TR/16/90 December 1990

Fluctuations in Atmospheric Contaminants

by

N. Mole and P.C. Chatwin

FLUCTUATIONS IN ATMOSPHERIC CONTAMINANTS

by

N. Mole* and P.C. Chatwin†

Department of Mathematics and Statistics, Brunei University,
Uxbridge, Middlesex, UB8 3PH.

Final Report written for the Chemical Defence Establishment
under Agreement No. 2066/71.

* Address from September 1990: Department of Mathematics, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ.

† Address from January 1991: Department of Applied and Computational Mathematics, University of Sheffield, Western Bank, Sheffield, S10 2TN.

October 1990

CONTENTS

1. Introduction

As a result of Dr. Mole leaving for a permanent post at the University of Essex in September 1990, and Professor Chatwin's impending transfer to the University of Sheffield, this agreement terminated in September 1990. Consequently this final report covers 19 months of work under the agreement. The material contained in the six monthly report of August 1989 and the annual report of February 1990 will not be repeated here, but only referred to where appropriate. Work carried out since February 1990 will be dealt with in the context of the outline of future work given in Section 5 of the annual report.

2. Work since February 1990

The immediate work outlined in February 1990 was the parametric estimation of probability density functions (pdfs) of concentration using maximum likelihood (ML) methods. Dr. AJ. Jakeman of CRES, ANU, Canberra supplied some software to us for applying ML to fitting the lognormal pdf, amongst others. This software calls routines from the International Mathematics and Statistics Library (IMSL). To use the software it was necessary to make a number of modifications to it (mainly because it was not written in standard FORTRAN), and to run it at the University of London Computing Centre (ULCC), in order to access the IMSL software library. (An alternative approach would be to rewrite the supplied software to make use of the much more widely available NAG, rather than IMSL, software library.) Problems of transmission over the network between Brunei University and ULCC were experienced in running this software on large datasets. These problems were being worked on at the time when the impending termination of the agreement became known. At this point it was agreed to spend the remaining tune on other problems (see below), so this software has still not been propertly tested.

Preparatory to including the truncated normal pdf in this ML estimation software, some theoretical work on applying ML to the truncated normal was carried out. This is included in Appendix A, and deals with the truncated normal treated as both a 2- and 3-parameter distribution.

As stated in the outline of future work in the annual report, one of the major interests was in the effect of source geometry. It had been hoped that experiments using the ion generator system would be carried out in 1990 to investigate this. Unfortunately this proved not to be possible, but the proposals for which experiments to perform are included as Appendix B.

When it became known that the agreement would terminate prematurely, it was agreed to devote the remaining time to digitising and analysing the data from the experiments referred to as set (6) in Section 5 of the annual report. The results of the analysis form Appendix C here.

Appendix D provides a guide to the datasets available, and the software developed for their analysis.

3. Work that would have been carried out during remainder of agreement

The remaining work envisaged under the agreement would have involved the fitting of a variety of pdfs to establish both how many parameters are required to produce a good fit, and which form of pdf, if any, gives good fits over a broad range of cases. Modelling of the pdf would then be possible through modelling of the parameters required to describe such a pdf. The range of cases of particular interest was discussed in Section 5 of the annual report. Principal among them are the effect of sensor properties and source geometry. The modelling component would have formed the major content of the third year of the project.

Appendix E of this report is a paper that has been separately submitted to CDE at their invitation, and is a proposal for work to be undertaken under a new three-year Agreement (supervised by Professor Chatwin at Sheffield). This paper therefore extends the summary in the previous paragraph.

References

PJ. Barry (1977) Stochastic properties of atmospheric diffusivity. *Sulphur and its Inorganic Derivatives in the Canadian Environment.* National Research Council of Canada, pp.313-358.

P.C. Chatwin & N.T. Hajian (1990) Concentration fluctuations in atmospheric dispersion. Final technical report on Agreement No. 2066/62, CDE.

P.C. Chatwin & PJ Sullivan (1989) The intermittency factor of scalars in turbulence *Phys Fluids A 1,* 761-763.

G.T. Csanady (1973) *Turbulent Diffusion in the Environment*, D. Reidel, Dordrecht.

AJ. Jakeman, JA. Taylor & R,W. Simpson (1986) Modeling distributions of air pollutant concentrations - II. Estimation of one and two parameter statistical distributions. *Atmospheric Environment* 12, 2435-2447.

M J. Lighthill (1958) Introduction to Fourier Analysis and Generalised Functions, CUP Cambridge.

KR. Mylne & PJ. Mason (1990) Concentration fluctuation measurements in a dispersing plume at a range of up to 1000m. Submitted to *Quart. J.R. Met. Soc.*

N. Mole and P.C. Chatwin (1989) Fluctuations in atmospheric contaminants. Report on work carried out under CDE Agreement No.2066/71, August 1989.

N. Mole and P.C. Chatwin (1990) Fluctuations in atmospheric contaminants. Annual report written for CDE under Agreement No.2066/71, February 1990.

S.B. Pope (1979) The statistical theory of turbulent flames. *Phil. Trans. Royal Soc.* A291, 529-568.

D.J. Ride (1987) Modelling fluctuations in the concentration of neutrally buoyant substances in the atmosphere. Ph.D. thesis. University of Liverpool.

G.S. Sanders (1987) 1986 field experiments - experimental conditions and arrangements. Report prepared at Pollution Research Unit, UMIST under CDE Agreement 2044/0129.

B.W. Silverman (1986) *Density estimation for statistics and data analysis.* Chapman and Hall, London.

KR. Sreenivasan (1985) On the fine-scale intermittency of turbulence. *J. Fluid Mech.* 151, 81-103

Appendix A Maximum likelihood applied to the truncated normal distribution

Here the probability that the concentration is non-zero is termed the intermittency, and denoted by γ. Although there are practical (because of the presence of noise) and theoretical problems associated with defining intermittency in this way (see Sreenivasan 1985, Chatwin and Sullivan 1989), it is nevertheless practically useful, especially if instrumentation effects can be made negligible. This definition leads to the following expression for the one-point p.d.f. of concentration $\mathrm{p}(\theta,\gamma,\phi)$ in the ideal case of zero noise:

$$\mathrm{p}(\theta,\gamma,\phi) - (1-\gamma)\,\delta(\theta) + \gamma\,\mathrm{f}(\theta;\phi),\ \theta \geq 0 \qquad \text{(Al)}$$

Here $\delta(\theta)$ is a delta-function and $\mathrm{f}(\theta,\phi)$ is the p.d.f. of non-zero concentration, or conditional p.d.f., where ϕ is a parameter (or possibly vector of parameters) of the distribution. Once $\mathrm{f}(\theta,\phi)$ has been chosen the problem is that of finding the ML estimates of γ and ϕ.

If there are n measured values of θ, say θ - ($\theta_1, \theta_2,\ldots \theta_n$), the likelihood function is

$$\mathrm{L}(\gamma,\phi;\theta) - \quad_{-1} \quad \theta_i : \gamma.\phi). \qquad \text{(A2)}$$

In order to proceed (Al) is rewritten as

$$\mathrm{p}(\phi,\gamma,\phi) = (1-\gamma)\mathrm{g}(\theta;\sigma) + \gamma\mathrm{f}(\theta;\phi) \qquad \text{(A3)}$$

Where

$$\mathrm{g}(\theta;\sigma) - \frac{1}{\sqrt{2\pi\sigma}}\mathrm{e}^{-\theta^2/2\sigma^2} \qquad \text{(A4)}$$

Since $\delta(\theta)$ is defined by the limit as σ 0 of $(\theta;\sigma)$ (p.17 of Lighthill 1958) the ML problem for (A3) can be solved with fixed a, and then letting $\sigma \rightarrow 0$.

Suppose that m of the measured values of 0 are non-zero, and for simplicity order θ so that these are the first m values. Then

$$\mathrm{In\ L}\,(\gamma,\phi\theta)\ \text{-(n-m) In}\left\{\frac{1-\gamma}{\sqrt{2\pi\sigma}} + \gamma\mathrm{f}(0;\phi)\right\} + \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{In}\{(1-\gamma)\mathrm{g}(\theta_i;\sigma) + \gamma\mathrm{f}(\theta_i;\phi)\}.$$

Thus

$$\frac{\partial}{\partial\gamma}\mathrm{In\,L}(\gamma,\phi;\theta) - (\mathrm{n}-\mathrm{m})\frac{\left\{\mathrm{f}(0;\phi) - \frac{1}{\sqrt{2\pi\sigma}}\right\}}{\frac{1-\gamma}{\sqrt{2\pi\sigma}} + \gamma\,(0;\phi)} + \sum_{=1}^{\mathrm{m}} \frac{\mathrm{f}(\theta;\phi) - g(\theta;\sigma)}{(1-\gamma)g(\theta;\sigma) + \gamma\mathrm{f}(\theta;\phi)}.$$

Putting ∂ L/∂ γ = 0 and letting σ→0 gives (since $\lim_{\sigma\to 0} g(\theta;\sigma) - 0$for $= 1......m$) the ML estimate:

$$\tilde{\gamma} \text{ - m/n} \tag{A5}$$

Thus the intuitive estimate of y is also the ML estimate.

Similarly, $\tilde{\phi}$ 5 is the limit as $\sigma \rightarrow 0$ of the solution of

$$\frac{\partial}{\partial\phi}\text{In}\,\text{L}_{\text{f,m}}(\phi;\theta) = -\sqrt{2\pi\sigma}(\text{n}-\text{m})\frac{\gamma}{1-\gamma}\frac{\partial}{\partial\phi}\text{f}(0;\phi).$$

Where

$$L \quad (\phi;\theta) - \prod_{-1} \ (\theta;\phi). \tag{A6}$$

So provided $\frac{\partial}{\partial\phi}\text{f}(0;\phi)$ is finite, $\tilde{\phi}$ is the solution of

$$\frac{\partial}{\partial\phi}\text{In}\,\text{L}_{\text{f.m}}(\phi;0) = 0 \tag{A7}$$

(Mathematically these calculations do not hold when γ - 0, but this is the irrelevant case of no pollutant.) In other words, the complete ML problem reduces to the ML problem for the conditional p.d.f. with only the positive data considered.

A number of suggestions have been made for the form of f(θ;Ø), among them the exponential (Barry 1977) and lognormal (see, for example, Csanady 1973). Both of these have well-known ML solutions (see, for example, Jakeman et al. 1986). Another p.d.f. which has recently been receiving much attention (e.g. Mylne and Mason 1990; Ride 1987; Pope 1979) is the truncated normal:

$$\text{f}(\theta,\mu,\sigma) - \frac{\exp\{-(\theta-\mu)^2/2\sigma^2\}}{\sqrt{\frac{\pi}{2}}\sigma\{1+\text{erf}(\mu/\sqrt{2}\sigma)\}}\theta \geq 0, \sigma > 0.$$

This gives

$$\text{In}\ \text{L}_{\text{f,m}}(u,\sigma,0) = \frac{-1}{2\sigma^2}\sum_{\text{i}=1}^{\text{m}}(\theta_i-\mu)^2 - \text{m}\ \ln\left[\sqrt{\frac{\pi}{2}}\sigma\{1+\text{erf}(\mu/\sqrt{2\ \sigma})\}\right].$$

From (A7), the ML estimates $\tilde{\mu}$ and $\tilde{\sigma}$ are then the solutions of the pair of equations:

$$\begin{cases} \mu + \sqrt{\dfrac{2}{\pi}}\dfrac{\sigma \exp}{1+\operatorname{erf}}\dfrac{\mu^2/2\sigma^2)}{\mu/\sqrt{2\sigma}} = \bar{\theta} & \text{(A8a)} \\ \\ \sigma^2 - \bar{\theta}(\bar{\theta}-\mu) + s^2 & \text{(A8b)} \end{cases}$$

Where

$$\bar{\theta} = \frac{1}{m}\sum_{i=1}^{m}\theta_i,\; s^2 = \frac{1}{m}\sum_{i=1}^{m}(\theta - \bar{\theta})^2.$$

It is often the case, however, that when the truncated normal is proposed, it is also proposed that the intermittency estimate y be determined by the area under the retained part of the normal distribution (Ride 1987; Mylne and Mason 1990), i.e.

$$\tilde{\gamma} - \tilde{\gamma}(\tilde{\mu},\tilde{\sigma}) = \frac{1}{2}\left\{1 + \operatorname{erf}(\tilde{\mu}/\sqrt{2\tilde{\sigma}})\right\}. \quad \text{(A9)}$$

This can then be compared with the estimate (A5). Ride (1987) and Mylne and Mason (1990) find some degree of agreement, but did not use ML to find $\tilde{\mu}$ fl and $\tilde{\sigma}$ a. It should be pointed out that an analogous estimate for γ could be obtained from any other chosen p.d.f. by similar truncation, possibly after translation of the origin.

From this viewpoint there are only two independent parameteters to be determined, namely μ and σ. In this case (A7) cannot be used; instead the solution must be derived from the full problem of maximising (A2) with γ- γ (ϕ). This leads to

$$\frac{\partial}{\partial\phi}\operatorname{InL}^1(\phi;\theta) = \left(\frac{n-m}{1-\gamma 0}\right)\frac{\partial\gamma}{\partial\phi}$$

where

$$L^1(\phi;\theta) = \gamma^m L_{f,m}(\phi;\theta).$$

For the truncated normal the relevant terms are:

$$\operatorname{In} L^1(\mu\mu\sigma,\theta) - m\operatorname{In}\left(\sqrt{2\pi}\,\sigma\right) - \frac{1}{2\sigma^2}\sum_{i-1}^{m}(\theta_i - \mu)^2$$

$$\frac{\partial\gamma}{\partial\mu} = \frac{1}{\sqrt{2\pi\sigma}}e^{-\mu^2/2\sigma^2}$$

$$\frac{\partial \gamma}{\partial \sigma} = \frac{-\mu}{\sqrt{2\pi\sigma^2}} e^{-\mu^2/2\sigma^2}$$

The equations to be solved are, therefore, the following:

$$\begin{cases} \mu + (\frac{n}{m} - 1)\sqrt{\frac{2}{\pi}} \frac{\sigma \exp(-\mu^2/2\sigma^2)}{1 - \mathrm{erf}(\mu\sqrt{2}\sigma} = \bar{\theta} & \text{(A10a)} \\ \\ \sigma^2 - \bar{\theta}(\bar{\theta} - \mu) + s^2 & \text{(A10b)} \end{cases}$$

Equations (A10) are identical to equations (A8) except that the second term in (AlOa) has an extra factor (l-m/n) γ /{(l- γ)m/n}. When m = n (i.e. all data are positive) then (A10) gives $\tilde{\mu} = \tilde{\theta}$ and $\tilde{\sigma}$ = s, and (A9) shows that the fitted values of intermittency for the two methods will agree to within 10% if $\bar{\theta}$ /s≥ 1.27, and to within 1% if $\bar{\theta}$ /s > 2.32. In general one might expect substantial differences between the two estimates of intermittency to exist. Use of (A5), (A8), (A9) and (A10) shows that the ratios of the two estimates for γ, μ and σ^2 reduce, respectively, to the following three functions (which then equal unity) in the limit as the estimates become identical:

$$\begin{cases} A(x,m/n) - \frac{2m/n}{1 + \mathrm{erf}(x/\sqrt{2})} & \text{(A11a)} \\ B(x,m/n) - \frac{1 + D(x)}{1 + E(x,m/n)} & \text{(A11b)} \\ C(x,m/n) - \frac{(1 + s^2/\ \bar{\theta})\{(1 + D(x\} - 1}{(1 + s^2/\ \bar{\theta}^2)\{1 + E(x,m/n\} - 1} & \text{(A11c)} \end{cases}$$

Where x=μ/σ and

$$D(x) - \sqrt{\frac{1}{\pi}} \frac{e^{-1}/2 \times^2}{\times\{1 + \mathrm{erf}(x/\sqrt{2})\}}$$

$$E(x, m/n) - (\frac{n}{m} - 1)\sqrt{\frac{2}{\pi}} \frac{e^{-1/2} \times^2}{\times (1 - \mathrm{erf} \times /\sqrt{2})\}}.$$

For the estimates given by the two methods to be close, therefore, the functions A, B and C should all be close to 1. Figure Al shows contour plots of these functions (in the case of C, for several values of $s^2/\bar{\theta}^2$). Note that in the limit $s^2/\bar{\theta}^2 \rightarrow \infty$,C becomes identical to B. If A(x,m/n)=1 then B and C are both automatically equal to 1, so either all, or none, of γ, μ and a are insensitive to which of the two methods is used. When m/n - 1 the two estimates of μ and of σ will agree to within 10% if $\bar{\theta}$ /s is greater than about 2, and to within 1% if $\bar{\theta}$ is greater than about 3.

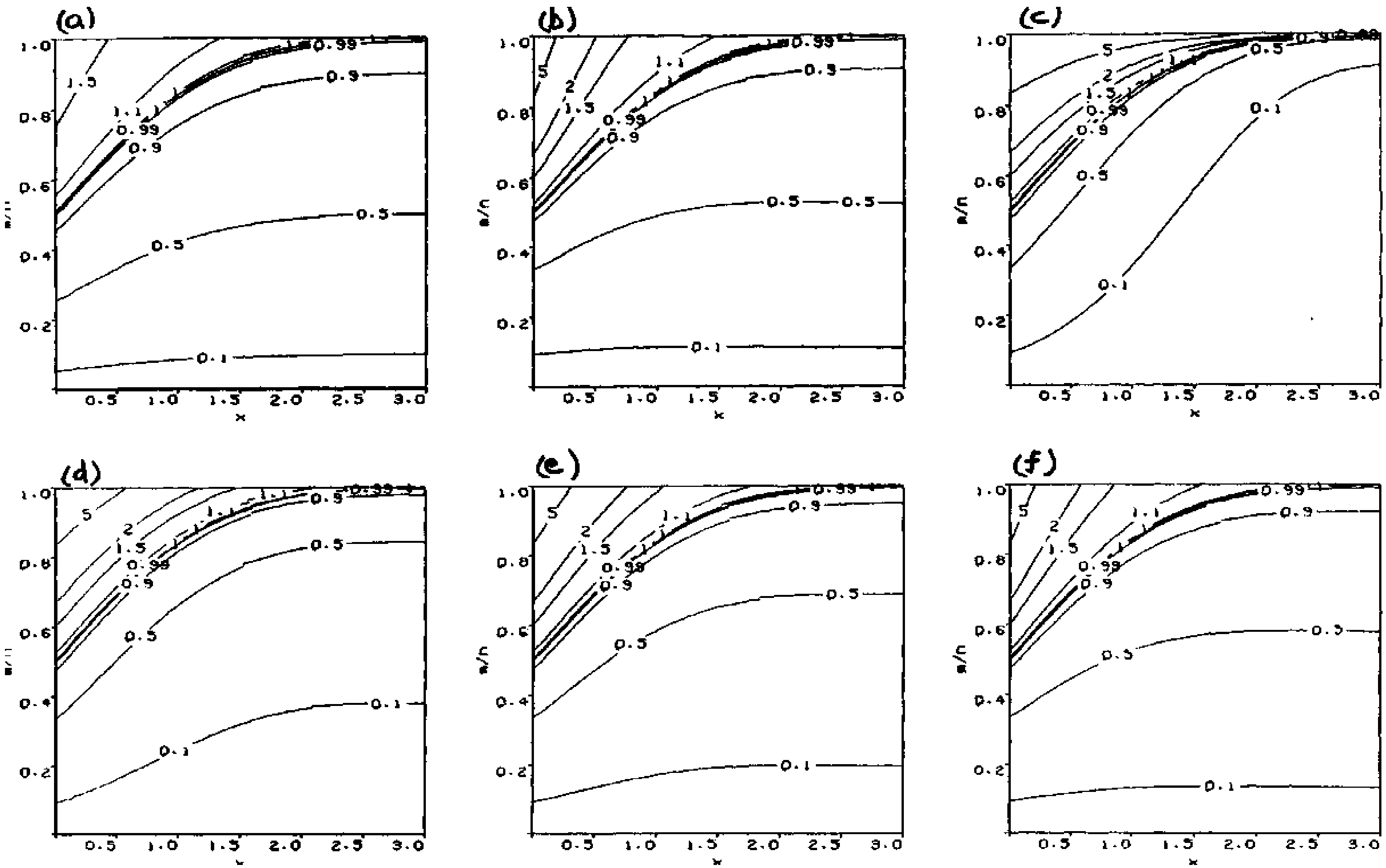
The method of moments gives the same estimates for μ and σ as the ML method when γ is estimated by ML, i.e. those satisfying (A8). When γ is estimated from (A9) the method of moments yields equations similar to (A10). With a truncated normal conditional p.d.f. it does not, therefore, allow any simplification over the ML calculations.

Figure Al. Contour plots of the functions defined in (All). In all cases the contour values plotted are 0.1, 0.5, 0.9, 0.99, 1.0, 1.01, 1.1, 1.5, 2.0 and 5.0

(a) A(x,m/n), (b) B(x,m/n).

(c) — (f) are of C(x,m/n), with the following values of $s/\bar{\theta}$: (c) 0.1, (d) 0.5, (e) 1.0, (f) 2.0.

FIGURE A1

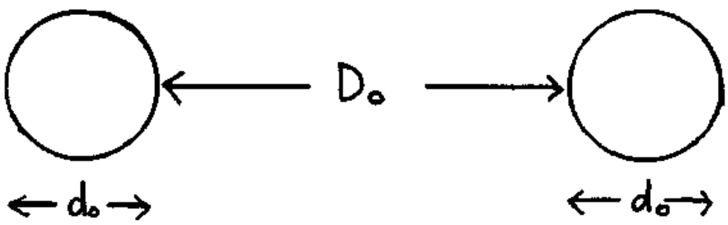


Appendix B Proposals for multiple source experiments

The aspect of source geometry which it is most desirable to study initially is the source size. However, with the ion generator system, what is available is several generators of the same size. While a configuration of several such sources can be studied explicitly, one also wants to know when, by siting the sources as close together as possible, they will behave like one larger source.

For two sources to behave as one larger source, a minimum requirement is that the instantaneous plumes from the two sources should almost completely overlap. A less stringent necessary condition is that the mean plumes should overlap. This condition is examined below.

Suppose we have 2 sources of diameter d , with centres separated by $D_o + d_o$



Taylor theory for mean plume from point source in homogeneous turbulence implies that at small (Lagrangian) times t the plume spread a_y satisfies

$$\sigma_y \sim \sigma_v t$$

or $$\sigma_y \sim \frac{\sigma_v}{\bar{u}} x$$ using Taylor's hypothesis.

Here we do not have a point source. Instead use

$$\sigma_y^2 \sim \frac{\sigma_V^2}{\bar{u}^2} x^2 + \frac{1}{8} d_0^2 .$$

We require

$$\sigma_y \gg D_0 + d_0$$

$$\Rightarrow \quad x \gg X_1$$

where

$$X_1 = \frac{\bar{u}}{\sigma_v}\left[(D_0 + d_0)^2 - \frac{1}{8} d_0^2\right]^{1/2} . \qquad \text{(B1)}$$

Even leaving aside the other assumptions involved, this might not work well if

1) The time is not small enough (i.e. we do not have $x \ll \bar{u}\int_O^\infty R_{22}(r)dr)$ for $\sigma_y \alpha\, t$ in the Taylor theory. This would result in overestimating σ_y and hence not being cautious enough.

2) Electrostatic repulsion increases the distance between the plumes by an amount comparable to the spread of the plumes.

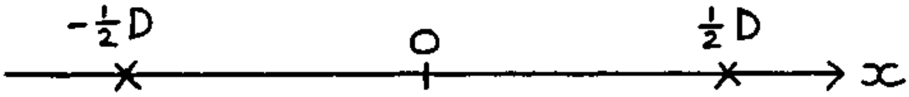
Point 2) can be tackled in a crude manner as follows.

Effect of electrostatic repulsion

Even tackling the problem of 2 infinite cylinders of constant cross-section (i.e. a 2D problem) and uniform charge density is difficult without resorting to a computer. This is because the repulsive effect causes the cylinders to distort from their cylindrical shape.

To get a crude idea of the effect, try 2 ways of further idealising this problem:

(a) Idealise the cylinders to lines of charge Q per unit length.



Line charge has $E = \frac{Q}{2\pi\pi_0}\frac{r}{r^2}$.

At $x - \frac{D}{2}$ the field due to the other line charge is $\frac{Q}{2\pi\varepsilon_0 D}\hat{x}, so$

$$\dot{D} = \frac{\mu Q}{\pi\varepsilon_0}\frac{1}{D}$$

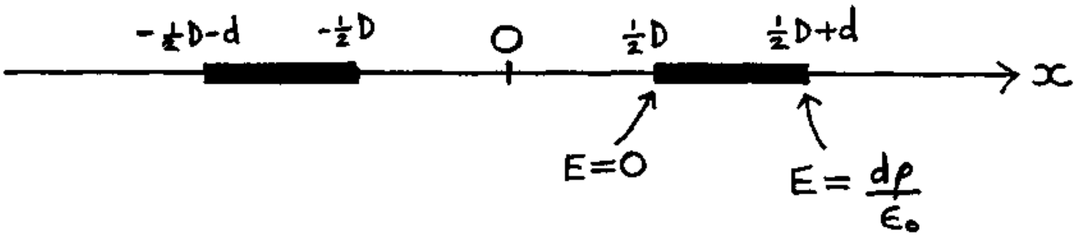
$$\rightarrow\ D^2 - D_0^2 + \frac{2\mu Q}{\pi\varepsilon_0}t.$$

For the problem of interest choose $Q = Q_0 = \pi\ {}_0^2\rho_0$ and write $T_{es} - \frac{\varepsilon_0}{\mu\rho_0}$.

