.96	Nov 1985			1990	3.19	-0.07	
4515	351.65	Mar 1984			1840	3.04	-0.23 *	
3758	353.90	Jul 1982	353.96	Nov 1984	1830	3.56	+0.26 *	
14779	355.11	May 1981	355.09	May 1983	1660	3.30	-0.02	
mean and standard deviation for all residuals							0.02	± 0.12
mean and standard deviation excluding outliers							0.00	± 0.05

Figure 6 shows a least squares fit line for the CO₂-in-air gas offsets D_i after rejection of the 3 furthest outliers (cylinders 2425, 4515 and 3758). The corresponding carrier gas correction (cf equation 17) is given by

$$k_1 = 0.0132 \pm 0.0019$$

$$k_2 = -1.38 \pm 0.66 \text{ ppmv.}$$

The repeatability of the analyser measurements can be summarized by the standard deviation of the $e_{i,i}$. For the CO₂-in-nitrogen gases this is 0.03 ppmv and for CO₂-in-air is 0.04 ppmv. (Note these values are smaller than suggested at first sight by the error bars shown in figures 5 and 6. This is because the overall value weights the errors shown for each calibration gas by the number of measurements made for that gas, and the gases which individually show poor repeatability tend to have fewer measurements.)

The consistency of the 24 CO₂-in-nitrogen calibration gases based on these measurements with the URAS-2T analyser can be summarized by the standard deviation of the offsets D_i . This is 0.11 ppmv, however, is significantly influenced by a few outliers. In fact 57% of the variance of the D_i comes from the 4 values marked as outliers in table 1. Removing these reduces the standard deviation of the cylinder offsets to 0.05 ppmv.

For the 15 CO₂-in-air calibration gases the consistency is estimated from the standard deviation of the residuals r_i after fitting a carrier gas correction line to the offsets $D_i + \langle F \rangle$. This standard deviation is 0.12 ppmv but again is strongly influenced by the 3 outliers marked in table 2 which account for 75% of the variance of the residuals. Removing these three outliers reduces the standard deviation to 0.05 ppmv.

Summarizing these results for the URAS-2T analyser, the repeatability of analyser measurements is estimated at 0.04 ppmv. The apparent consistency in the calibration gases

is strongly influenced by a few anomalous gases. Including these gives a consistency of 0.12 ppmv, whereas excluding them improves the figure to 0.06 ppmv.

8 Analysis of gases used with the Siemens analyser

A Siemens Ultramat 3 analyser was installed at Baring Head in the latter part of 1985, and replaced the URAS-2T from that time. As the Baring Head station is not manned and only visited once a week, the opportunity was taken to increase the amount of automatic control of the NDIR analyser system.

The new analyser and an inlet manifold with 11 on-line calibration gases is controlled by a Digital Equipment Corporation LSI-11/2 microprocessor and software developed by the authors. In addition a cryocooler air drying unit is used which switches sampled air regularly between two twin traps, heating the off-duty trap to purge it of ice and water. As the Siemens analyser is less sensitive to water vapour than the URAS-2T and as calibration gases have very low dew points, the calibration gases for the Siemens analyser are not passed through the cryogenic drying system.

The computer control program monitors analyser voltage every second and when calibration gases are being used, switches to the next gas as soon as 60 seconds of stable data are obtained. In practice this allows a saving in gas usage. Note that each calibration gas is measured over a number of 60 second "jogs" separated by measurements of other gases as described in section 4. The data analysis procedures described in section 5 monitor variation between individual measurements of the same gas and reject outliers.

Calibration gases and sample airlines are selected by the control program according to a number of predetermined sequences. The choice of sequence is based on a combination of fixed scheduling and dynamic adaptation to the ambient CO₂ concentrations. It has been mentioned that this provides the ability to run more calibration gases during steady CO₂ concentration periods. However the ability to schedule cycles that run all 11 gases at regular intervals also provides the ability to accumulate large amounts of calibration data very easily. As a result this report contains much more data from the 1986-1987 period than for the 1983-1985 period.

In the 1986-1987 period 117 independent calibrations were carried out for the Siemens analyser involving 11 CO₂-in-nitrogen and 14 CO₂-in-air calibration gases at various times. After some data selection a total of 1253 deviation values $E_{t,i}$ are obtained. Details of the calibration gases are given in tables 3 and 4, and in figures 3 and 4. The values for the deviations $E_{t,i}$ together with the separation into time-dependent and gas dependent factors are given in the computer output in appendix B.

Note that none of the calibration gases used in the 1986-1987 period have yet had a CCL calibration after use.