Then

$$D^2 = D_0^2 + \frac{2\ {}_0^2}{T_{es}\bar{u}}x.$$

(b) Idealise to planar slabs of thickness d (i.e. a ID problem).2



Charge Q per unit area

$$Q = d_0\rho_0 = d\rho \quad \text{(change conservation)}$$

$\therefore$ Have

$$\begin{cases} \dfrac{d}{dt}(\dfrac{1}{2}D + d) = \dfrac{\mu Q}{\varepsilon_0} \\ \\ \dfrac{d}{dt}(\dfrac{1}{2}D) = 0 \end{cases}$$

$$\Rightarrow \quad \begin{cases} D = D_0 \\ \\ \dfrac{d}{d_0} = 1 + \dfrac{x}{\bar{u}T_{es}}. \end{cases}$$

Require these estimates to give $\dfrac{d}{dx}(D).\dfrac{d}{dx}(d) \ll \dfrac{d}{dx}\sigma_Y$.

(a) We require

$$\frac{d_0}{T_{es}\sigma_v} \ll \frac{\sigma_v}{\overline{\cup}}.\frac{x}{d_0}$$

i.e

$$x \gg X_3 = d_0\left(\frac{d_0}{\sigma_v}\right)\left(\frac{\bar{u}}{\sigma}\right)$$

For the values given in the table below $X_3 < X_1$, so this does not provide an extra constraint.

(b) Require

$$-\frac{d_0}{T_{es}\sigma_v} \ll 1 \tag{B2}$$

To avoid charge leakage to the ground we also require $d \ll H_o$, where H_O is the source height.

For a single plume electrostatic repulsion alone gives

$$d^4 = d_0^4 + \frac{2d_0^4 \mu \rho_0}{\overline{u} \in_0} x$$

(see Chatwin and Hajian (1990) - final report to CDE).

If the ionised air is blown out of the generators at V_0 (ms^{-1}) then the generator current i_0 and charge density ρ_0 are related by

$$i_0 = \pi d_0^2 \; p_0 v_0$$

Thus we derive the condition

$$x \ll x_2 - \frac{\pi \overline{u} V_0 \in_0 (H_0^4 - d_0^4)}{2d_0^2 \mu i_0}$$

(B3)

and (B2) becomes

$$F = \frac{\mu i_0}{\pi d_0 V_0 \in_0 \sigma_v} \ll 1.$$

(B4)

So we want to satisfy (Bl), (B3) and (B4) simultaneously. For the present apparatus we have

$V_0 = 3ms^{-1}$, d_0 - 0.1m and a minimum possible of $D_0 = 0.2$m.

If we choose $H_0 = 2$m, and use $\in_0 = 8.85 \times 10^{-12}$ CV^{-1} m^{-1}, $\mu = 10^{-4}$ V^{-1} $m^2 s^{-1}$ then we find

$$X_1 \approx 0.30 \left(\overline{u} / \sigma_v\right) m$$

$$X_2 \approx 6.7 \times 10^{-4} \left(\overline{u}/i_0\right) m$$

$$- 6.7 \times 10^5 \; (\overline{u} / 10^9 i_0) \; m$$

$$F \approx 1.2X10^{-2} \left(10^9 i_0/\sigma_v\right)$$

and we require

$$X_2 \gg \times \gg X_1,$$

$$F \ll 1.$$

$\bar{u}\ (ms^{-1})$	$\sigma\sqrt{u}$	$i_0(10^{-9}A)$	F	$X_1(m)$	$X_2(m)$
1	0.1	5	0.6	3.0	1.3×10^5
5	0.1	25	0.6	3.0	1.3×10^5
10	0.1	50	0.6	3.0	1.3×10^5
5	0.1	50	1.2	3.0	6.7×10^4
5	0.1	10	0.24	3.0	3.4×10^5
5	0.5	50	0.24	0.6	6.7×10^4
5	0.05	50	2.4	6.0	6.7×10^4
5	0.05	5	0.24	0.6	6.7×10^5

Conclusions

1) The condition on X_2 does not provide a constraint in practice.

2) If $\sigma\sqrt{u} \sim 0.1$, measurements will not become single source - like until at least 10m downstream, and probably rather further.

3) The maximum possible i_0 (in nA) is about 50 σ_v (with σ_v measured in ms^{-1}), and preferably less.

Even if the plume becomes single source - like downstream, it is still not clear what the effective source size is.

Source and collector arrangements

The suggested source arrangements are as follows:

If have 3 sources, in order of priority:

z y

1) *o*
oo

2) *o o o*

3) *o*
o
o

In all cases use minimum sensible spacing (assumed to be 20cm gaps in the above calculations).

If have 2 sources:

1) $o \quad o$

2) $\begin{matrix} o \\ o \end{matrix}$

The arrangements other than the primary one are probably not worth much effort - at most make short tests to see whether the arrangement makes a significant difference. Experiments should be performed in as near identical conditions as possible for the 1, 2 and 3 source cases.

The favoured arrangement for the ion collectors is to set them out as close as possible to the mean centreline, at different downstream distances. Ideally one collector would be closer than the distance at which the plume might become single source - like (e.g. 5m if $\sigma\sqrt{u} \sim 0.1$), one would be in the intermediate range (e.g. 10m if $\sigma\sqrt{u} \sim 0.1$), and the rest further downstream. The latter would be the priority if there are not enough collectors. Arranging the collectors crosswind at each of these distances in turn would also be acceptable.

It is desirable to measure u, both so that the turbulence statistics are available for general analysis purposes, and so they can be used to check the satisfaction of the above conditions.

Appendix C Analysis of experiments carried out under convectively stable and unstable conditions

These experiments were conducted by Dr. CD. Jones in November 1989 at the US Army Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico. His field notes on the experiments are included here. The data were recorded in analog form on magnetic tape.

They were then digitised at l0Hz using the ISC-67 software, by passing the signals through a 5Hz filter and then through a 12 bit A-D converter. The voltage range of the A-D converter was (-2.5,2.5)V, so the bin width is about 1.2mV. The actual range of the data was about (0,2.2)V, so about 1800 bins are utilised.

The zero levels on the tape recorder used to playback the data when digitising are displaced by small amounts, which appear from visual inspection to be roughly: Channel l-140mV, Channel 2-30mV, Channel 3-20mV, Channel 4-10mV.

If the ion collector current used in the experiment is I (in pA), then the voltage V is related to the charge density Γ (in nCm^{-3}) as follows:

$$\Gamma = 0.354 \text{ IV}.$$

The digitisation is such that 0 corresponds to -2.5V, 2048 corresponds to 0V, and 4096 corresponds to 2.5V. Thus, if the digitised value is n, the voltage is given by:

$$V = 2.5\left(\frac{n-2048}{2048}\right).$$

Source details: The source diameter was 7.5cm, the generator output current was about 30nA, and the ions were expelled at a velocity of $3ms^{-1}$. This gives a source concentration of $\theta_0 \approx 570nCm^{-3}$.

Before and after most experiments, a few minutes of measurements were recorded while the ion generator was switched off, to provide statistics for the background noise. Comparing these for periods immediately before and after experiments shows that there is some drift. This is not large enough to be important for most of the results shown, but is significant for the mean and intensity (and probably also the intermittency) in those cases when the mean is small. No attempt has yet been made to remove the drift.

The time series of concentration and the pdfs for the periods of noise which are presented below are based on the raw data. All other statistics have been calculated after:

1) Subtracting the means of the periods of noise immediately before the experiment (or after in the case of experiments 11 and 16) from the time series of concentration, and then
2) Excluding occasional spurious negative spikes by ignoring any concentration values below $0.22nCm^{-3}$.

Experiments 8, 9, 10, 11 and 13 were conducted in convective daytime conditions (stability class ~ B), while experiments 16 and 17 were carried out in stable conditions after sunset (stability class ~ D/E).

In all cases except experiment 13, the channel number is the same as the ion collector number. In experiment 13:

Channel 1 = IC1, Channel 2 = IC3, Channel 3 = IC4, Channel 4 - IC2.
Concentration units in all cases are nCm^{-3}.

1-

TRIALS DATA 8-11-89

(1) Weather (pn)

Cloud loss N-N∊ and gusting 2-6ms^{-1}.

Estimated 'B' stab

Not data available.

TAPE:

TRIAL (1) 0000 — 0165

TRIAL (2) 0B1 — 0273 -0283

↑

IG off

(2) Tape recorder

(1) Voice + WWV

(2) IC1

(3) Flutter amp

(4) IC2 — Ch(4) approves to fail during Trail (2)

(5) IC3

(6) IC4

(7) —

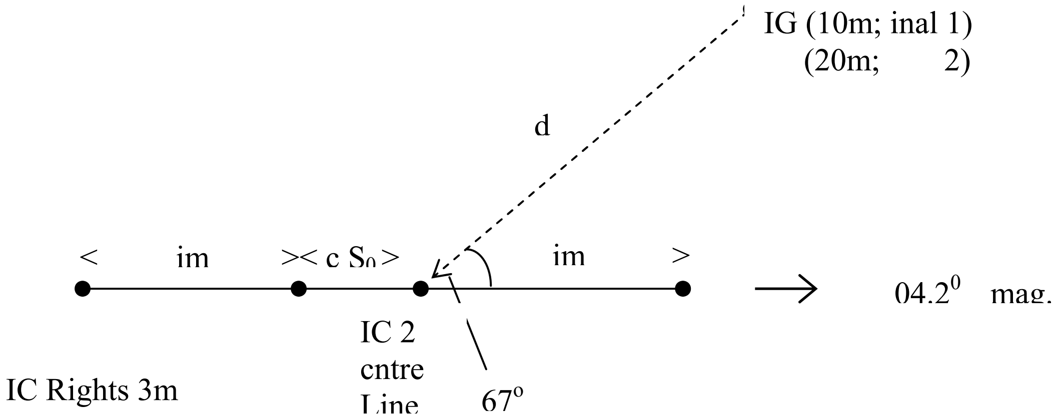
(8) US?

(3) Commenced ~ 1400 ; 5m background – 10m D that -5min background - 20 mD inter finish ~ 1535.

(4) IC's set to I think (100pA) then miues to 30 pA – check Voice channel.

(5) IG – free running at – 3kv – control circuit satisfied.

<u>TRIALS 1,2</u> 8-11-89



<u>TRIALS DATA</u> <u>9-11-89</u>

(1) ~ cloudloss. A little G later.

Wind generally from N-E but highly conveitue at tends.

'B' stab.

<u>Tral 3</u>

Array as Trals 1,2.

IC's 30pA

IC1	Ch2		IG on ~ 3kv. un controlled Tape 0296 – 0481	pababl not very much dater
IC2	Ch7	(Ch 4 defecticve)		
IC3	Ch5	Check it correct ?		
IC4	Ch6			

16^0C 17% 1029 am. 5mins IG off to start with. — some flucti @039C

IC2 → Ch7 Tape:0297 IG15mu

<u>Tral 4</u> array exactly as before line 006^0E(n os) and IC boom rotated to be ⊥.(ie 10mD IC's, 3mH)

Tape No:- <u>0490</u> , 1122 – 5mins IG off

18^0C 15%

IC's 30pA (18pA peak currents seen)

Start: 0486

Generator on-0499
Generator off-0761

<u>Tral 5</u> A 0770 (1245pm)

Short Tral with zero array but log rattings on IC's – abandoned because of returation effects.

DO NOT DIGITIZE 2nd PRIORITY

<u>Tral 5B</u> C822 (1258pa)
as ↑ but IC's as 100pA setting.
Started 1258, shapped 1309.

DO NOT DIGITIZE 2nd PRIORITY

<u>Tral 6</u> 0859(1310pm) — 0201350)

as ↑ but IC's on 10pA setting
IG (at – 3xV) u = 1-3ms^{-1}. very convective

Observed that most of IC noise is at ~ 200Hz – should be aseranble to filtering.

DO NOT 2nd PRIORITY

<u>Tral 7</u> 1070 (1404pm) 20^{0}C 13%

IG 12.5m array as ↑

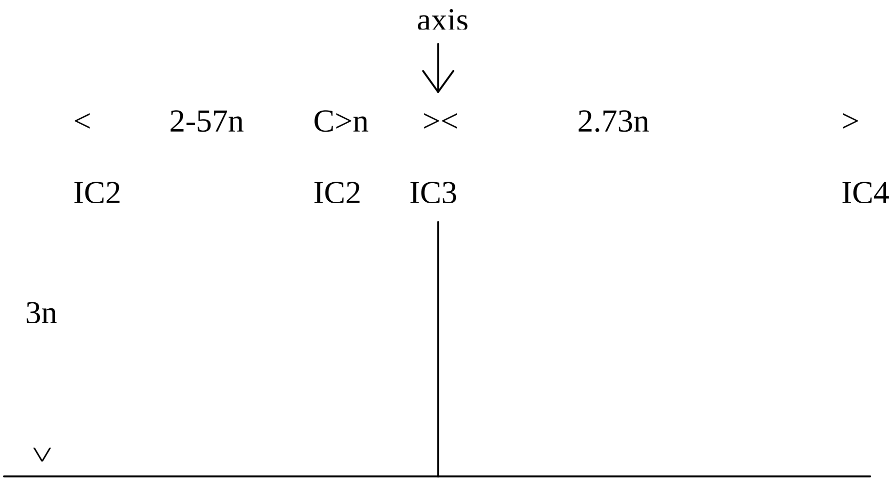
Run till end of tape (1399)

Very convective probably only occasional busts recorded.

DO NOT DIGITIZE 2nd PRIORITY

<u>TRIALS DATA</u> <u>10th NOV 89</u>

1. <u>New IC array</u> - looking down wind:



<u>Tral 8</u>

1. array axis as ↑ 006^0 may (0.18^0) True. Wid uare / areno aligned on True North

IG R.Sm U 3mH, fixed at -3.07kv. OK

Start 1005 Tape 0000
IG on 1009 Tape 0011

u ~ $2ms^{-1}$ <u>NNε</u>
Cloudlos

IG off at 1051. Tape 0130

IC's 10pA range.

DIGITIZE – some noise present – will need filtering.

17 Tape laud saye 1046.

array as ↑

IG at 15mU , on at 1056 (0143 tape)

IC's 10pA OK

IG off 1143 (0274 tape) ~~DIGITIZE~~ -same noise spikes.

T10

array as ↑

OK

IG at 10mU on at 1151 (0302 tape)

IC's 10pA

IG off @ 1236 (0443 tape). DIGITIZE

T11

array as ↑ OK

IG 75mU IC's 30pA DIGITIZE

1243pm (0462) – 1344pm (0665)

T12

array as ↑

IG 5mU IC's 30pA

Probably no good too convective.

1349 IG on (Tape 0683)

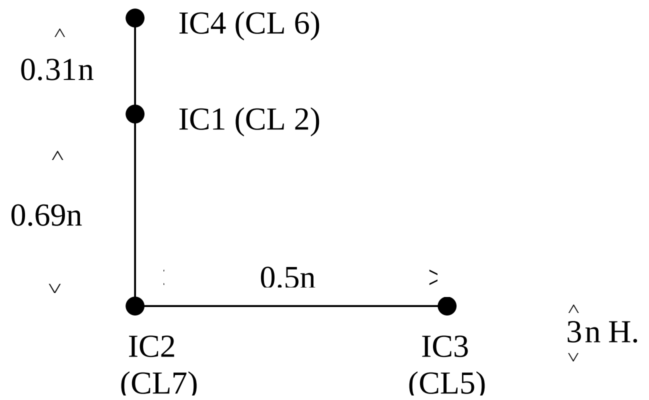
1423 IG off (Tape 0804)

1427 (Tape 0820) Trailed

DO NOT DIGITIZE

T 13 — 13 Nov 89

New IC array – looking up ward: (45 tape scale !)



IG 3.5H 5mU uncontrolled at ~3.05kV.

Convective u 1-3ms^2 generally from N.

(750 G)

Tape start 0828 (1105am) — DIGITISE

IG on 0841

IC's 30pA. 20^{o}C 17% ~12noon.

Off at 1159, 1227pm

T14

as ↑ IG at 10nU On at 1170 , 1229 pm

IC' at 30pA

Pave last at 1304, (1333 on tape).

Pave out again at ~ 1330. (1317 on tape)

NIHIL DIGITATIS

Very convective – probably so good.

14 Nov 1989

T15-18 (stable conditions)

IC array as T13,14. No that tewe dator _ as a result of pave failure.

Same @ 1702

T15

IG 5mU, 3.5mH

IG off 0000 1715pm, IG on at 1715 (tape 0009)

array as ↑ IC's @30pA.

NIHIL DIGITATIS

Bearing 223^0 mag

Trial stepped to realign array. (0037)

T16

Center line now 243^0 (mag) IC array not rotated. DIGITIZE

Wrd W ~ 4,5ms^{-1}. 4 oktor a 18^{0}C 23% @ 1757

0039 tape on at 1730 } Local peak core ~ 26nCm^{-3}

0123 tape: off at 1801hrs } Ture

T17

Tape 0126 at 1803pm, IG on at 1805pm 131 tape

IG 10mU, 3.5mH. IC's as ↑ @ 30pA DIGITIZE

G off at 1831pm 204tape.

18 Nov 89 cont

T18

as ↑ IG 15mU, IC's10pA

0226 1839pm start

0234 1841pm IG on.

Not much seen.

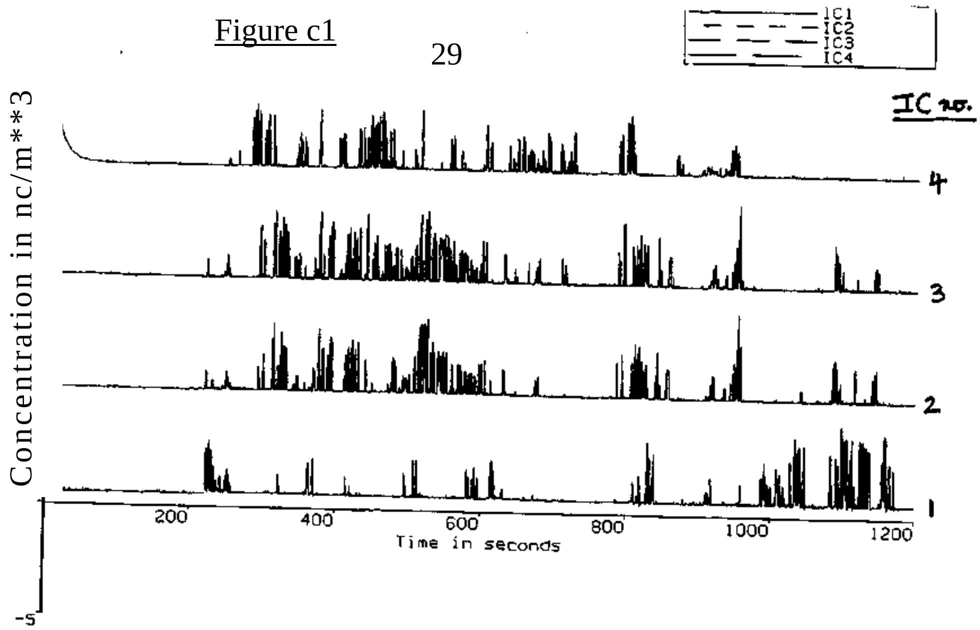
Stopped 0327, 1912pm

Time series of concentration (in nCm^{-3}) are shown for all the experiments in Figures C1-C6. In all cases the all cases the whole file (including the final part which is padded out with zeros) is plotted.

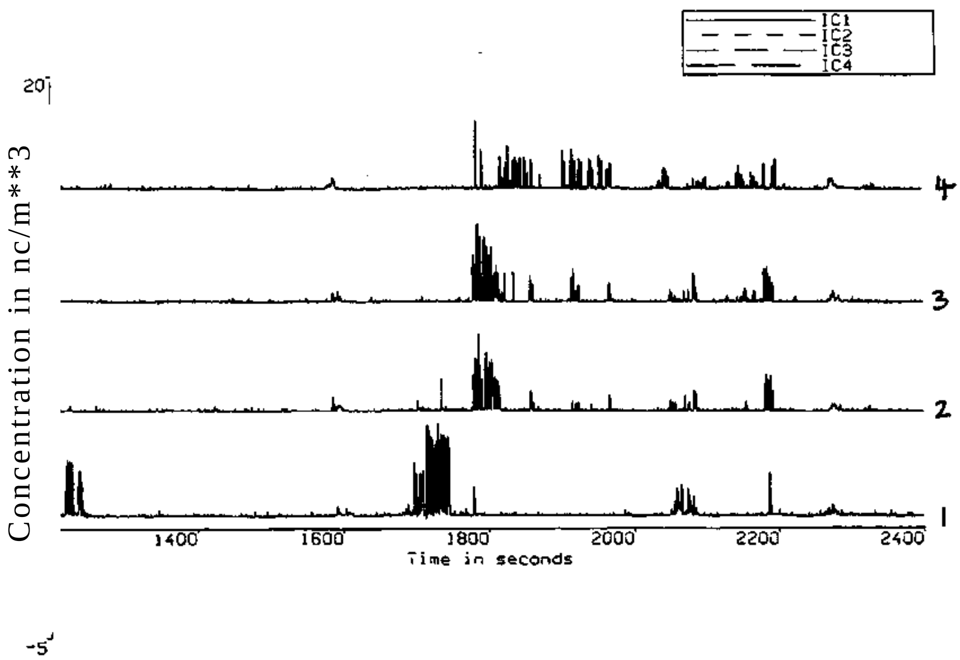
Figure C7 shows pdfs for various periods representative of measurement system noise (i.e. during which the ion generator was off). The pdfs have been estimated by the Gaussian kernel method described in the Annual Report of February 1990.

Figure C8 shows pdfs for the actual experiments, after the means of the appropriate periods of noise have been subtracted.

<u>Figure c1</u>

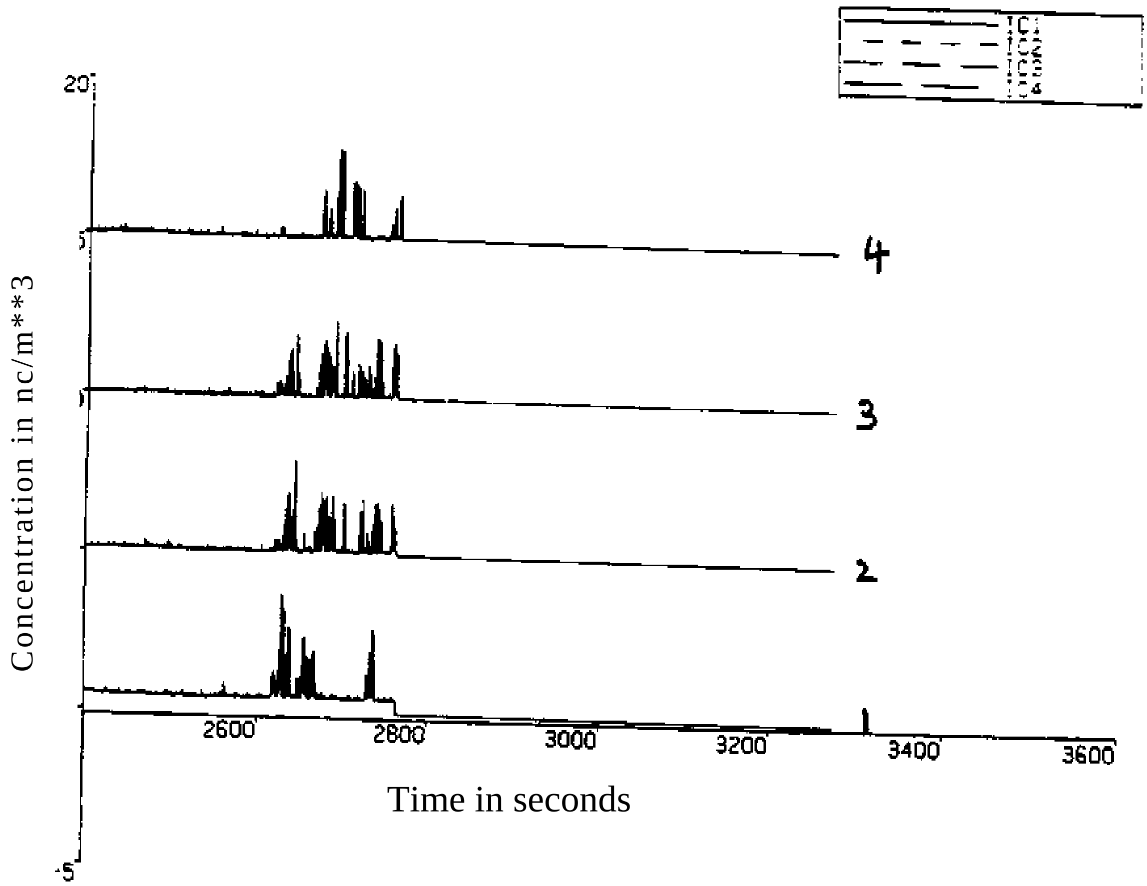


as 108. num

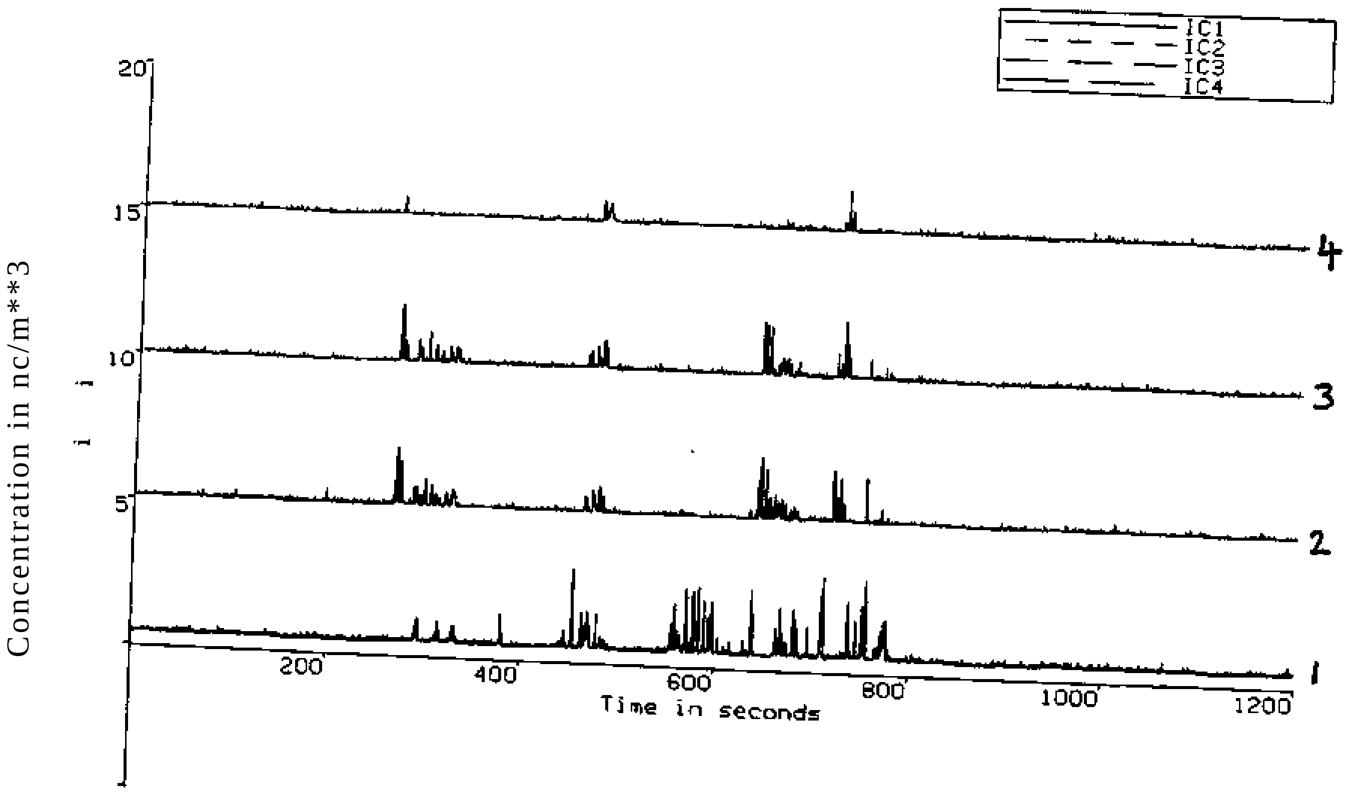


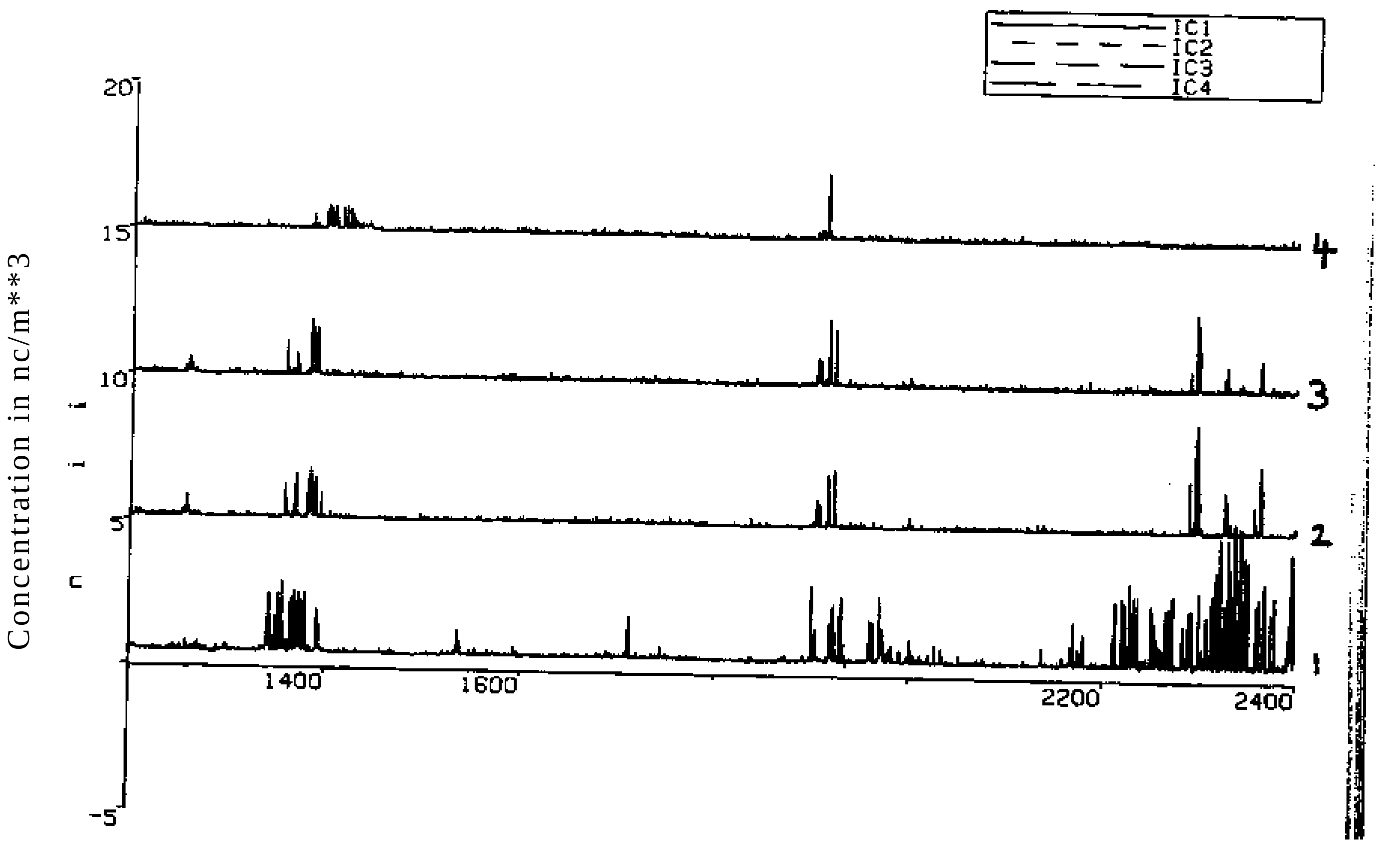
as 108. num

FIGURE C1

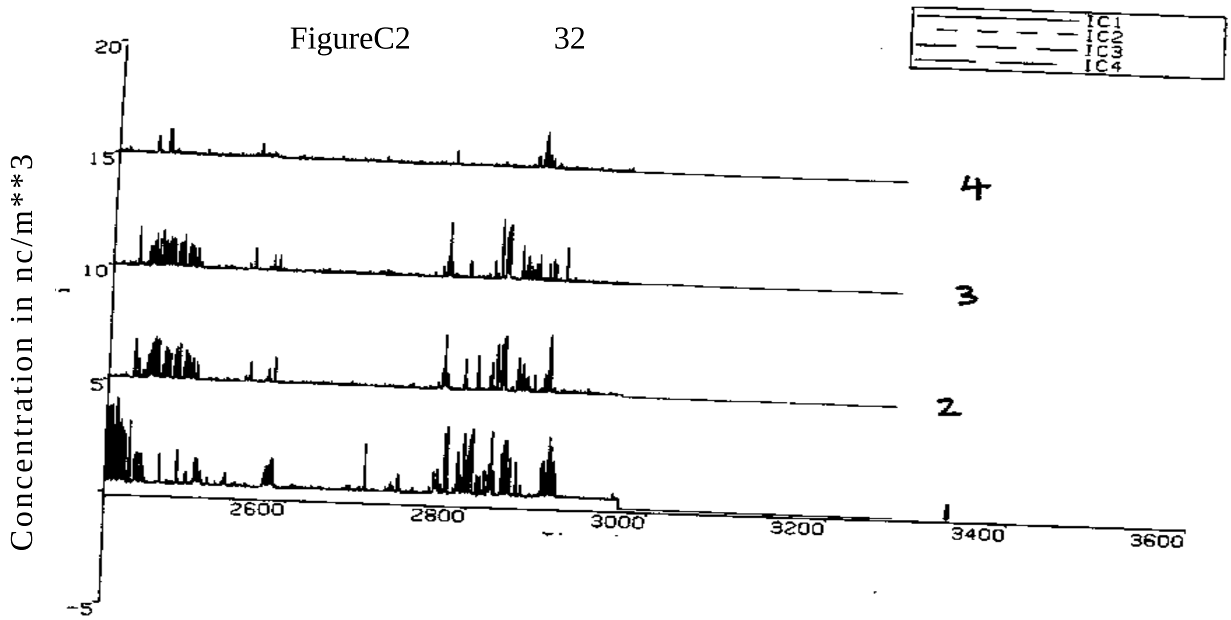

20
4
3
2
1
2600
2800
3000
3200
3400
3600
-5
Concentration in nc/m**3
Time in seconds

FigureC2


IC1
IC2
IC3
IC4
Concentration in nc/m**3
20
15
10
5
200
400
600
800
1000
1200
Time in seconds
4
3
2
1

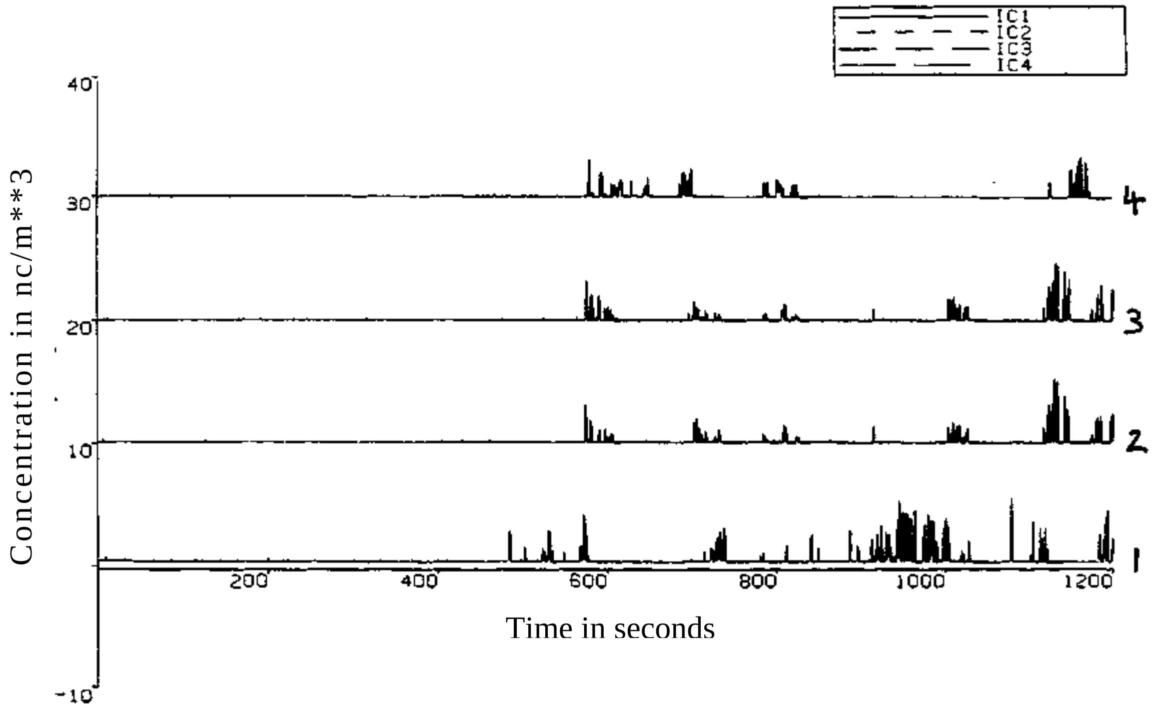

IC1
IC2
IC3
IC4
Concentration in nc/m**3
20
15
10
5
-5
1400
1600
2200
2400
4
3
2
1