Figures 7 and 8 show plots of the gas offsets D_i against concentration for the CO₂-in-nitrogen and CO₂-in-air gases respectively with the same conventions as used for figures 5 and 6 above. Once again the 3 furthest outliers (cylinders 1205, 619 and 4515) have been discarded in determining the least squares line shown in figure 8. The corresponding carrier gas correction term is given by

$$k_1 = 0.0128 \pm 0.0025$$

$$k_2 = -1.506 \pm 0.51 \text{ ppmv}$$

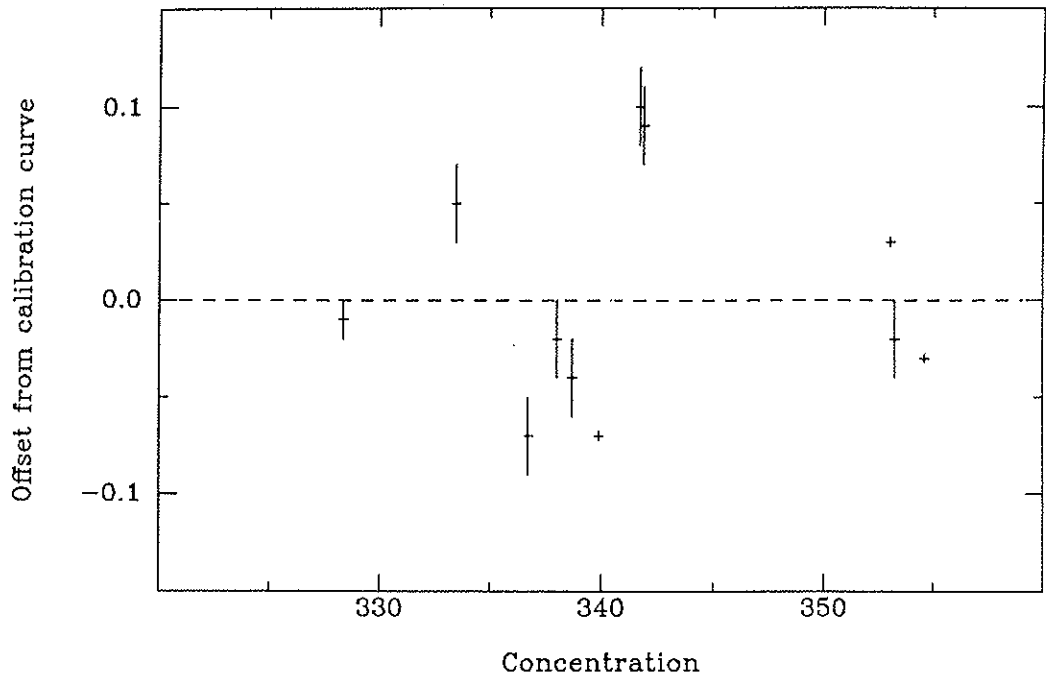


Figure 7: Calibration gas offsets for CO_2 -in-nitrogen gases measured with the Siemens analyser between 1986 and 1987.