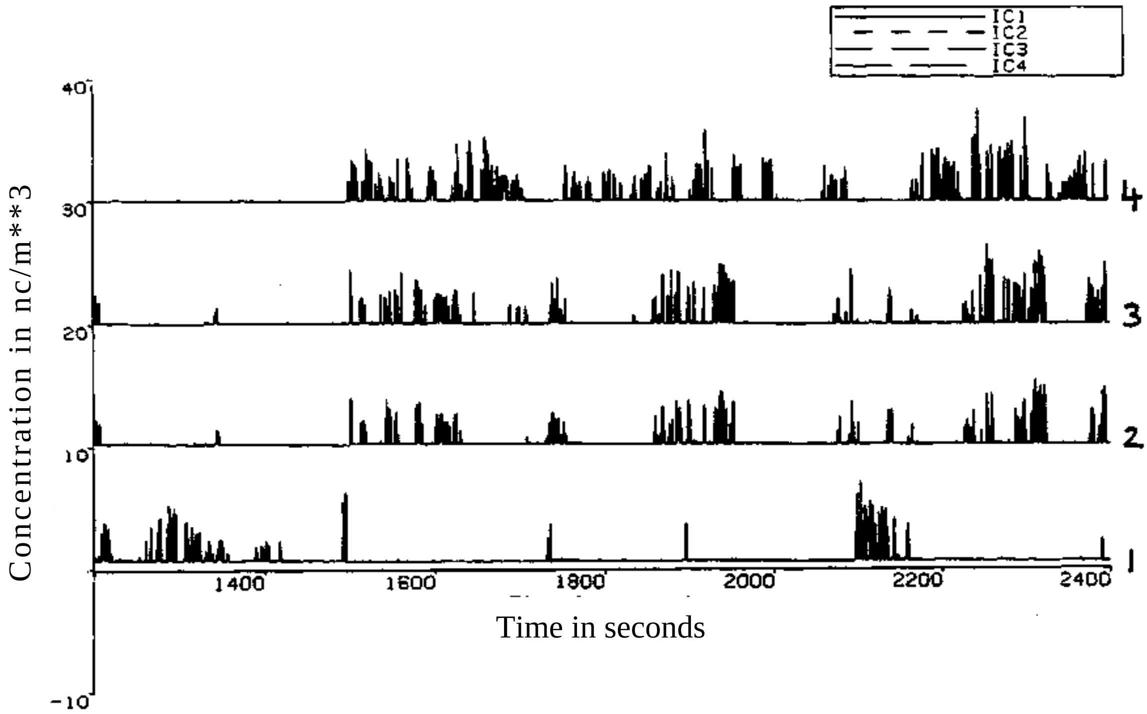
FigureC2 32



as 109.num

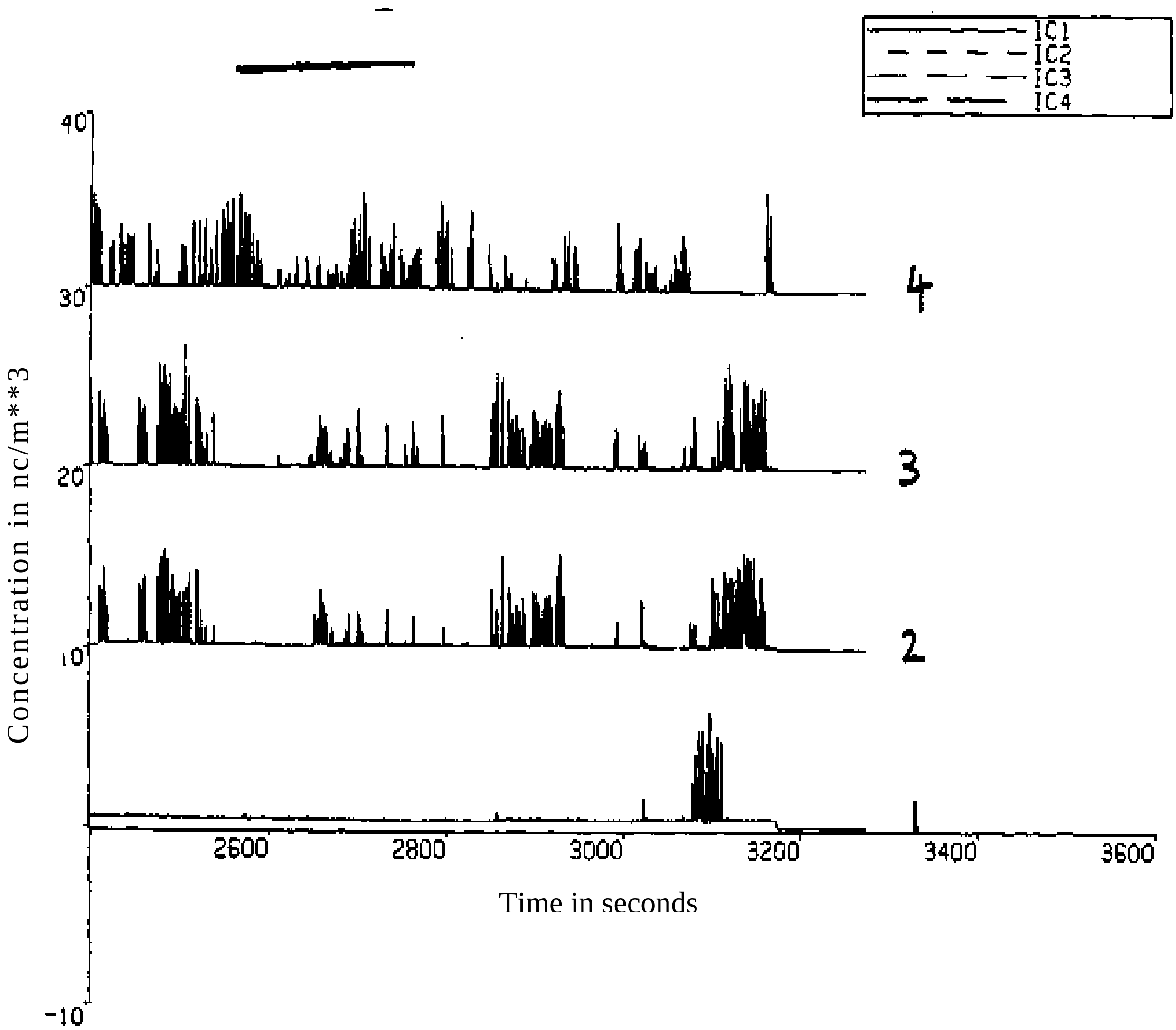
FIGURE C3





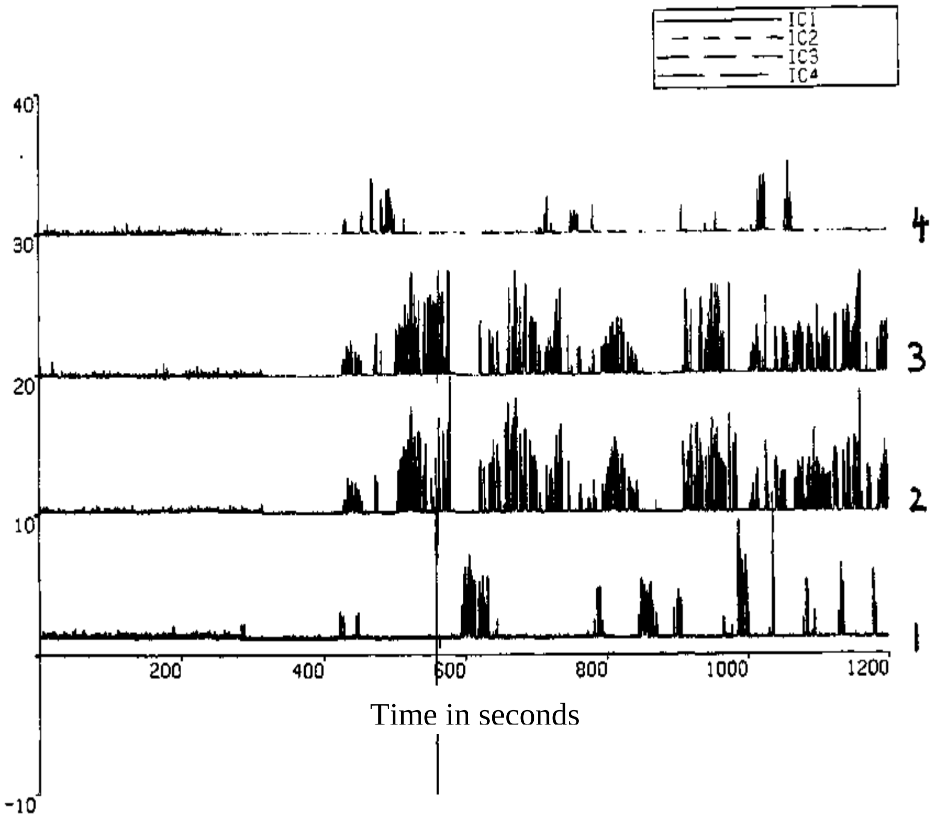
as 110.num

FIGURE C3



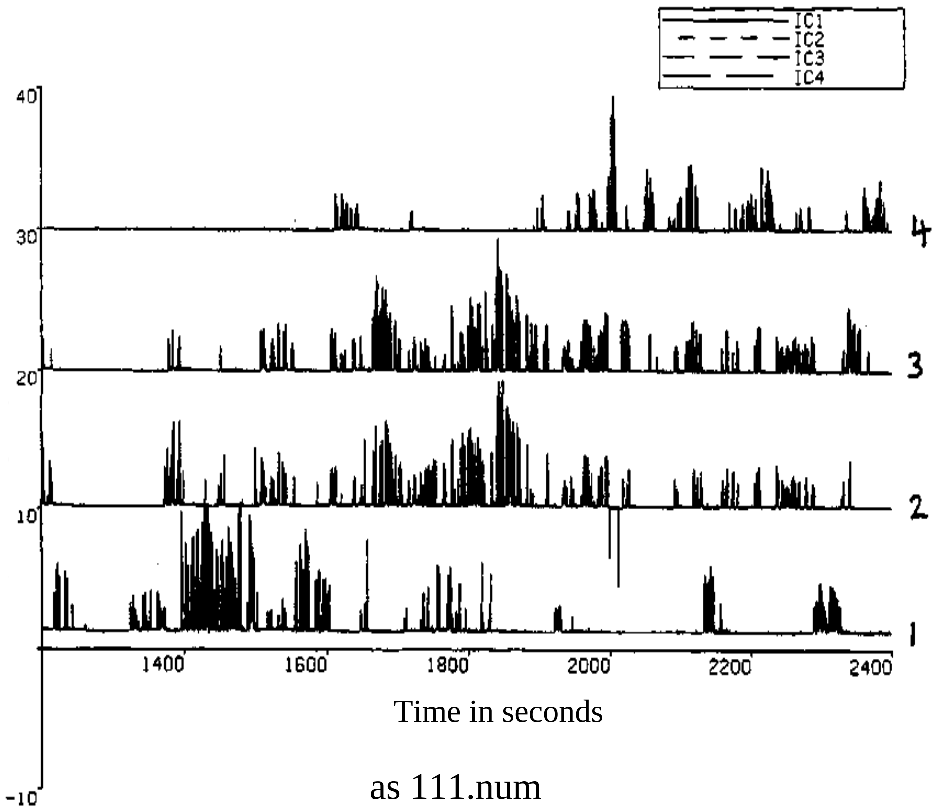
as 110.num

FIGURE C4



Concentration in nc/m**3

Time in seconds

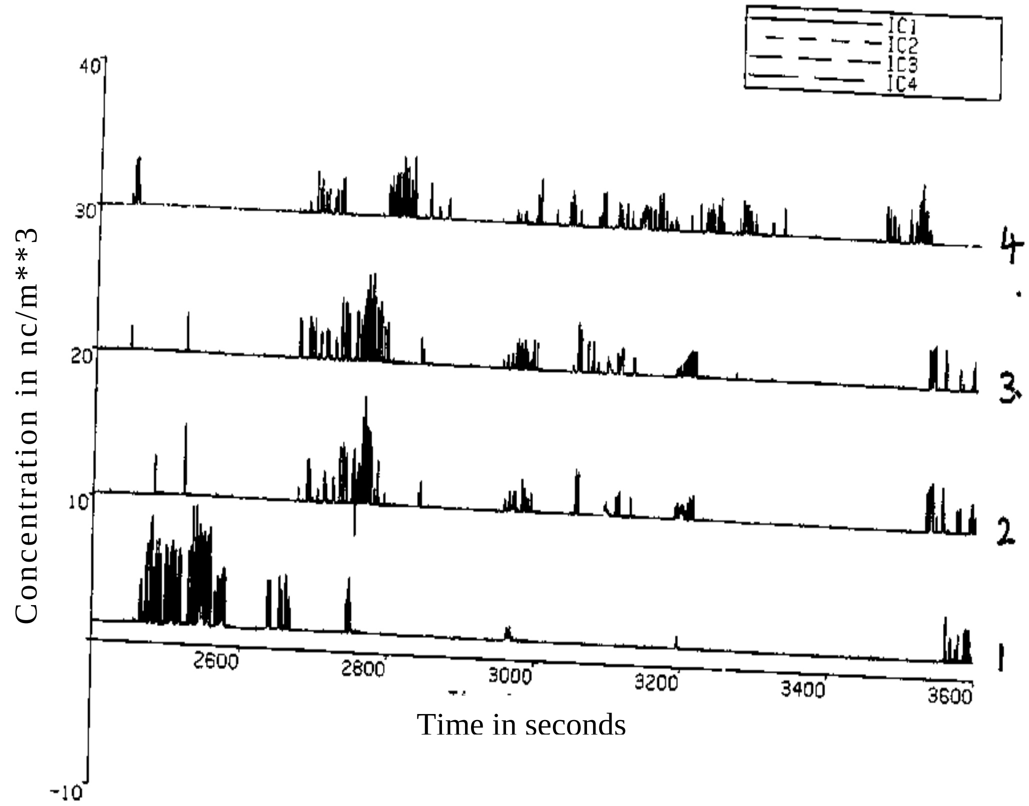


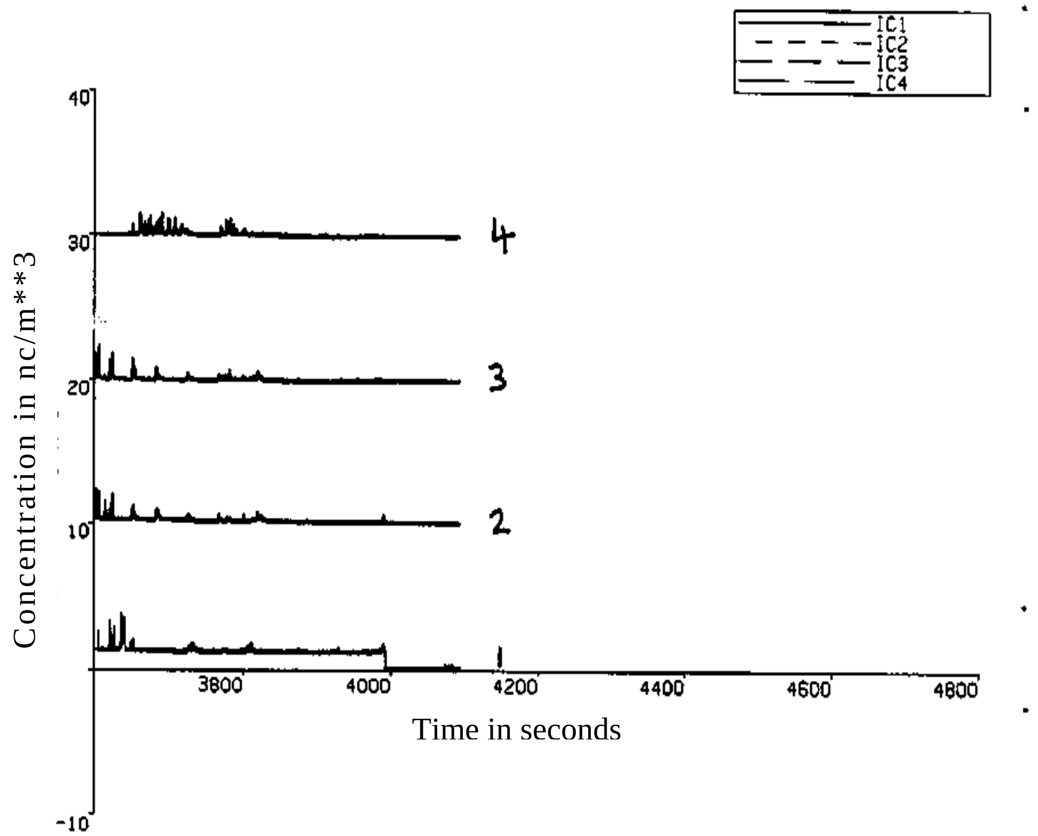
Concentration in nc/m**3

Time in seconds

as 111.num

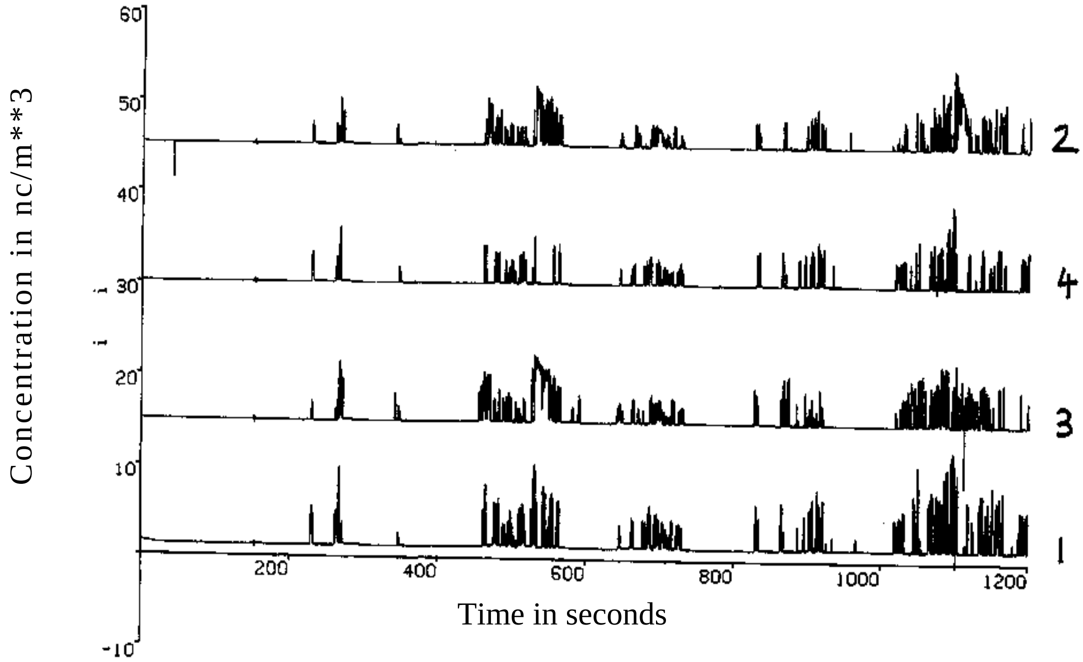
FIGURE C4

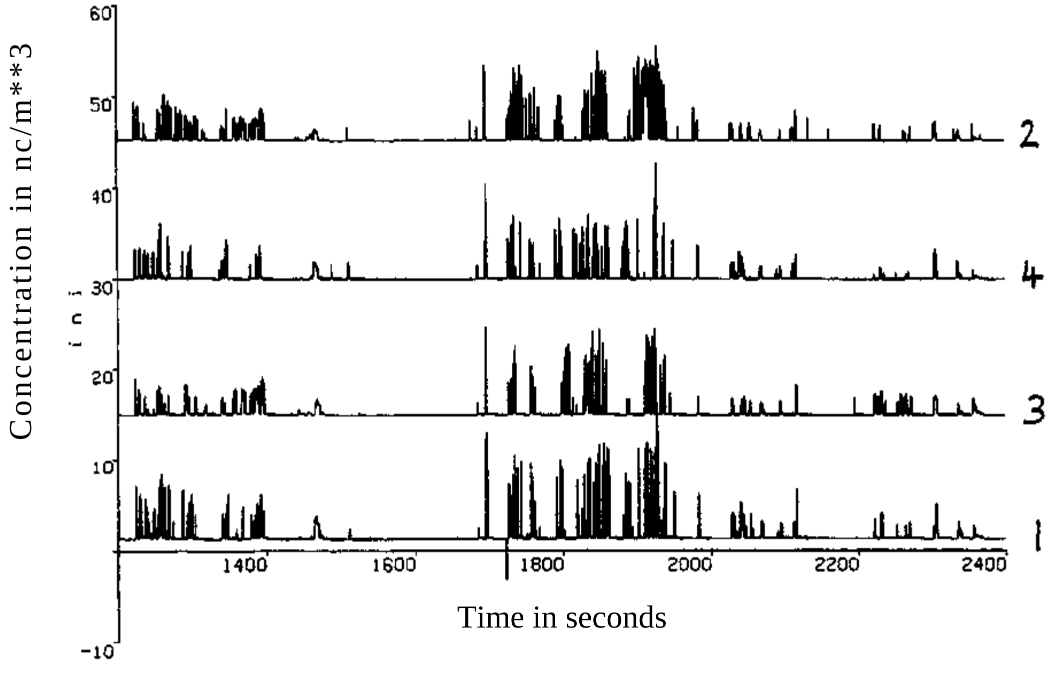




as 111.num

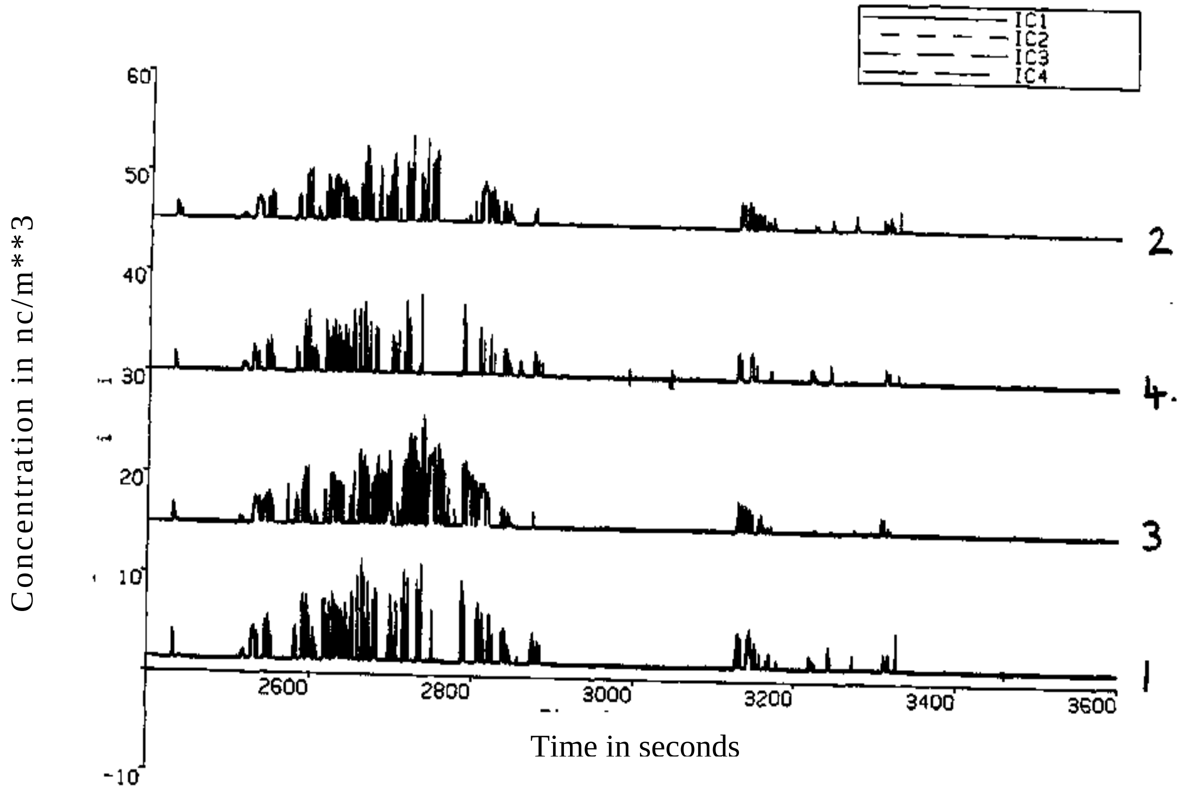
FIGURE C5

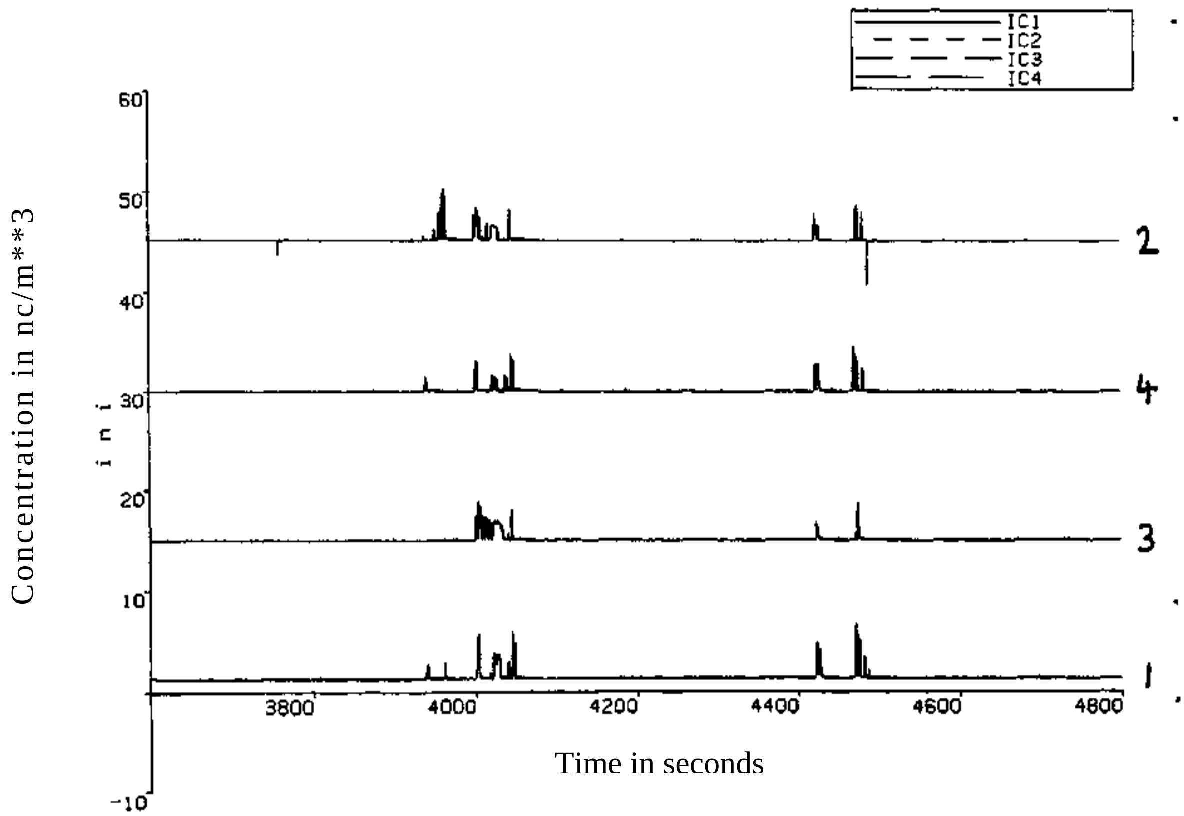




as 113.num

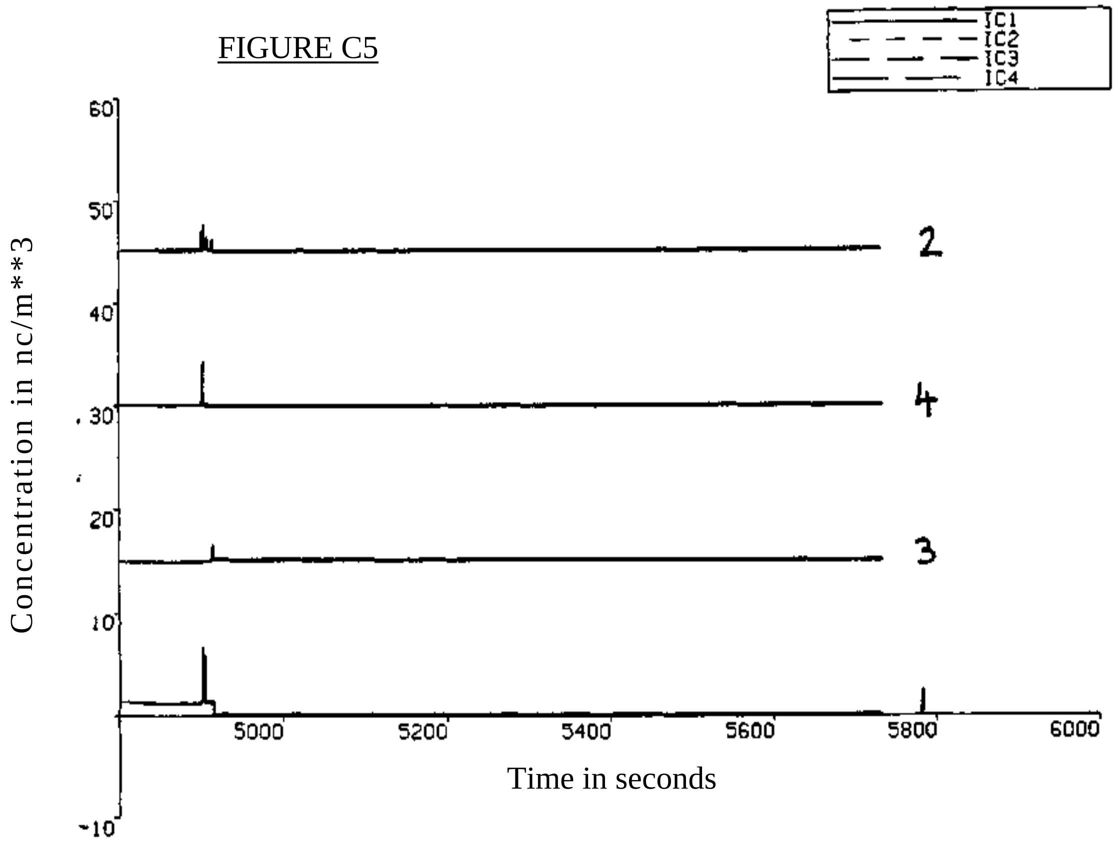
FIGURE C5



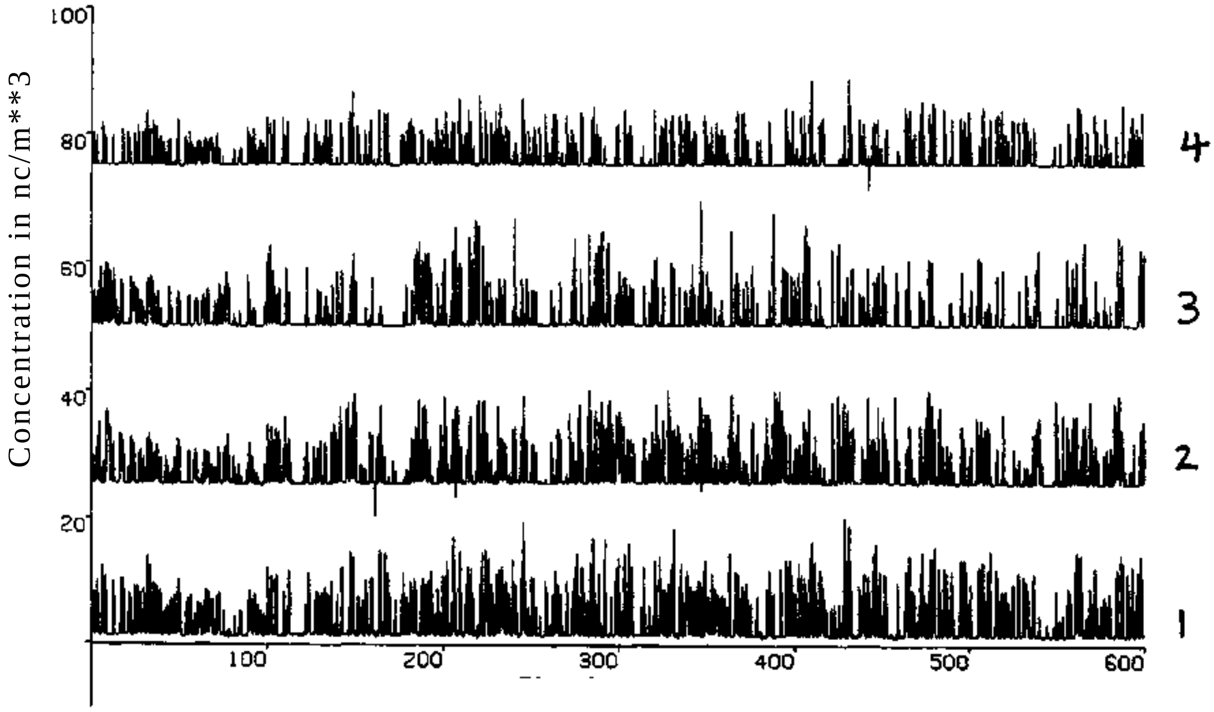


as 113.num

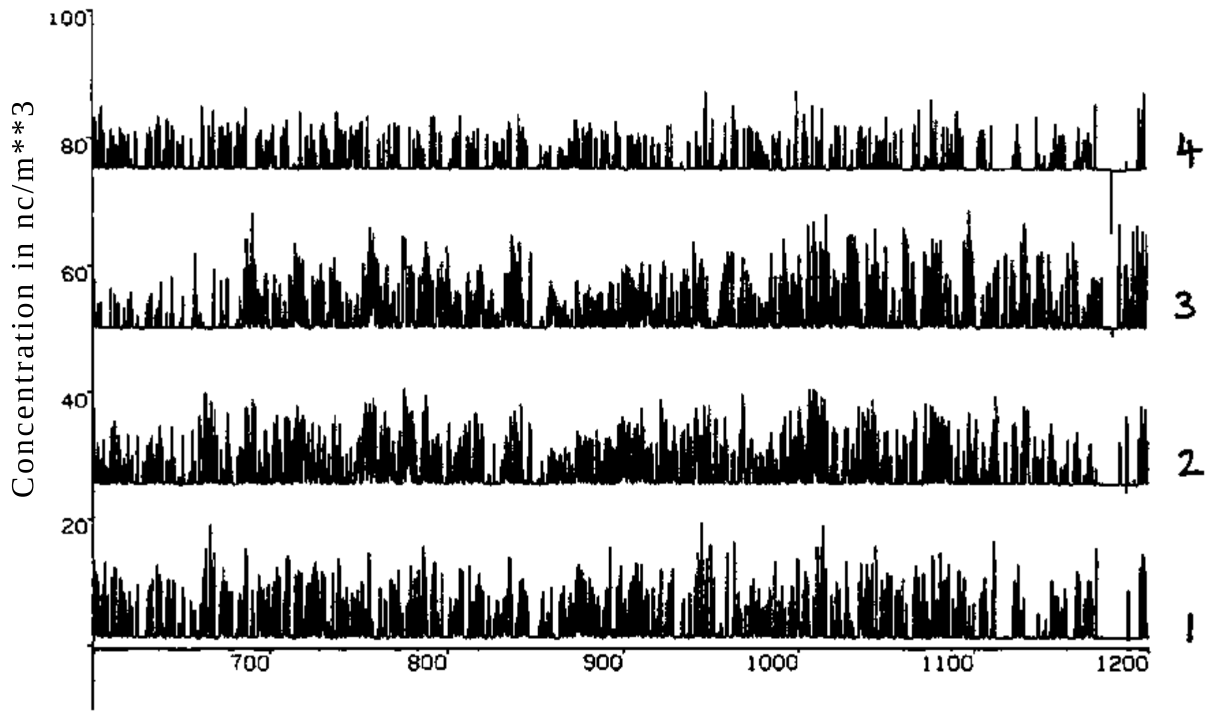
FIGURE C5



as 113.num



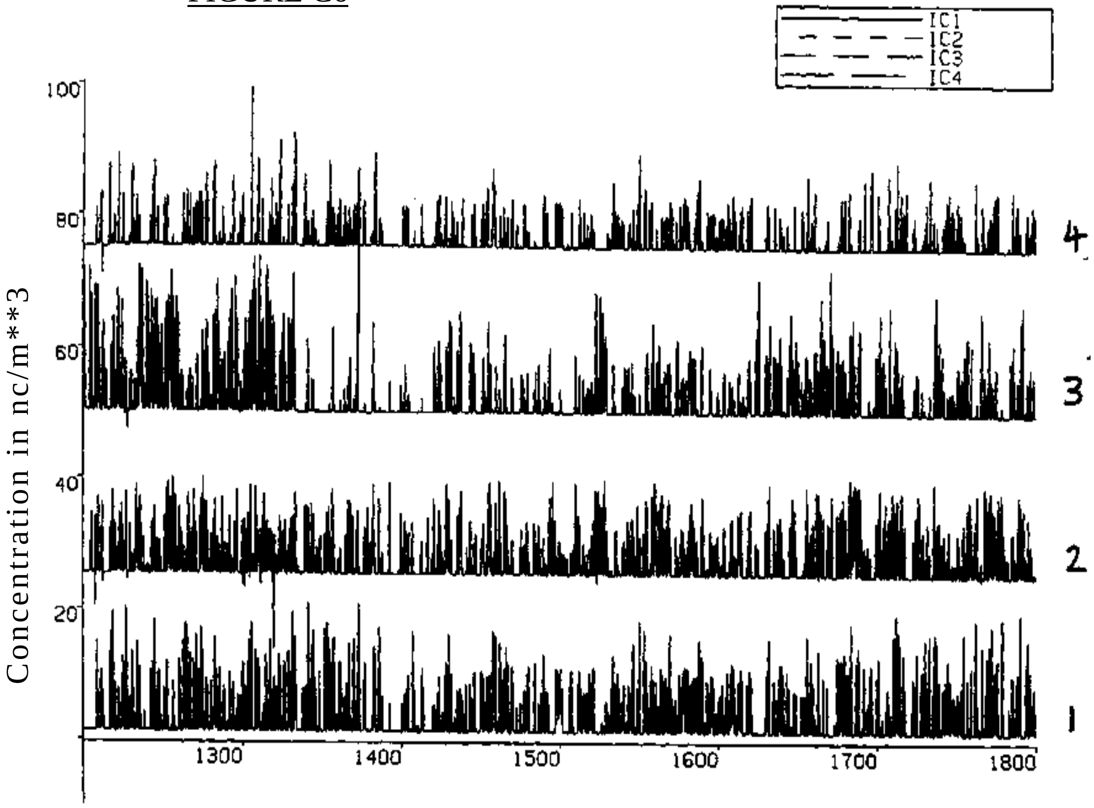
Time in seconds



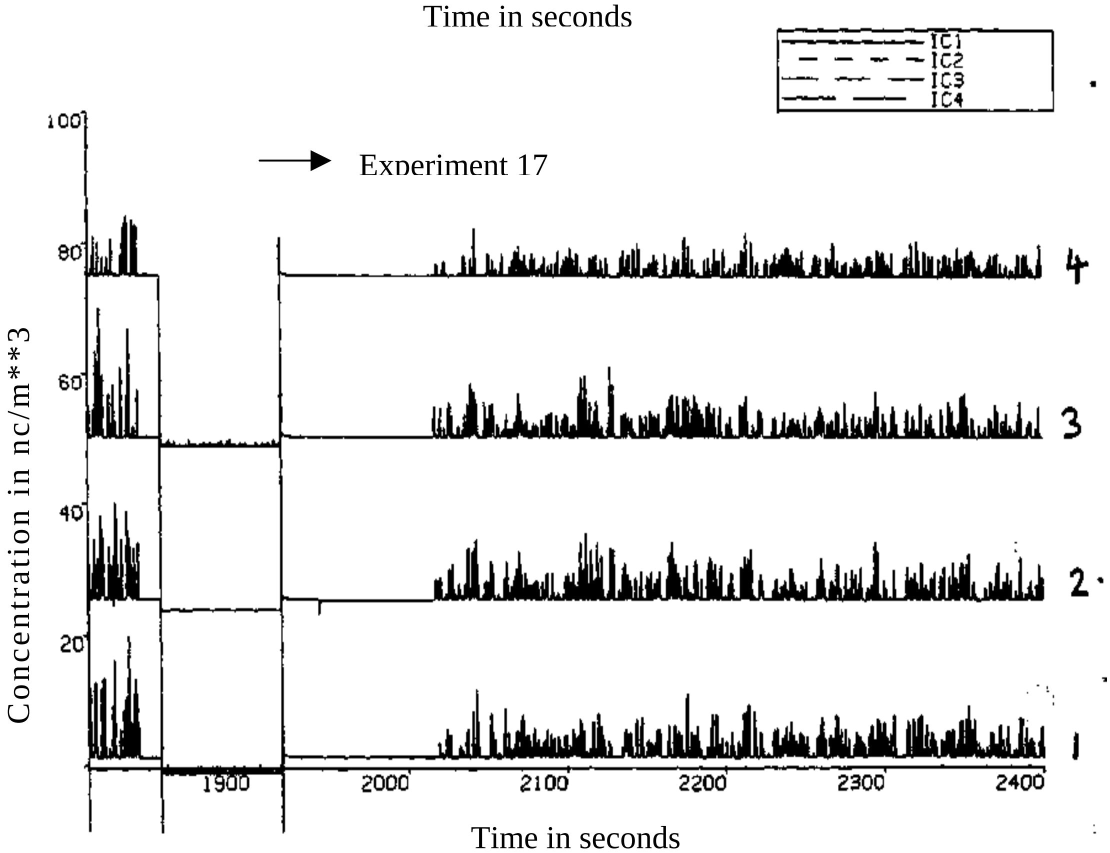
Time in seconds

as 116.num

FIGURE C6



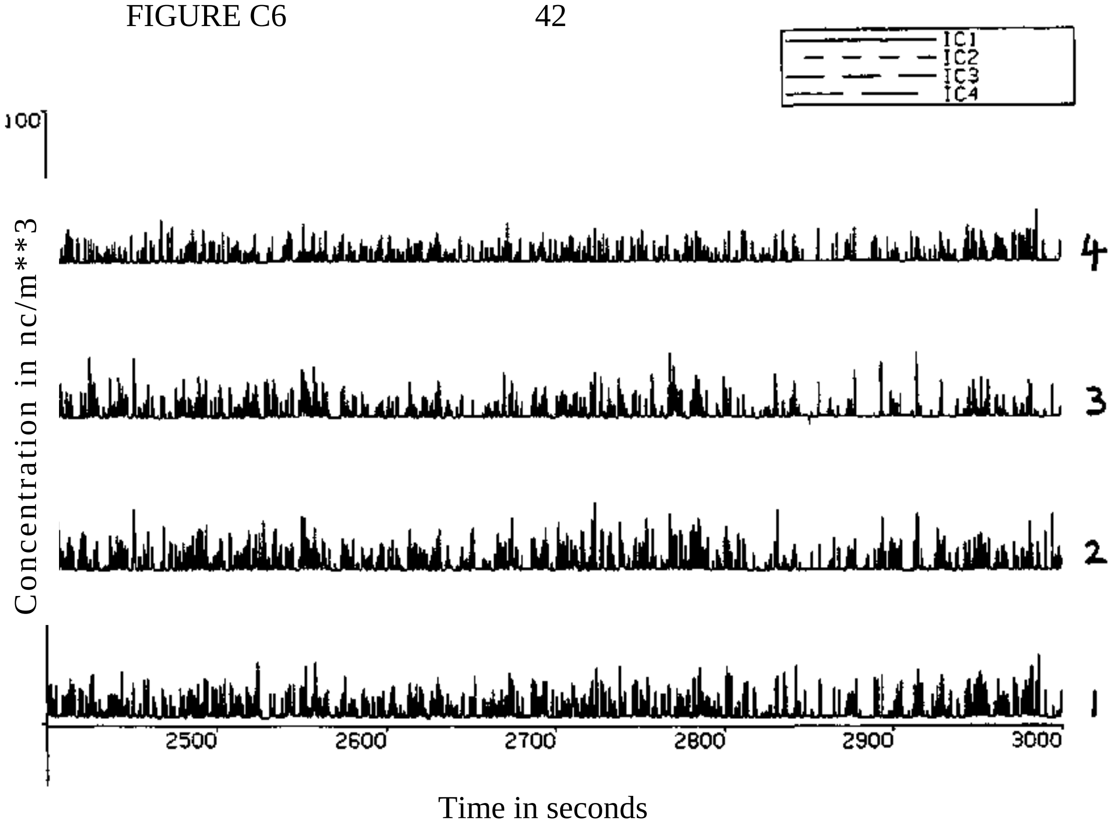
Time in seconds

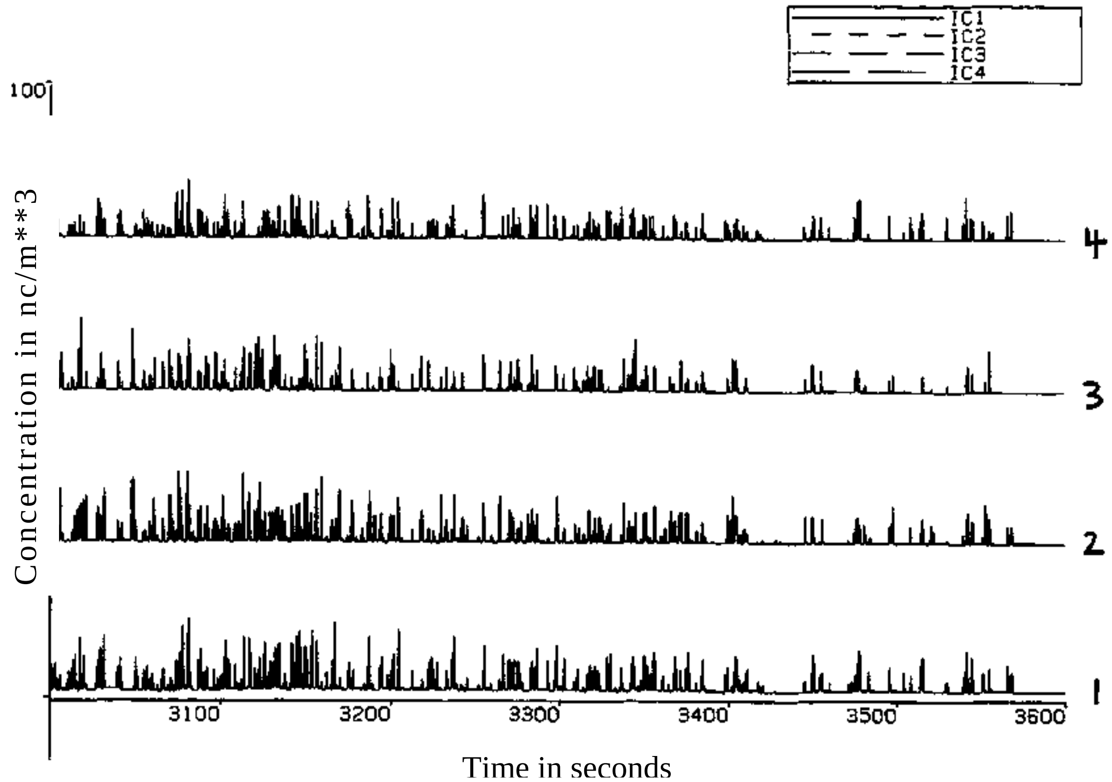


Time in seconds

as 116.num

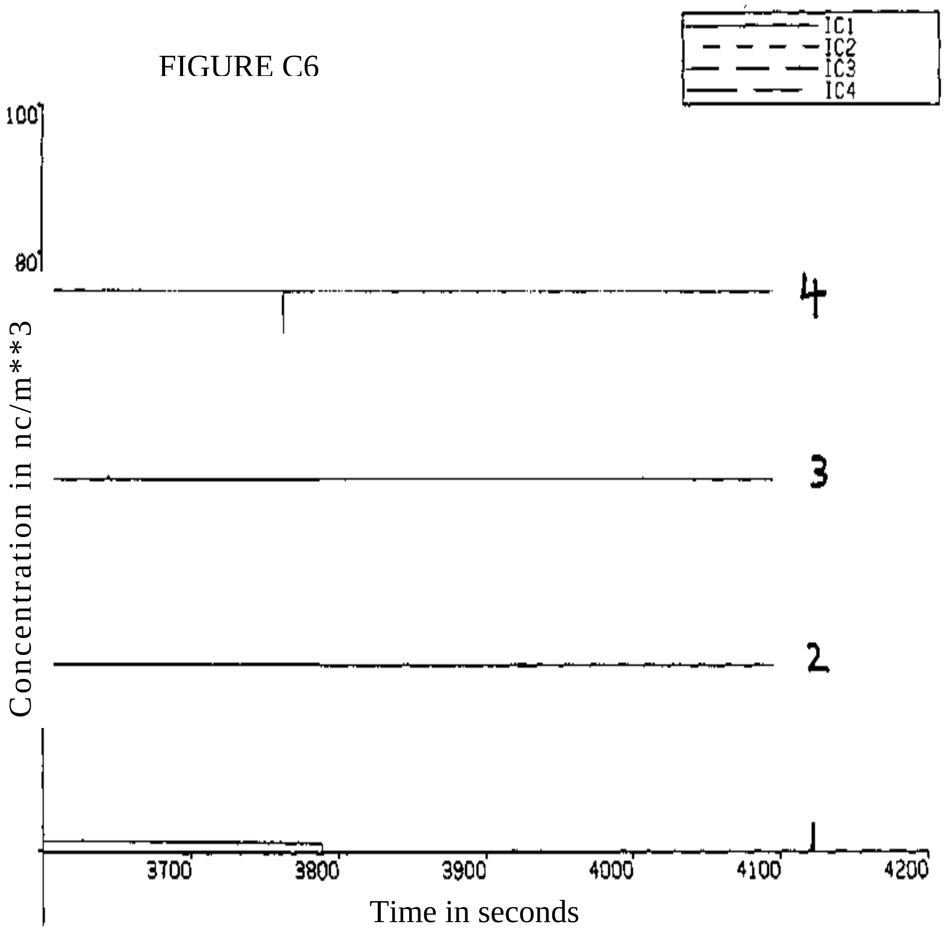
FIGURE C6





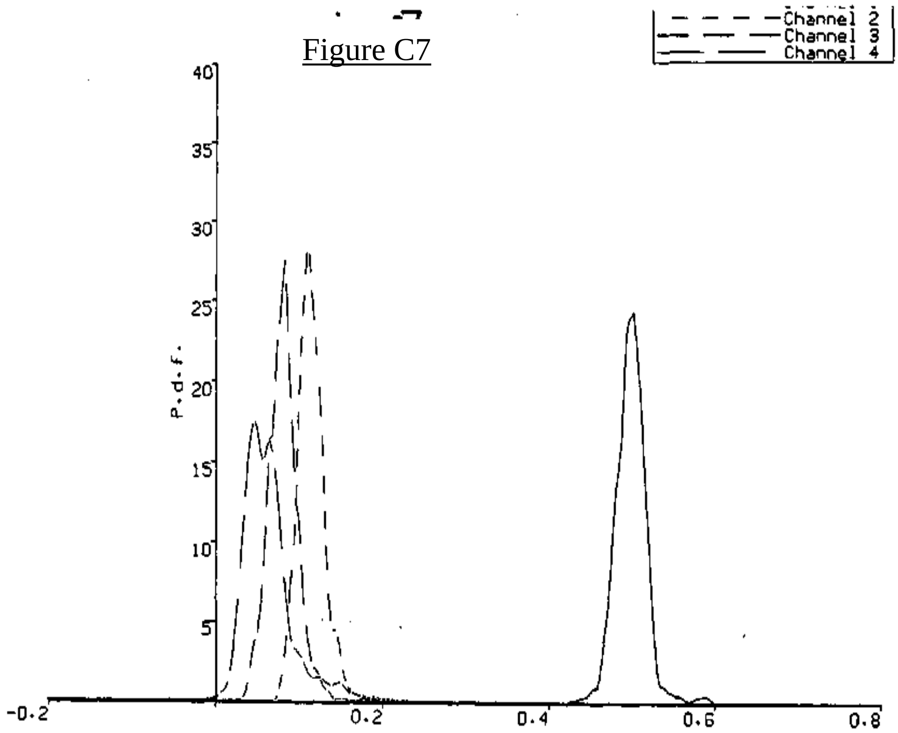
as 116.num

FIGURE C6



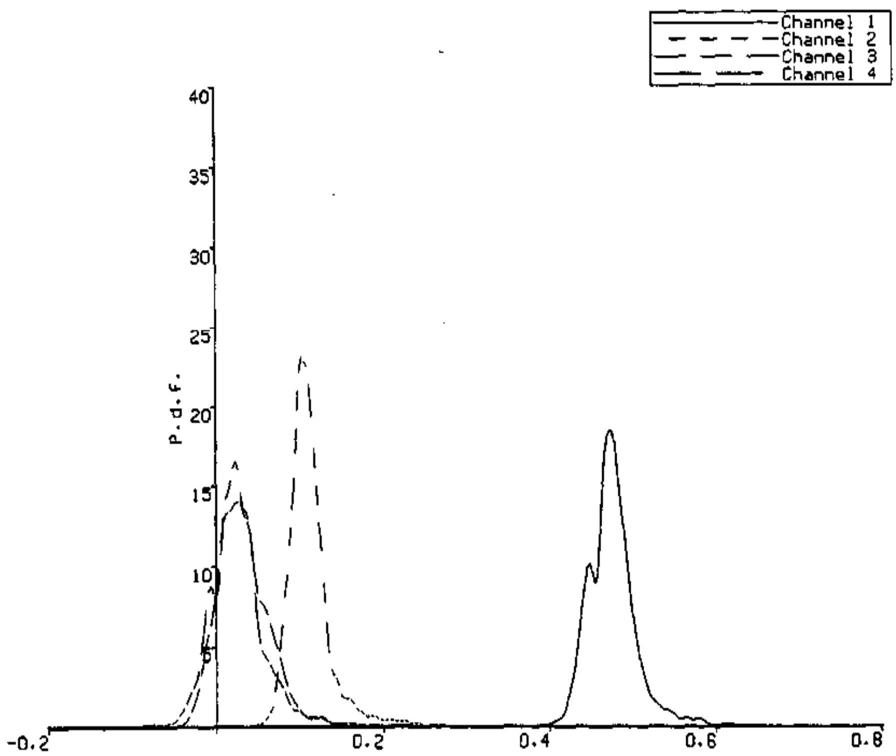
as 116.num

Figure C7



Concentration

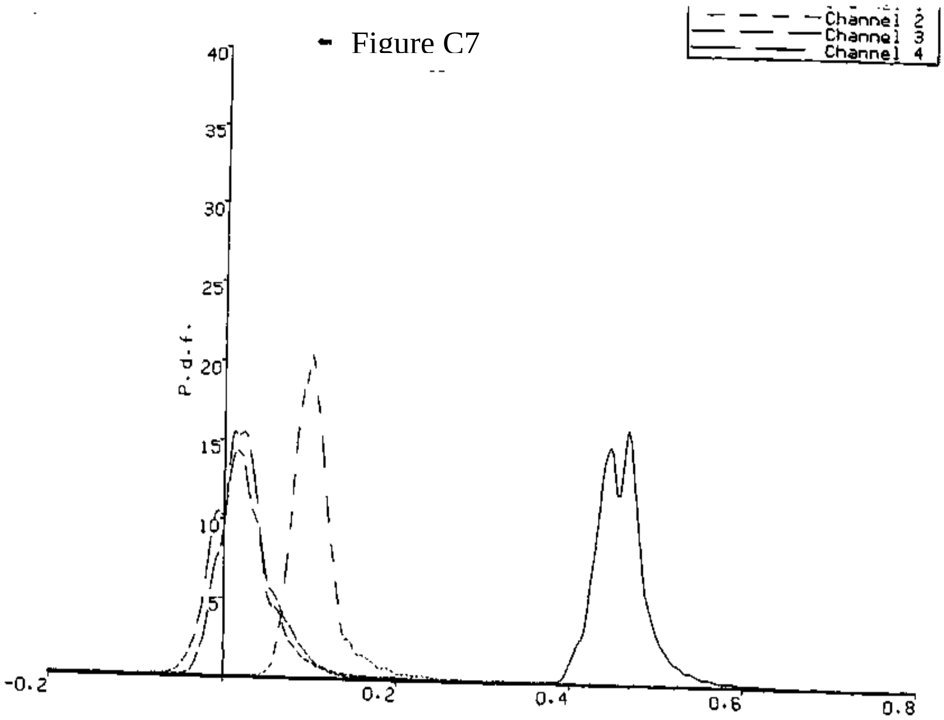
pdf m 1024 as 108 noise



Concentration

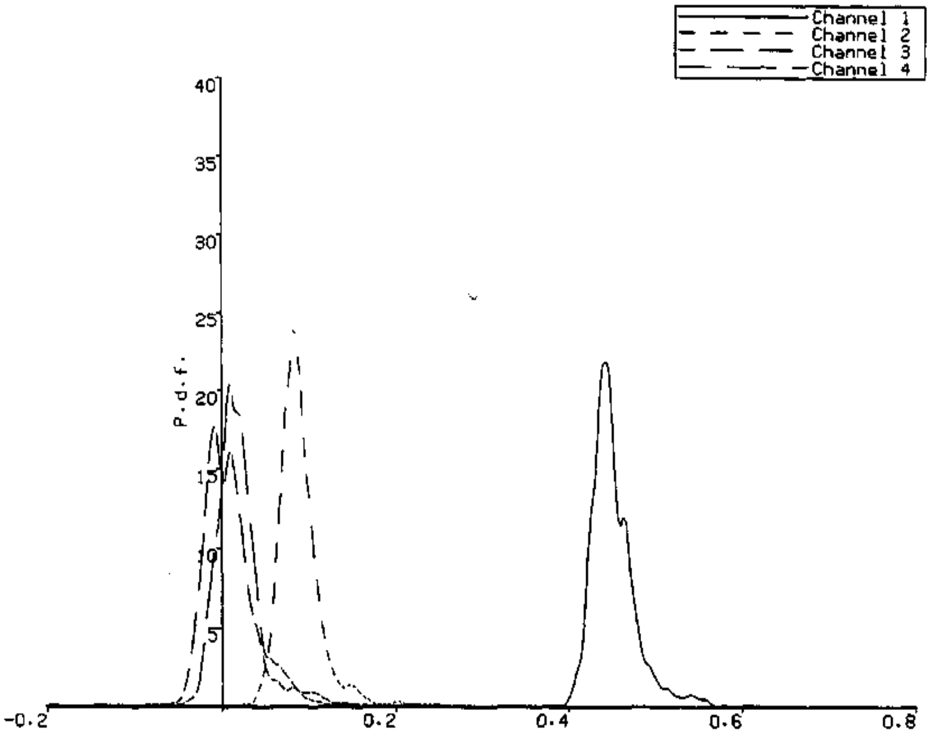
pdf m 1024 as 109 noise

Figure C7



Concentration

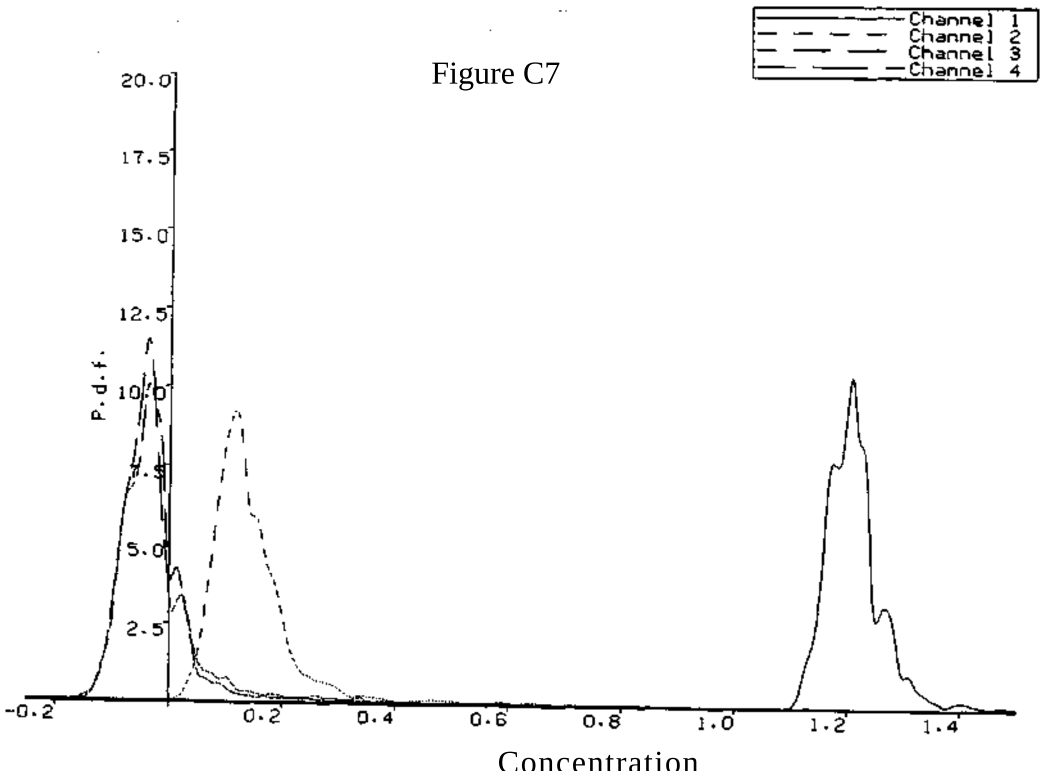
pdf m 1024 as 110 noise



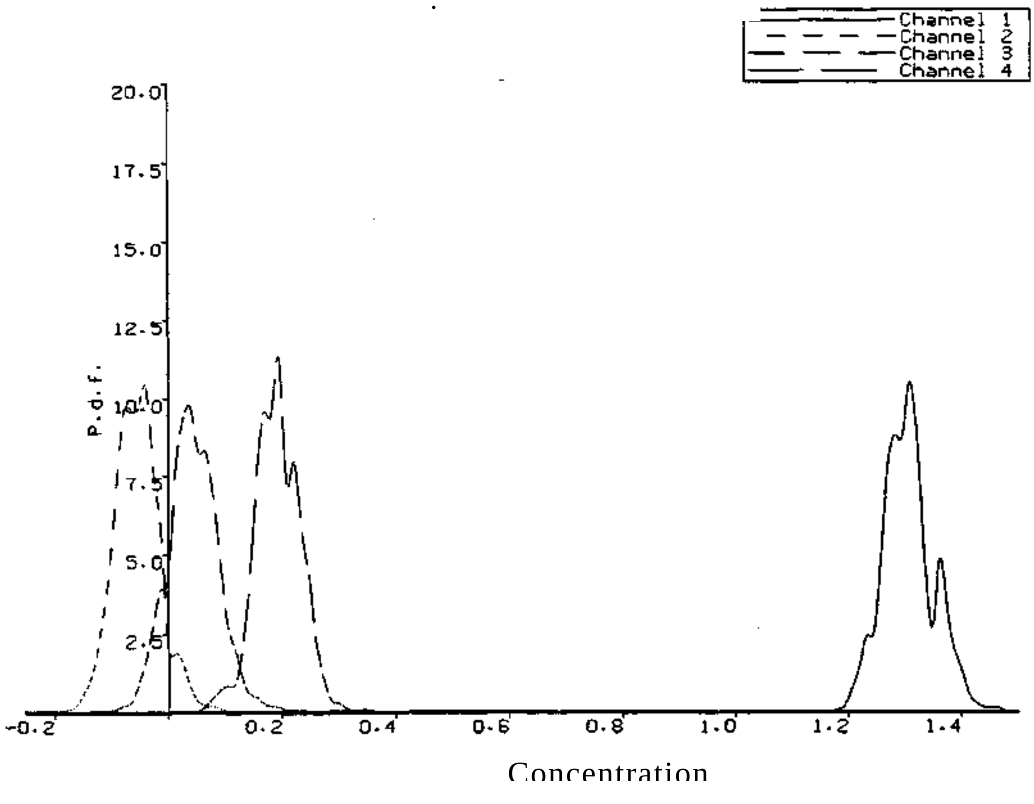
Concentration

pdf m 1024 as 111 noise2

Figure C7

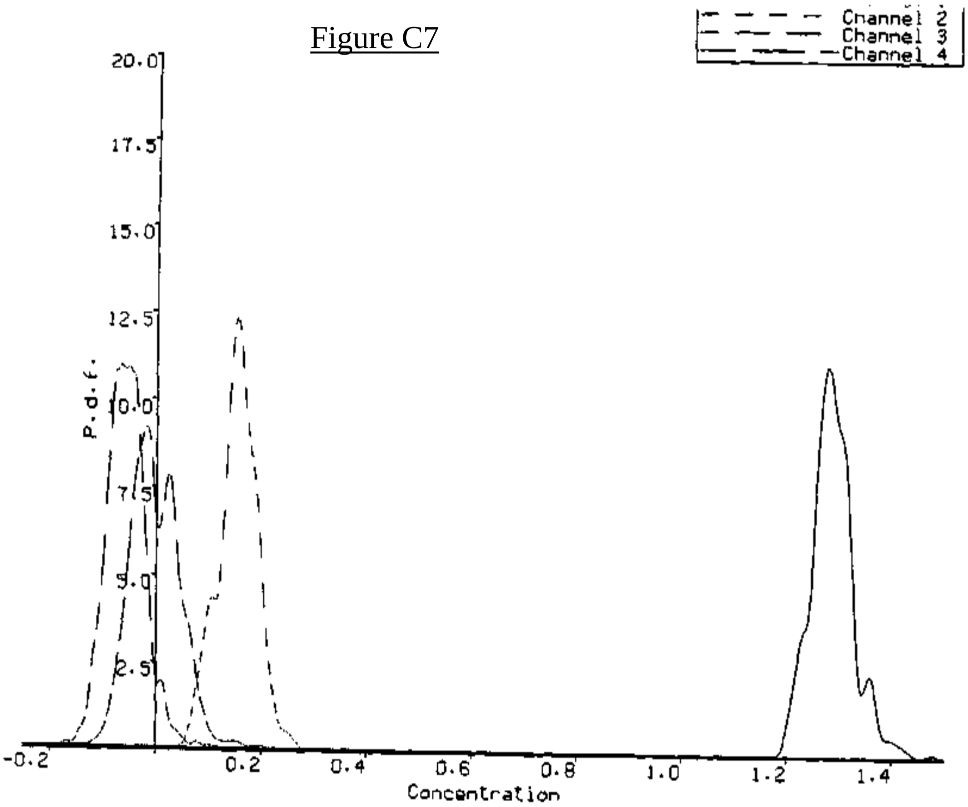


pdf m 1024 as 111

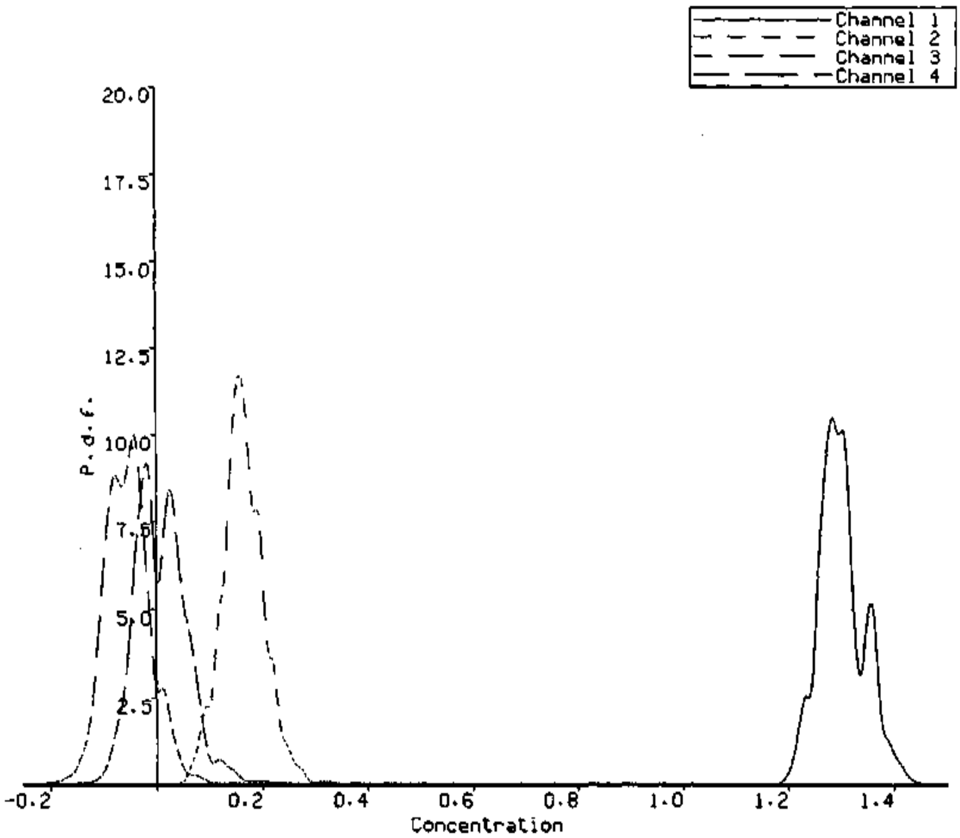


pdf m 1024 as 113 noise2

Figure C7