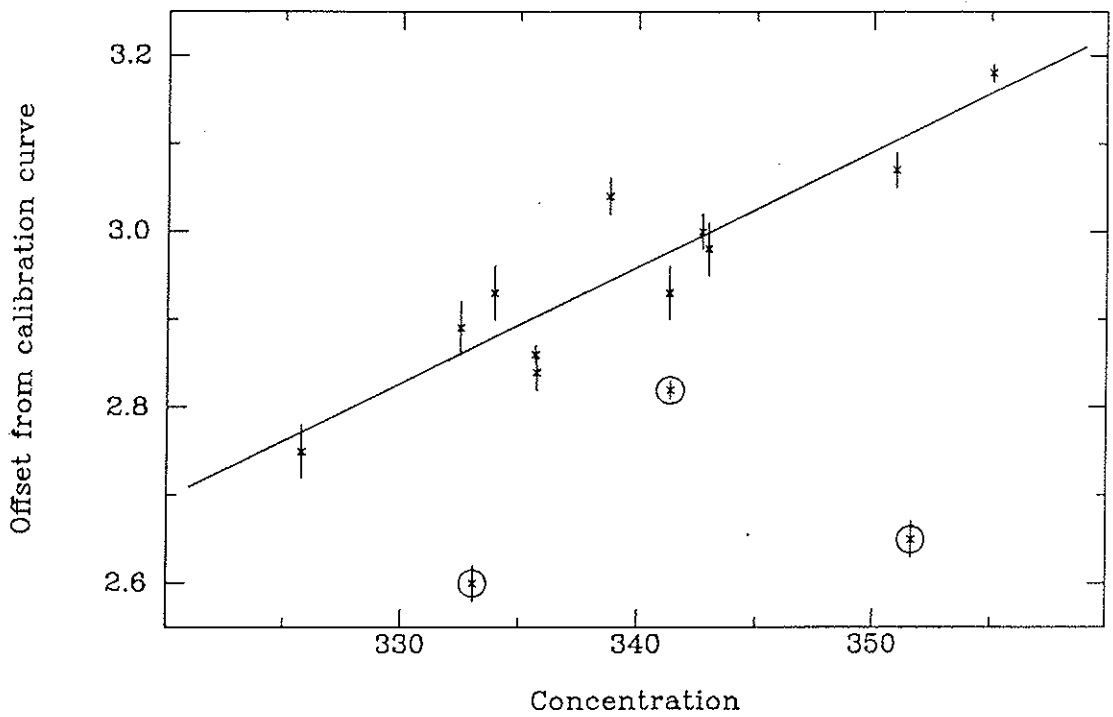


Figure 8: Calibration gas offsets for CO_2 -in-air gases measured with the Siemens analyser between 1986 and 1987. The outlier points discussed in the text are ringed.

Table 3: CO₂-in-nitrogen calibration gases used with the Siemens analyser in the 1986-1987 period. The definition of offset values is given in the text.

Cylinder No	First Calibration		Lowest pressure (psi)	Offset D_i (ppmv)
	Conc (ppmv)	Date		
4296	328.36	Nov1983	460	-0.01
11078	333.49	Oct1984	1400	+0.05
18206	336.74	Oct1986	1750	-0.07
18208	338.04	Sep1984	1380	-0.02
4274	338.73	Sep1984	470	-0.04
4283	339.90	Oct1986	1400	-0.07
7364	341.74	Nov1984	1750	+0.10
4284	341.91	Oct1984	1860	+0.09
2419	353.09	Oct1984	1800	+0.03
6051	353.20	Oct1986	1980	-0.02
6071	354.64	Oct1983	620	-0.03
mean and standard deviation for all offsets				0.00 ± 0.06

Table 4: CO₂-in-air calibration gases used with the Siemens analyser in the 1986-1987 period. The definition of offset values and discussion of the outliers marked with * is given in the text.

Cylinder No	First Calibration		Lowest pressure (psi)	Offset $D_i + \langle F \rangle$ (ppmv)	Residual r_i (ppmv)
	Conc (ppmv)	Date			
14112	325.75	Sep 1985	2070	2.75	-0.02
3241	332.50	May1983	1580	2.89	+0.03
1205	333.08	Jan 1984	1680	2.60	-0.27 *
10660	333.94	Sep 1985	2020	2.93	+0.05
2441	335.69	Sep 1985	2010	2.86	-0.04
18003	335.76	Nov1985	1550	2.84	-0.06
4231	338.79	Jan 1984	1650	3.04	+0.10
14361	341.31	Sep 1985	1780	2.93	-0.05
619	341.37	Feb 1984	1840	2.82	-0.16 *
16995	342.67	Nov1985	1610	3.00	+0.01
15246	342.96	May1983	1590	2.98	-0.02
18020	350.96	Nov1985	1780	3.07	-0.03
4515	351.65	Mar1984	1930	2.65	-0.46 *
14779	355.09	May1983	1230	3.18	+0.02
mean and standard deviation for all residuals					-0.06 ± 0.15
mean and standard deviation excluding outliers					0.00 ± 0.05

The repeatability of the analyser measurements can be summarized by the standard deviation of the $e_{t,i}$ residuals as before. For both the 11 CO₂-in-nitrogen gases and the 14 CO₂-in-air gases this figure is 0.02 ppmv.

The consistency of the 11 CO₂-in-nitrogen calibration gases based on these measurements with the Siemens analyser can be summarized by the standard deviation of the offsets D_i . This value is 0.06 ppmv. For the 14 CO₂-in-air gases the consistency estimated as the standard deviation of the residuals r_i is 0.15 ppmv, but once again this figure is strongly influenced by a few outliers. The 3 outliers marked in table 4 account for 73% of the variance of the residuals and removing these reduces the standard deviation to 0.05 ppmv.

Summarizing these results for the Siemens analyser, the estimated repeatability is about

0.02 ppmv and the apparent consistency of the calibration gas concentrations is about 0.05 ppmv after removal of obvious outliers.

9 Comparison of the results for the two analysers

This section gives a comparison of the results for the gases common to calibrations of both analysers and provides a measure of the repeatability of the NZ NDIR calibration scale across two analysers.

The 10 CO_2 -*in-nitrogen* calibration gases common to both sets of analyser calibrations are shown in table 5 with summary statistics of the difference between the two sets of offsets. There are clearly three anomalous gases (11078, 18208 and 4274) which appear high on the URAS-2T analyser, and these were identified earlier as outliers in the URAS-2T data. For the remaining 7 gases the difference between offsets on the two analysers is just $0.03 \text{ ppmv} \pm 0.04 \text{ ppmv}$.