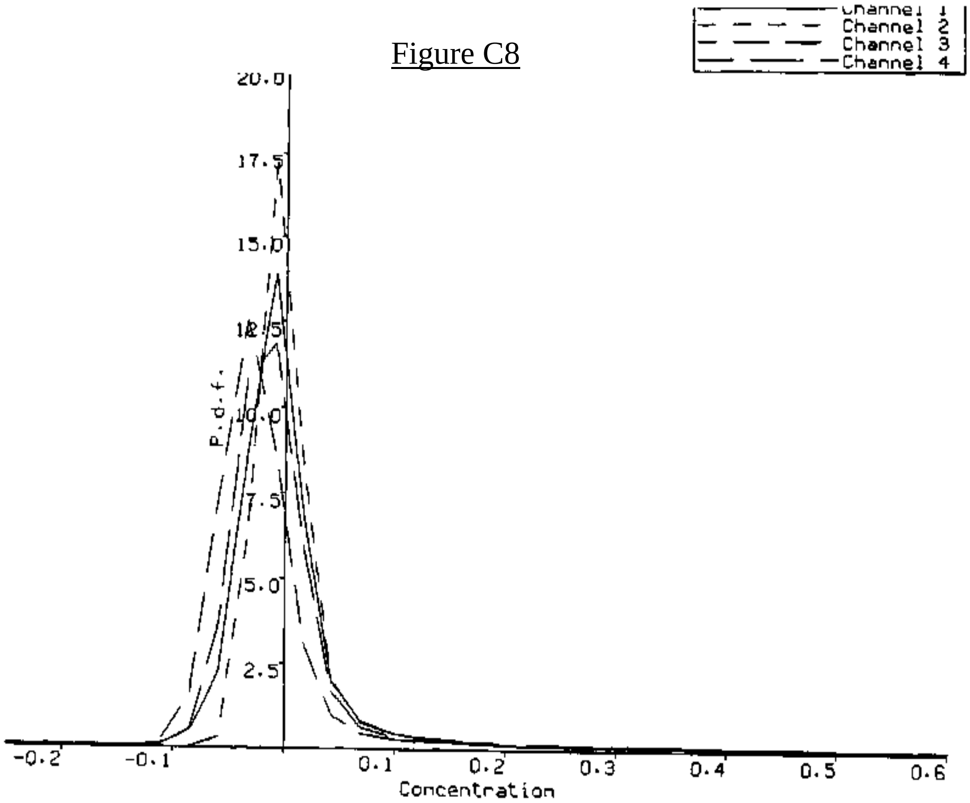


pdfm1024asl16noise

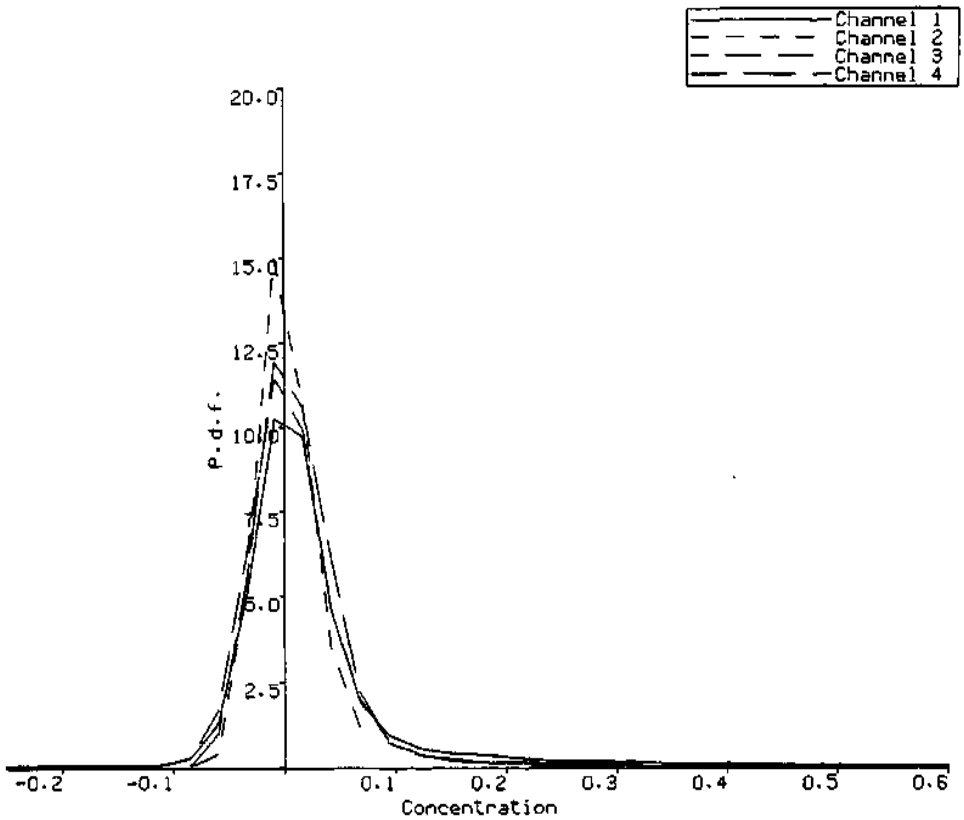


pdfm1024asl17noise

Figure C8

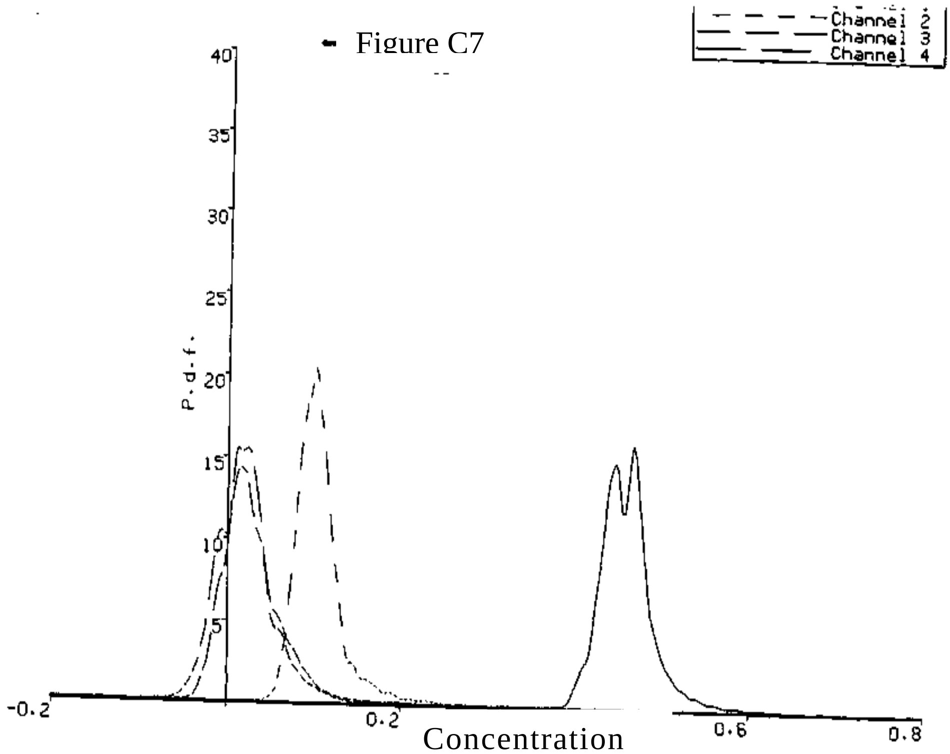


pdfm1024asl08



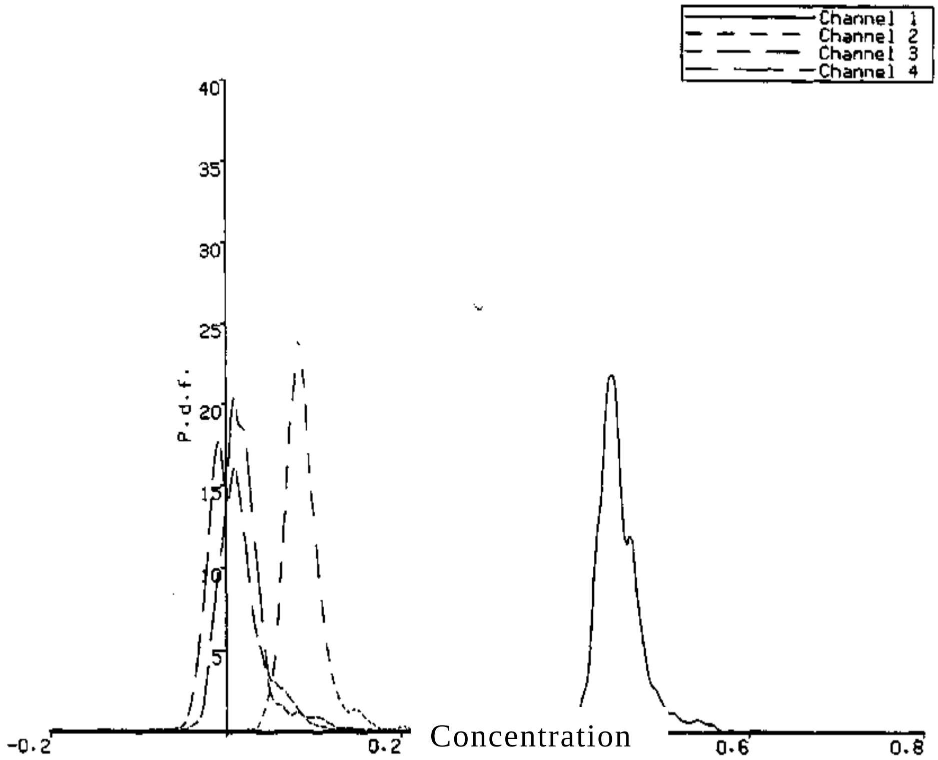
pdfm1024asl09

Figure C7



Concentration

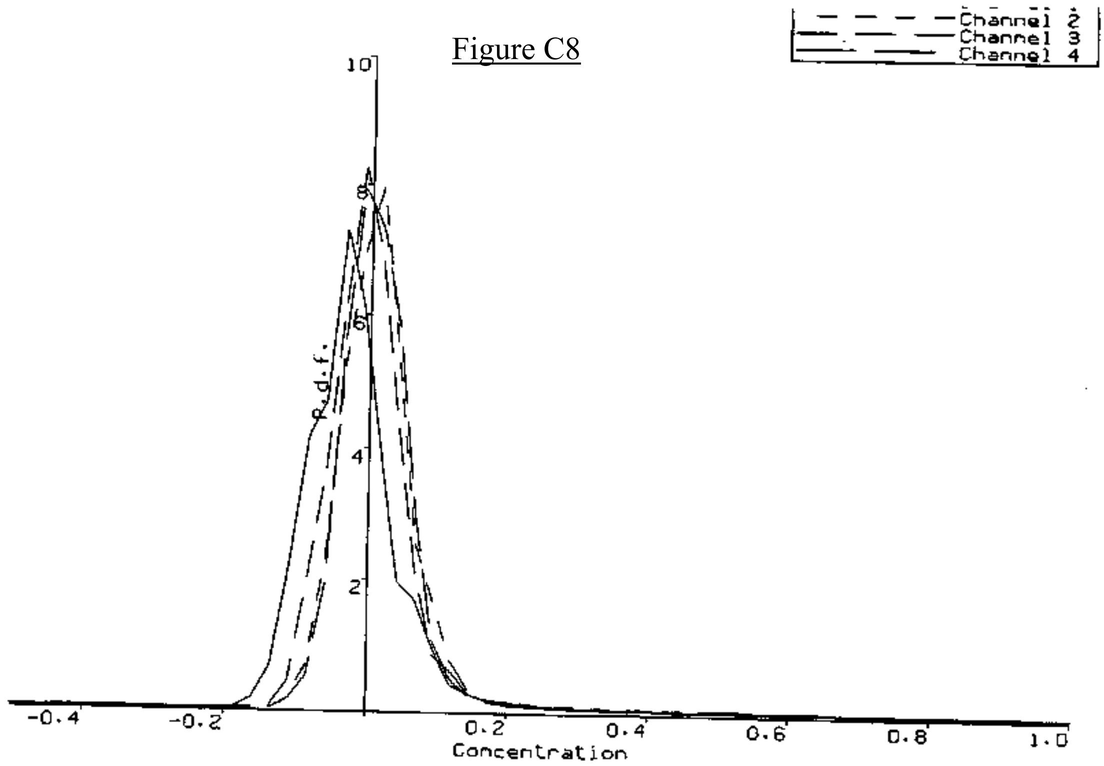
pdf m 1024 as 110 noise



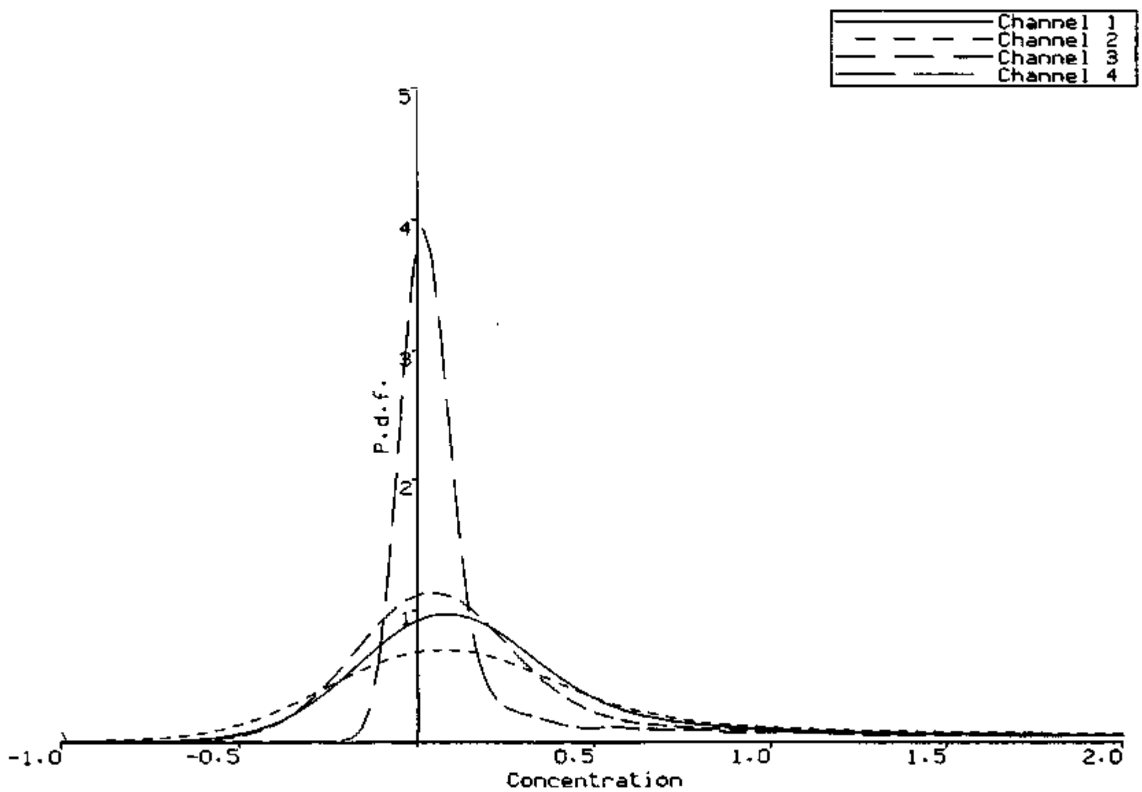
Concentration

pdf m 1024 as 111 noise2

Figure C8

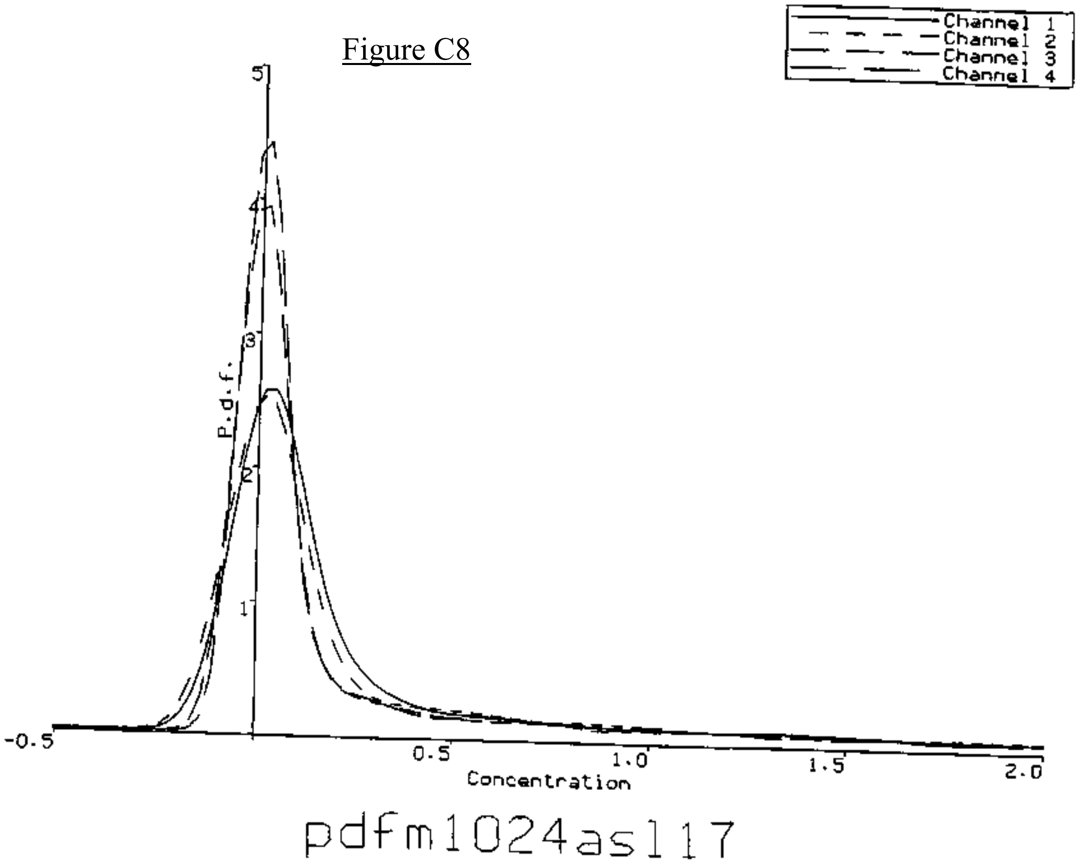


pdfm1024asl13



pdfm1024asl16

Figure C8


Channel 1
Channel 2
Channel 3
Channel 4
5
4
3
2
1
P.d.f.
-0.5
0.5
1.0
1.5
2.0
Concentration
pdfm1024asl17

Intermittency

The intermittency π has been calculated as:

$$\pi = \Pr ob(\Gamma > \theta_\Gamma)$$

where Γ is the concentration (after removal of the noise means) and θ_Γ is a threshold. The results, for various choices of θ_Γ, are shown in Figure C9.

To aid interpretation, the following table gives the ratios of $(\overline{c_2})^{1/2}$ for the actual experiment (σ) to that for the appropriate period of noise (σ_N), and the ratio of σ_N, to the bin size $\Delta\theta$ of the A-D conversion:

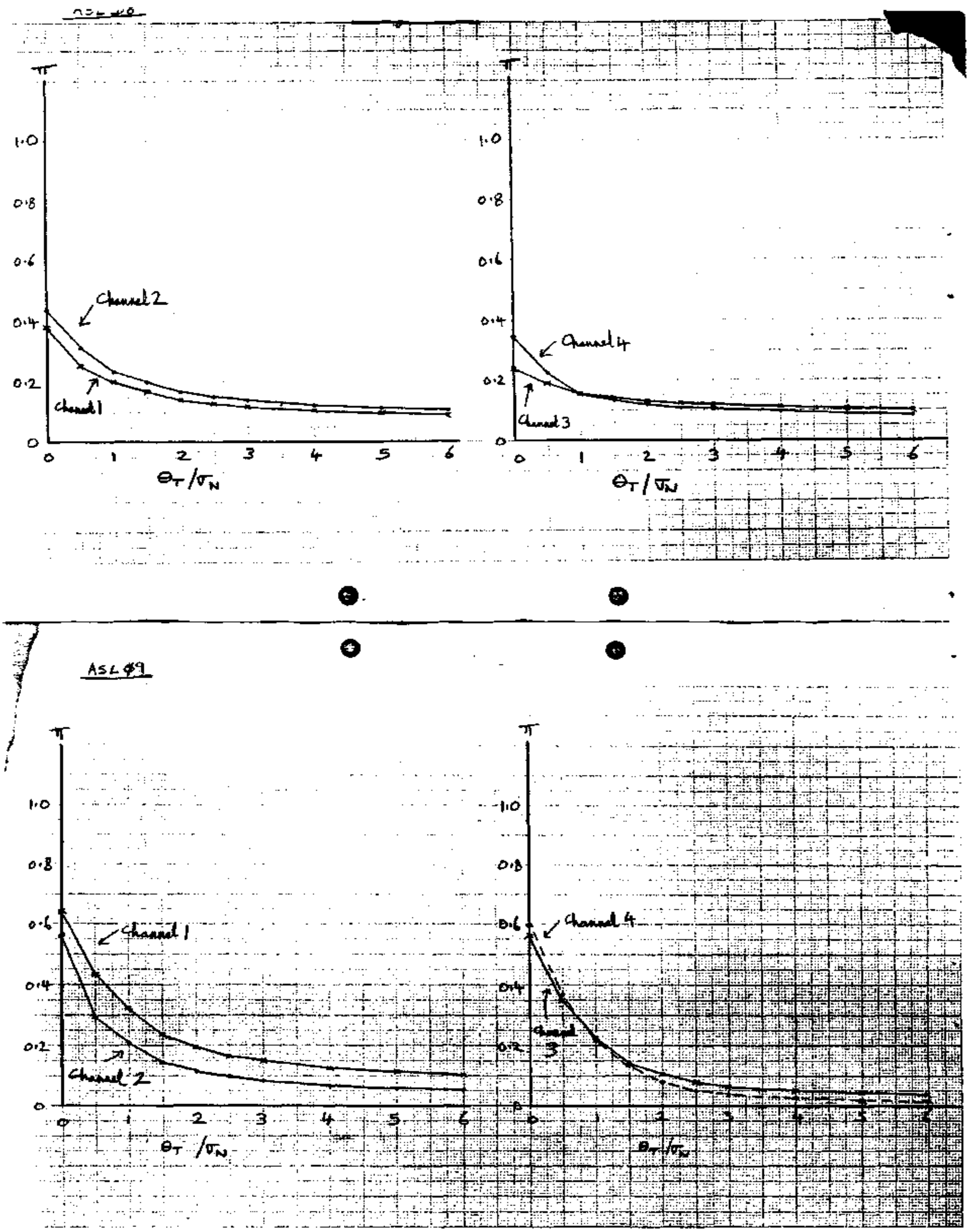
Experiment	8				9			
Channel	1	2	3	4	1	2	3	4
σ/σ_N	17.5	17.7	17.5	9.3	12.3	6.7	5.1	2.1
$\sigma_{N/\Delta\theta}$	5.0	4.1	4.5	6.7	7.0	6.5	7.5	7.5

Experiment	10				11			
Channel	1	2	3	4	1	2	3	4
σ/σ_N	13.9	18.6	18.1	18.5	13.3	11.5	10.1	7.8
$\sigma_{N/\Delta\theta}$	8.0	6.8	8.0	7.8	5.2	6.1	6.3	4.5

Experiment	13				16			
Channel	1	2	3	4	1	2	3	4
σ/σ_N	21.7	26.7	14.6	23.1	75.6	85.6	90.3	42.0
$\sigma_{N/\Delta\theta}$	3.8	3.1	3.5	3.3	3.2	2.8	2.8	3.5

Experiment	17			
Channel	1	2	3	4
σ/σ_N	34.1	40.0	35.5	22.0
$\sigma_{N/\Delta\theta}$	3.2	2.8	2.8	3.5

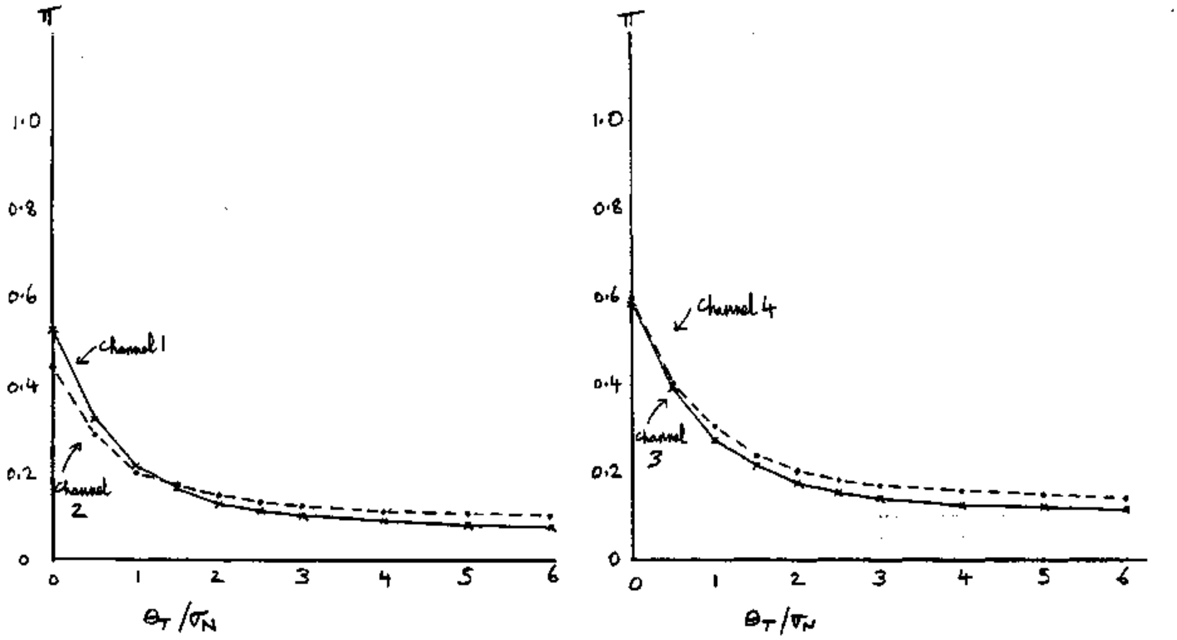
Figure C9



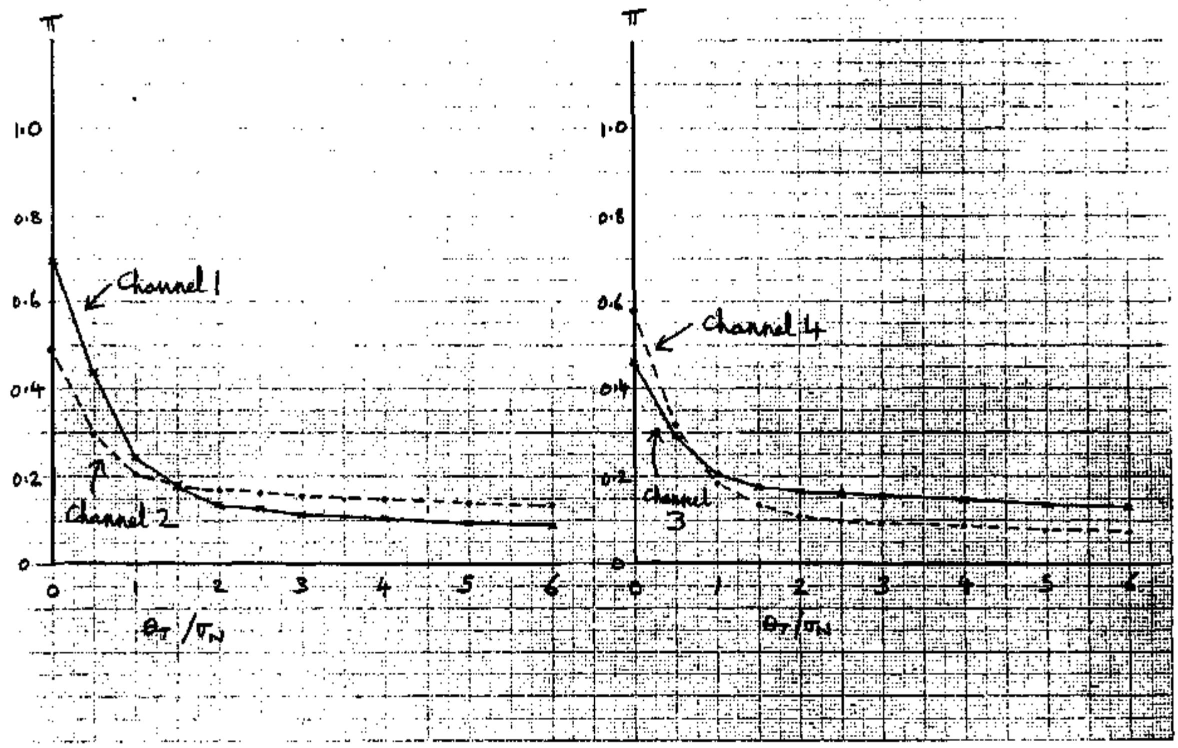
π
1.0
0.8
0.6
0.4
0.2
0
Channel 2
Channel 1
0 1 2 3 4 5 6
θT / σN
π
1.0
0.8
0.6
0.4
0.2
0
Channel 4
Channel 3
0 1 2 3 4 5 6
θT / σN
ASL Ø9
π
1.0
0.8
0.6
0.4
0.2
0
Channel 1
Channel 2
0 1 2 3 4 5 6
θT / σN
π
1.0
0.8
0.6
0.4
0.2
0
Channel 4
Channel 3
θT / σN

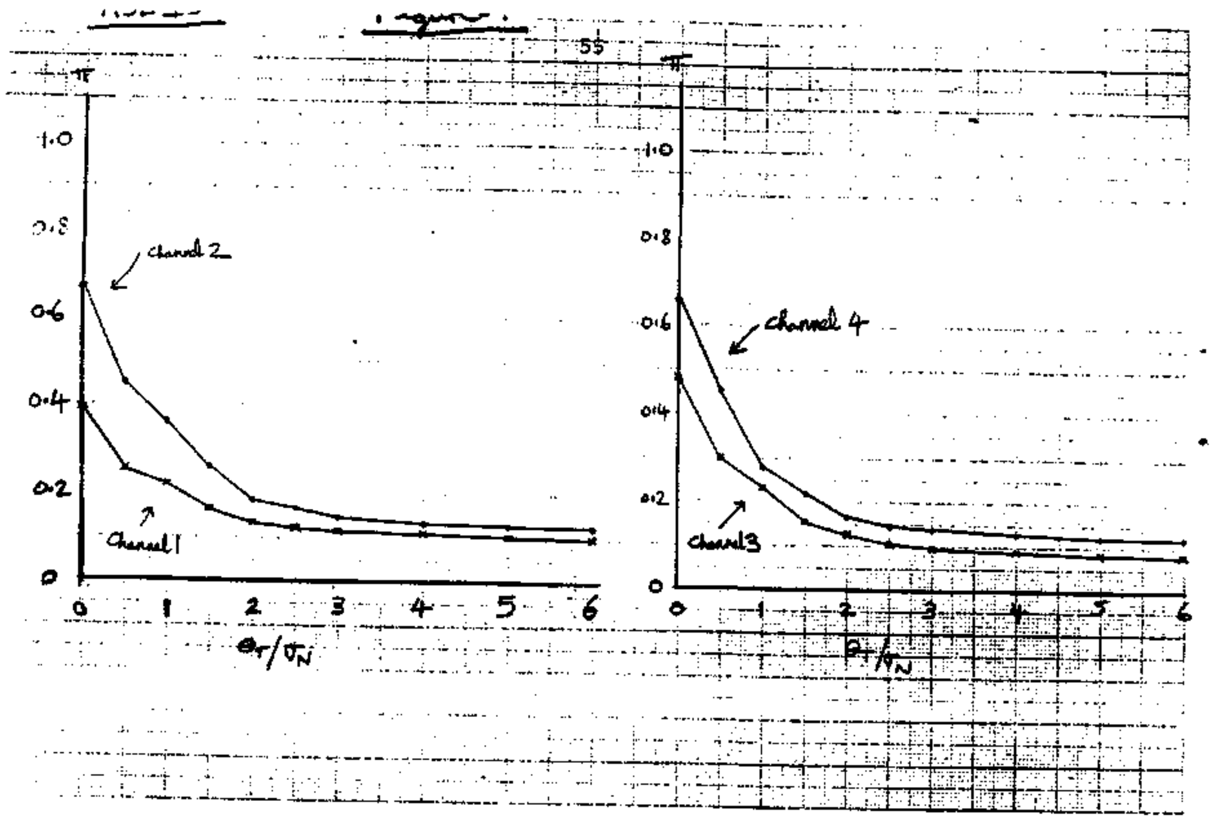
Figure C9

ASL 10



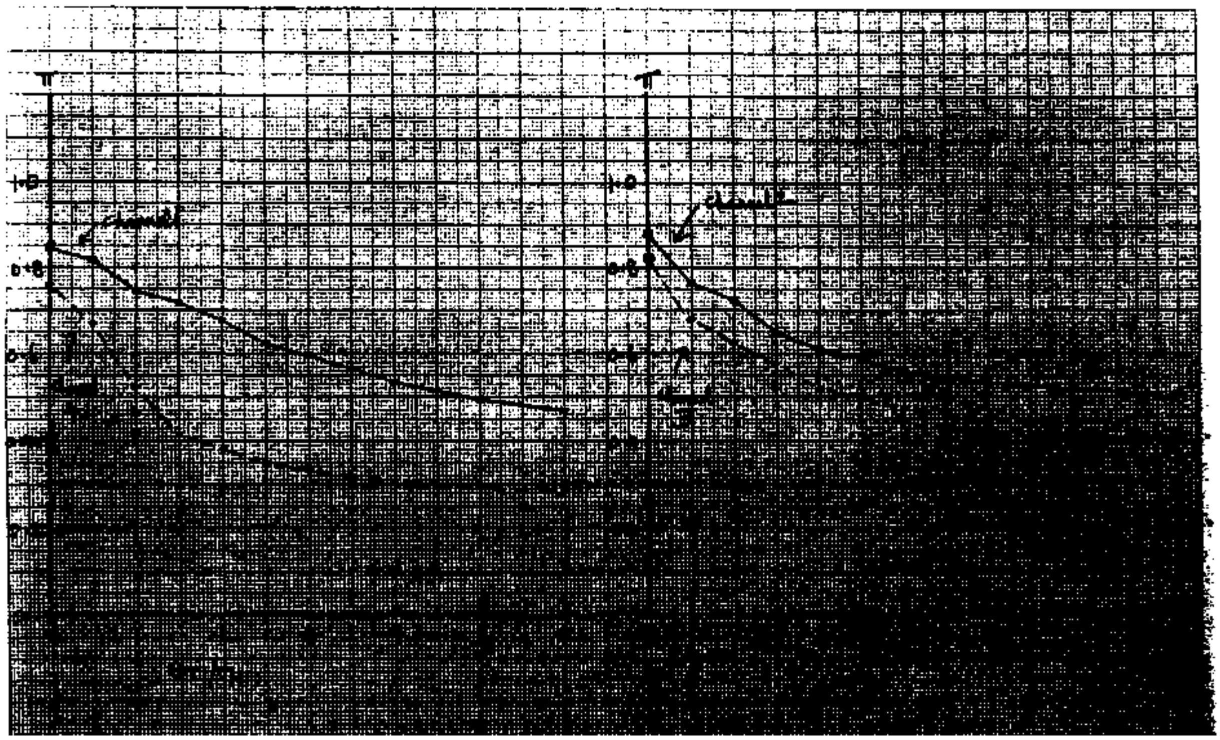
ASL 11





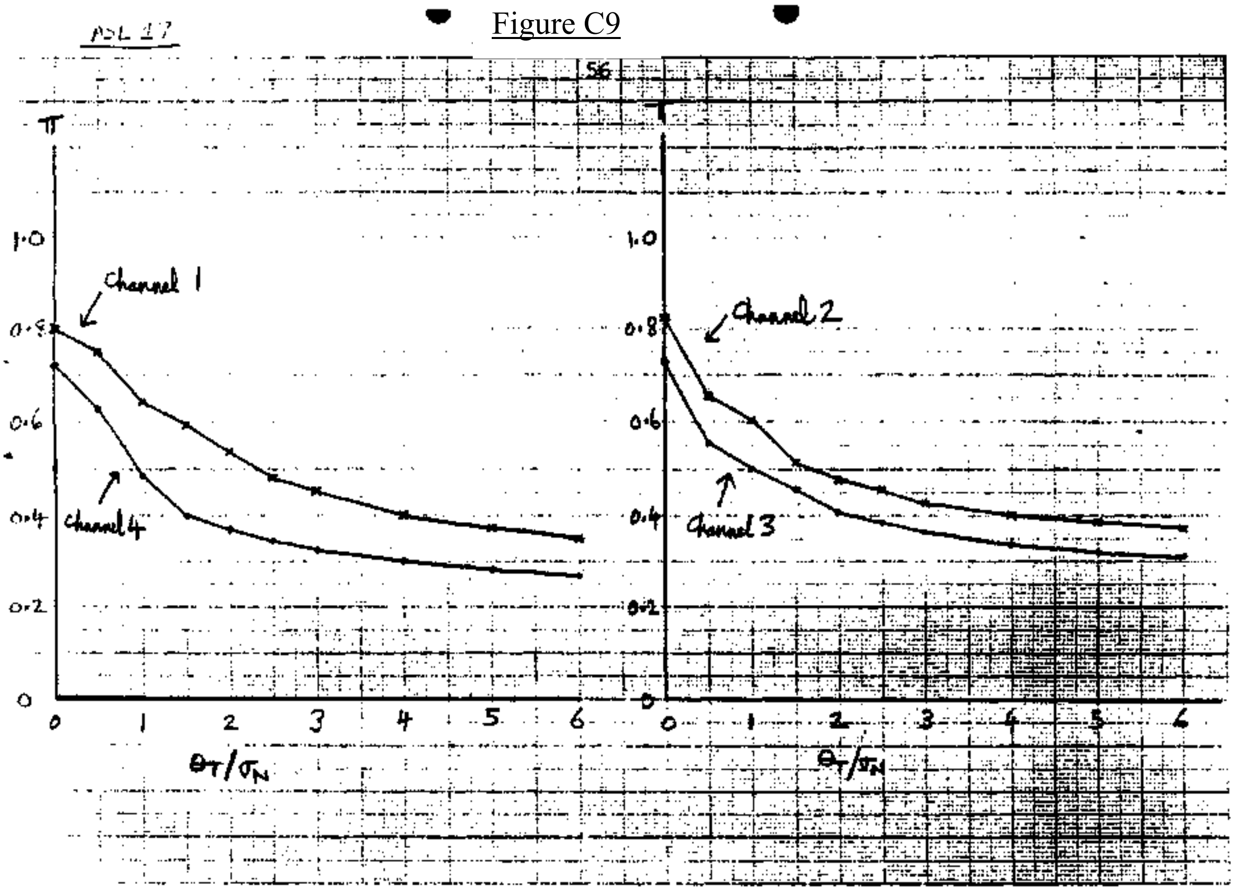
π
1.0
0.8
0.6
0.4
0.2
0
channel 2
Channel 1
channel 4
channel3
0 1 2 3 4 5 6
θ_T/σ_N

ASL 16



π
1.0
0.8
0.6

Figure C9


ASL 17
56
Π
1.0
0.8
0.6
0.4
0.2
0
Channel 1
channel 4
0
1
2
3
4
5
6
θ_T/σ_N
Channel 2
Channel 3
θ_T/σ_N

Moments

Table Cl gives statistics for periods of noise (i.e. when the ion generator was off), and Table C2 gives statistics for the actual experiments. Seconds skipped - no. of seconds from the start of the file that are not used. Seconds analysed - no. of seconds that are analysed to give the statistics. Points excluded - no. of points ignored because they fall below -0.22nCm^{-3}.

Figure C10 plots the downstream variation of these statistics, together with the intermittency for $\theta_T = 2\sigma_N$. In each experiment the channel with the largest value of C has been used. Since there are only four ion collectors in the cross-wind plane only a very crude indication of downstream variation results. Also plotted, for comparison, are results of experiments carried out with the same equipment at Cardington, UK, in stability class ~ C. Details of these experiments can be found in Chatwin and Hajian (1990).

Figure C11 plots the same results, but with the downstream distance d normalised by the mean wind speed U. The figures used for U are only a crude estimate, since the wind data was not available at the time of performing the analysis. The values used were:

Stable experiments	2ms^{-1}
Convective	4.5ms^{-1}
Cardington experiments	3.5ms^{-1}

Although the picture is by no means completely clearcut, the tendency seems to be for increased atmospheric stability to give larger mean and variance (and hence intermittency), and lower skewness, kurtosis and intensity.