Table 5: Comparison of offsets of CO_2 -*in-nitrogen* gases measured with two analysers. Outliers marked with * are discussed in the text.

Cylinder No	URAS-2T offset	Siemens offset	Difference (Siemens - URAS-2T) in offset	
4296	+0.01	-0.01	-0.02	
11078	+0.25	+0.5	-0.20	*
18206	+0.00	-0.07	-0.07	
18208	+0.17	-0.02	-0.19	*
4274	+0.20	-0.04	-0.24	*
4283	-0.06	-0.07	-0.01	
7364	+0.09	+0.10	+0.01	
4284	+0.18	+0.09	-0.09	
2419	+0.02	+0.03	+0.01	
6071	+0.00	-0.03	-0.03	
mean and standard deviation for all differences			-0.08	± 0.09
mean and standard deviation excluding outliers			-0.03	± 0.04

It should be noted that 6 of the gases shown in table 5 and all 3 of the anomalous ones were measured only twice with the URAS-2T analyser.

Table 6 shows the 10 CO_2 -*in-air* gases common to both sets of analyser calibrations and a comparison of the residuals r_i after fitting the (different) carrier gas correction lines to the cylinder offsets. Again there are three anomalous gases (1205, 619, and 4515) which now appear low on the Siemens analyser and these gases have already been identified as outliers from the analyser calibration curves. Removing these 3 gases reduces the difference between the two sets of residuals to a value consistent with the repeatability of the individual analysers.

Each of the 6 calibration gases identified in tables 5 and 6 as having an anomalously large difference in effective concentration between the two analysers, has previously been identified as an outlier from the calibration curve for one or other of the analysers. Furthermore there is only one gas (4284) identified as inconsistent on a single analyser which did not show a large difference in apparent concentration between analysers. Of course the inconsistency of a calibration gas relative to others on a single analyser can be interpreted as a variation in effective concentration between the CCL analyser and the local one. The variation in

Table 6: Comparison of residuals of CO₂-in-air gases measured with two analysers. Outliers marked with * are discussed in the text.

Cylinder No	URAS-2T residual	Siemens residual	Difference (Siemens - URAS-2T) in residual	
3241	-0.03	+0.03	0.06	
1205	+0.04	-0.27	-0.31	*
18003	-0.06	-0.06	0.00	
4231	+0.08	+0.10	+0.02	
619	+0.01	-0.16	-0.17	*
16995	0.00	+0.01	+0.01	
15246	-0.02	-0.02	0.00	
18020	-0.07	-0.03	+0.04	
4515	-0.23	-0.46	-0.23	*
14779	-0.02	+0.02	+0.04	
mean and standard deviation for all differences			-0.05	± 0.13
mean and standard deviation excluding outliers			+0.02	± 0.02

effective concentration with analyser appears to be as large as 0.5 ppmv (4525 between the CCL analyser and the NZ Siemens analyser).

The analysis presented here therefore consistently identifies a small number (6 out of 45) of calibration gases as showing different effective concentrations on different analysers with no clear pattern. Gases with this type of behaviour can not be used for intercalibration purposes, and will be referred to as “rogues”.

Several possible explanations for “rogue” gases have been advanced. One possibility (Keeling, private communication, 1987) is the presence of “rogue” pressure regulators inducing a bias in gas concentration, however we believe this can be discounted. Prior to 1986 the regulators used were changed frequently but with no deliberate pattern, while since 1986 regulators have tended to remain attached to specific cylinders for long periods. If the anomalous behaviour of the “rogue” gases were due to “rogue” regulators the biases that they induced in gas concentrations would have been seen as a more or less random component of error in the repeatability of the URAS-2T analyser. This suggests any regulator bias is less than 0.03 ppmv. However we are proceeding to do a series of controlled exchanges of regulators to confirm this directly.

An alternative explanation is that the ¹³C/¹²C ratio of the CO₂ gas is anomalous for some calibration gases, and that NDIR analysers have different sensitivities to this ratio (Tans, private communication, 1987). We hope to make measurements of the ¹³C/¹²C ratios of calibration gases in future to investigate this possibility.

A third possible explanation is the presence of variable amounts of trace gases with infra-red absorption in the frequency range selected by the NDIR detector. As different NDIR instruments involve different infra-red optical filtering at the sample cell and detector, this frequency spectrum cannot be assumed to be identical.

10 Conclusions

The internal consistency of calibration gases determined by their offset from analyser calibration curves has been shown to be better than 0.10 ppmv (1 σ) for both CO₂-in-nitrogen and CO₂-in-air gases for two different NDIR analysers. Data with better repeatability from the