Table C1

Experiment	Channel	Seconds skipped	Seconds analysed	C	$(\overline{C}^2)^{1/2}$	Skewness	Kurtosis	Points excluded
8	1	50	170	0.499	0.0215	1.99	17.7	0
	2			0.115	0.0178	1.50	8.98	0
	3			0.0806	0.0194	0.853	10.1	0
	4			0.0623	0.0291	1.26	5.91	0
9	1	5	260	0.475	0.0301	1.43	8.41	0
	2			0.112	0.0280	3.58	36.4	0
	3			0.0252	0.0326	1.05	7.39	0
	4			0.0330	0.0322	1.21	6.65	0
10	1	5	475	0.461	0.0347	2.14	13.9	0
	2			0.105	0.0292	2.39	15.6	
	3			0.0206	0.0344	1.12	7.01	0
	4			0.0288	0.0355	1.50	8.04	0
11	1	5	200	0.452	0.0284	2.21	12.9	0
	2			0.0883	0.0237	1.72	9.28	0
	3			0.0077	0.0306	2.54	21.2	0
	4			0.0204	0.0294	2.15	13.0	0
11	1	3800	180	1.22	0.0679	3.34	22.0	0
	2			0.144	0.0787	3.29	21.2	0
	3			-0.0815	0.0819	3.31	19.9	0
	4			-0.0289	0.0589	2.43	13.7	0
13	1	5	215	1.31	0.0490	-0.337	14.0	0
	2			-0.0548	0.0408	0.858	8.40	4
	3			0.0435	0.0460	0.708	7.80	3
	4			0.191	0.0430	-0.111	11.6	3
16	1	1925	80	1.29	0.0411	0.713	4.53	0
	2			0.154	0.0365	-0.220	3.57	1
	3			-0.0566	0.0359	0.412	4.00	0
	4			-0.0008	0.0451	0.523	3.72	0
17	1	3580	200	1.30	0.0438	0.605	4.28	0
	2			0.163	0.0400	0.631	7.41	0
	3			-0.0534	0.0443	1.82	26.1	0
	4			0.0087	0.0474	0.503	3.69	2

Table C2

Experimnet	Channel	Seconds skipped	Seconds analysed	C	$(\overline{C}^2)^{1/2}$	Skewness	Kurtosis	Intensity	Points Excluded
8	1	220	2500	0.0779	0.376	5.57	37.4	4.83	0
	2			0.0782	0.316	5.17	34.2	4.04	0
	3			0.0600	0.340	4.90	30.2	5.67	0
	4			0.0539	0.27	5.08	32.9	5.01	0
9	1	265	2600	0.0992	0.372	6.19	51.8	3.75	0
	2			0.0400	0.187	7.34	73.1	4.69	0
	3			0.0333	0.166	7.53	74.4	4.99	0
	4			0.0163	0.0677	9.34	168	4.17	0
10	1	480	2600	0.108	0.481	5.94	44.1	4.46	0
	2			0.139	0.544	4.98	30.9	3.91	0
	3			0.177	0.624	4.62	27.1	3.52	0
	4			0.198	0.619	4.30	24.8	3.13	0
11	1	420	3300	0.252	0.906	5.05	32.2	3.59	4
	2			0.294	0.903	3.97	21.3	3.07	15
	3			0.266	0.831	3.94	21.5	3.13	1
	4			0.116	0.458	6.22	59.1	3.96	0
13	1	220	4690	0.256	1.06	5.03	33.7	4.15	5
	2			0.331	1.09	4.26	23.2	3.30	2
	3			0.157	0.669	5.67	45.5	4.25	14
	4			0.289	0.993	4.68	28.0	3.44	3
16	1	5	1800	1.83	3.11	2.05	6.88	1.70	29
	2			2.04	3.12	1.72	5.25	1.53	26
	3			1.78	3.24	2.36	8.99	1.82	4
	4			0.872	1.89	2.88	13.1	2.17	7
17	1	2015	1500	0.733	1.40	2.66	10.9	1.91	1
	2			0.733	1.46	2.60	10.6	1.89	1
	3			0.604	1.27	2.93	13.0	2.11	1
	4			0.468	0.994	2.87	12.4	2.13	0

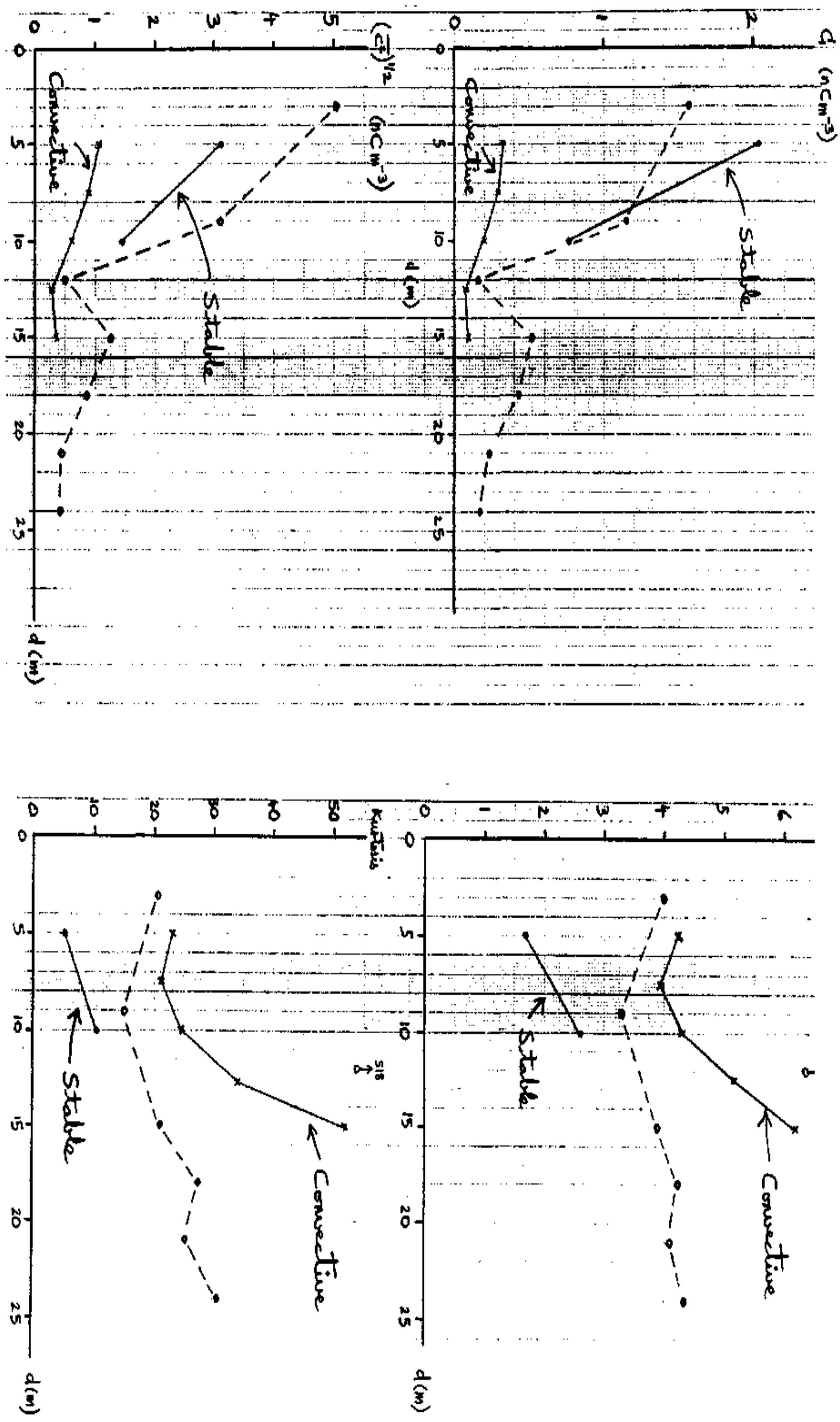
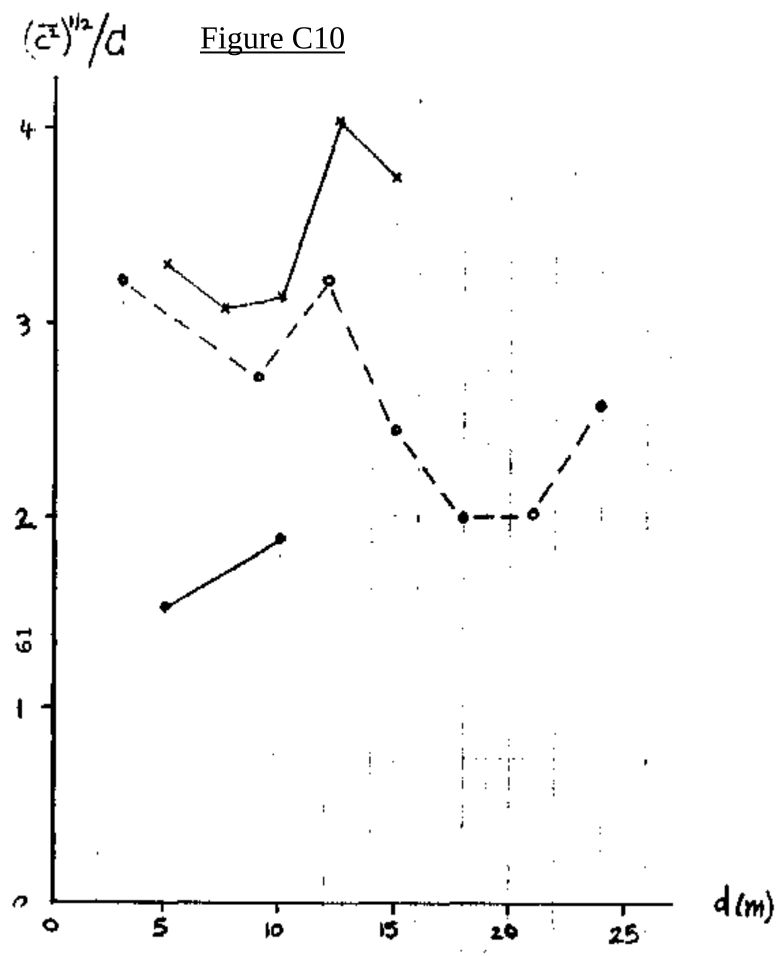
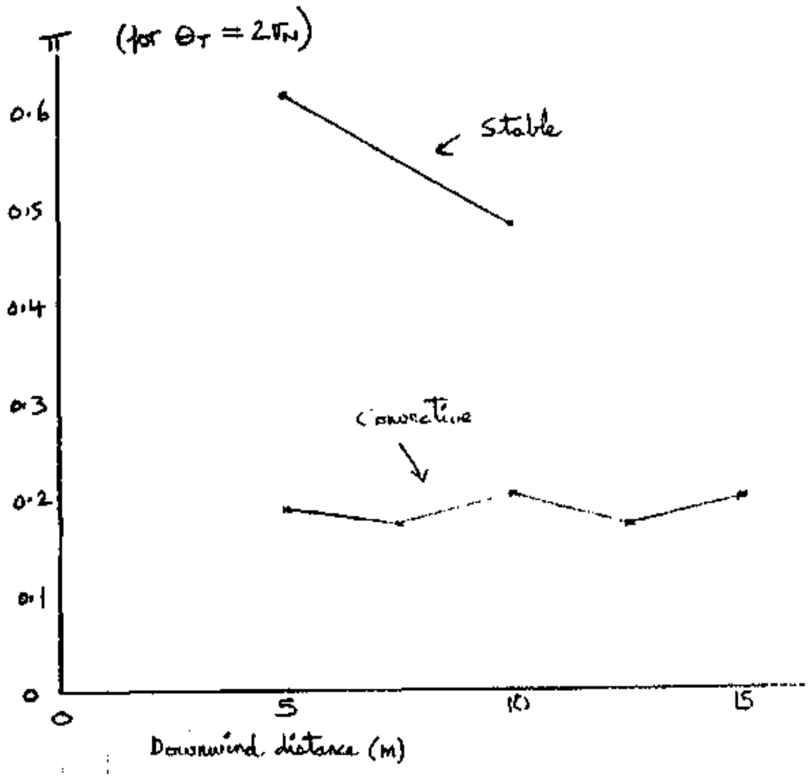


Figure C10

Figure C10



$(\bar{c}^{\prime})^{1/2}/\bar{c}$
d (m)



Π (for $\Theta_T = 2\sigma_N$)
Stable
Convective
Downwind distance (m)

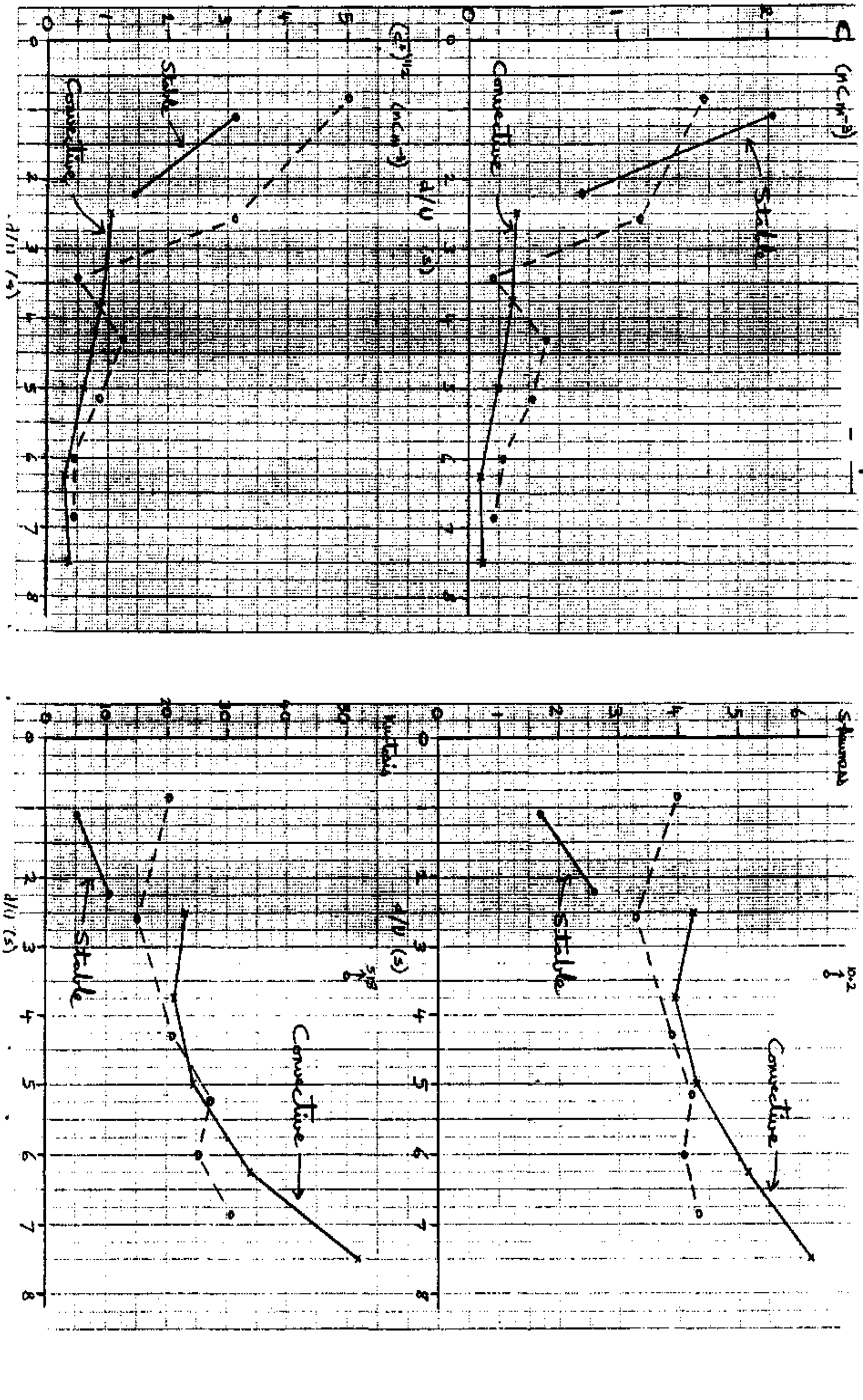
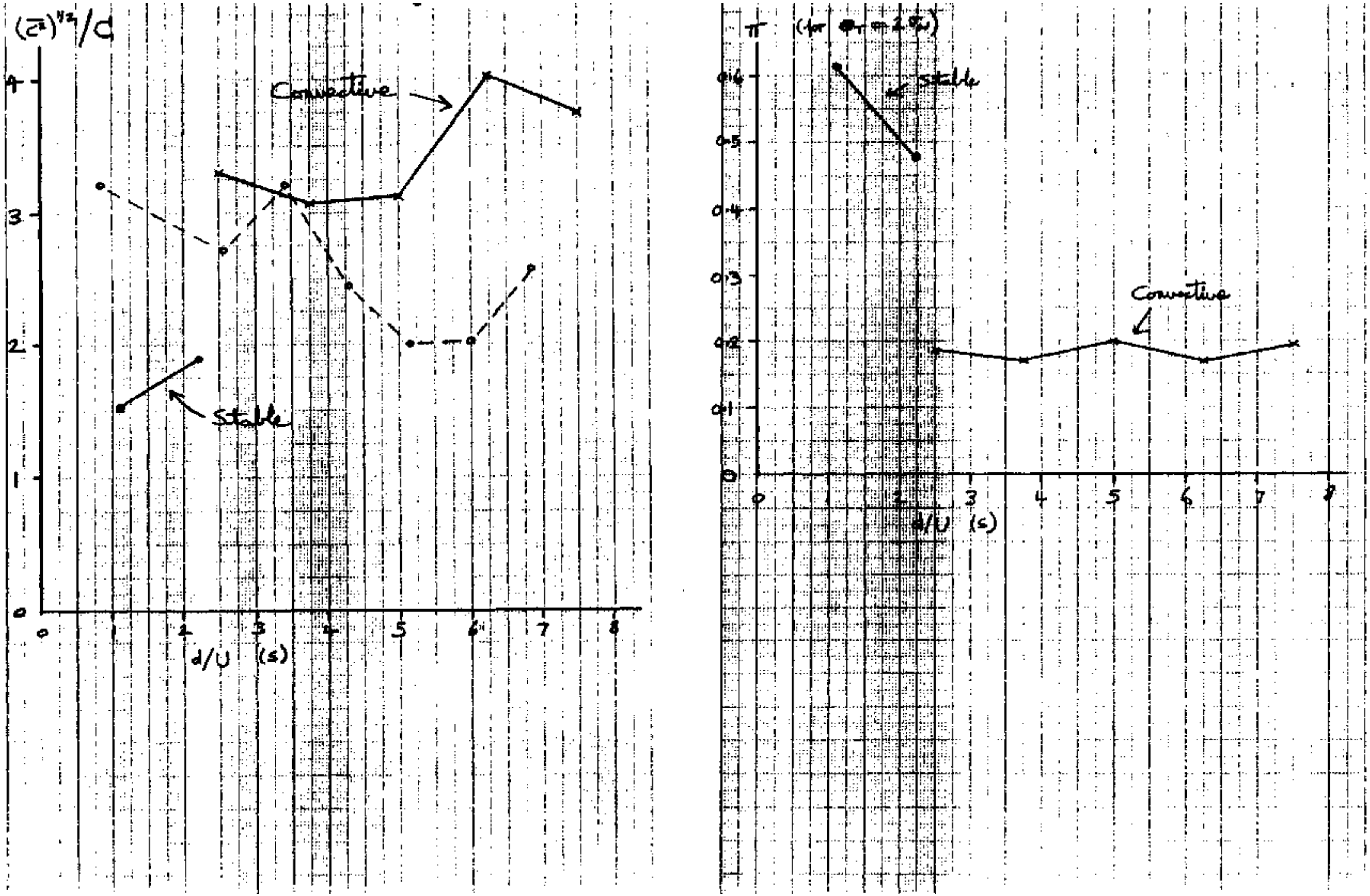


Figure C11

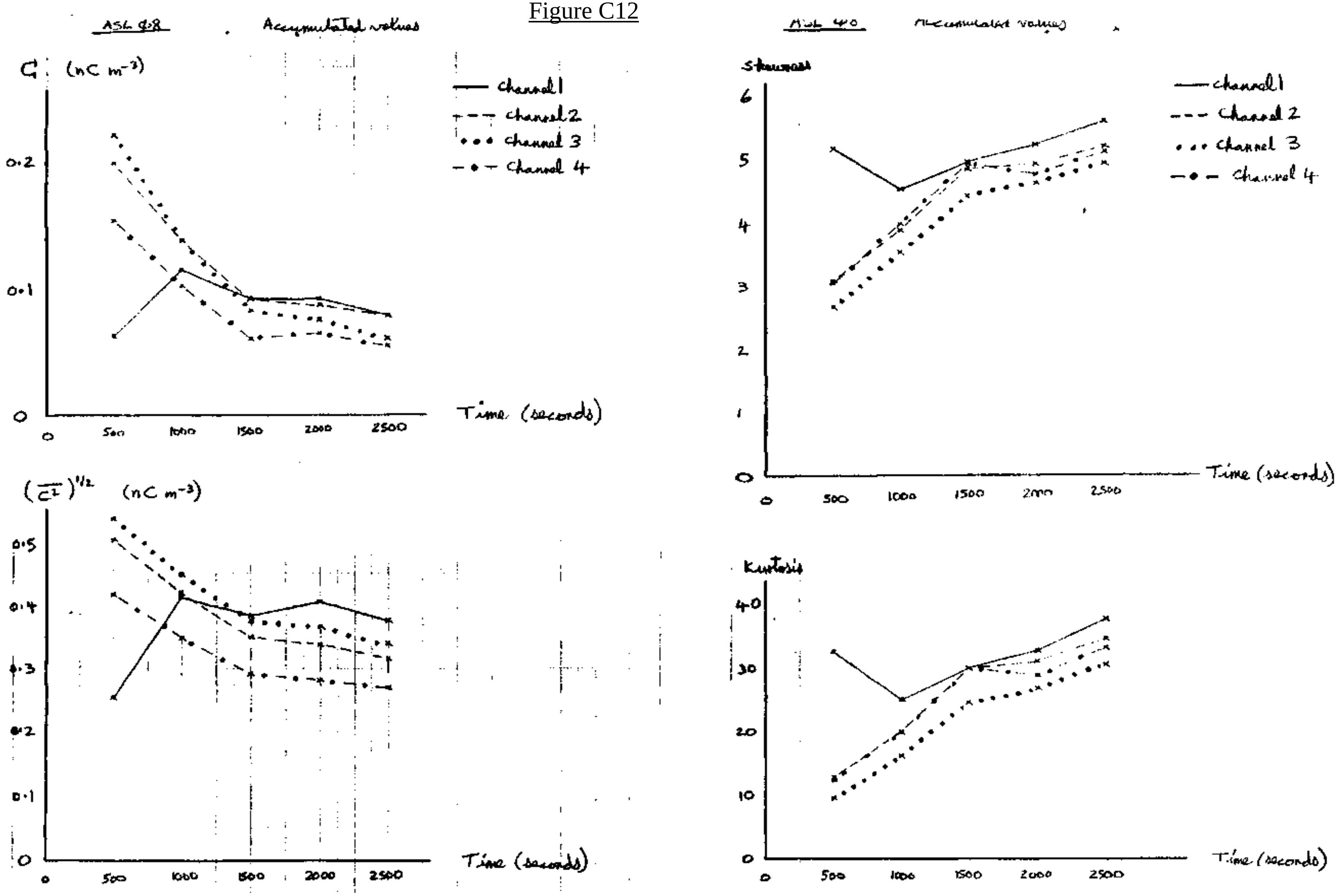
Figure C11



Stationarity and Convergence of Statistics

Some idea of the degree of Stationarity and convergence of statistics is provided by Figures C12-C18. These show the statistics calculated for an increasing (cumulative) length of time and, in the case of experiments 13 and 16, for successive periods of length 500 seconds.

Figure C12


Accumulated values
C (nC m^{-3})
0.2
0.1
0
500
1000
1500
2000
2500
Time (seconds)
channel 1
channel 2
channel 3
channel 4
$(\overline{c^2})^{1/2}$ (nC m^{-3})
0.5
0.4
0.3
0.2
0.1
Skewness
6
5
4
3
2
1
Kurtosis
40
30
20
10

ASL ϕ8 Accumulated values FigureC12

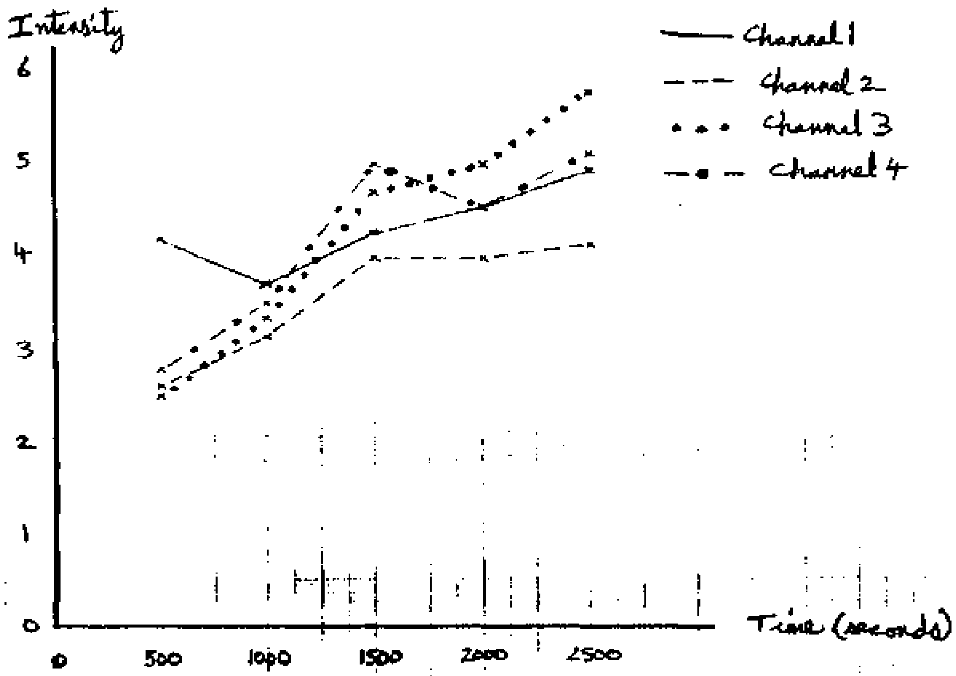
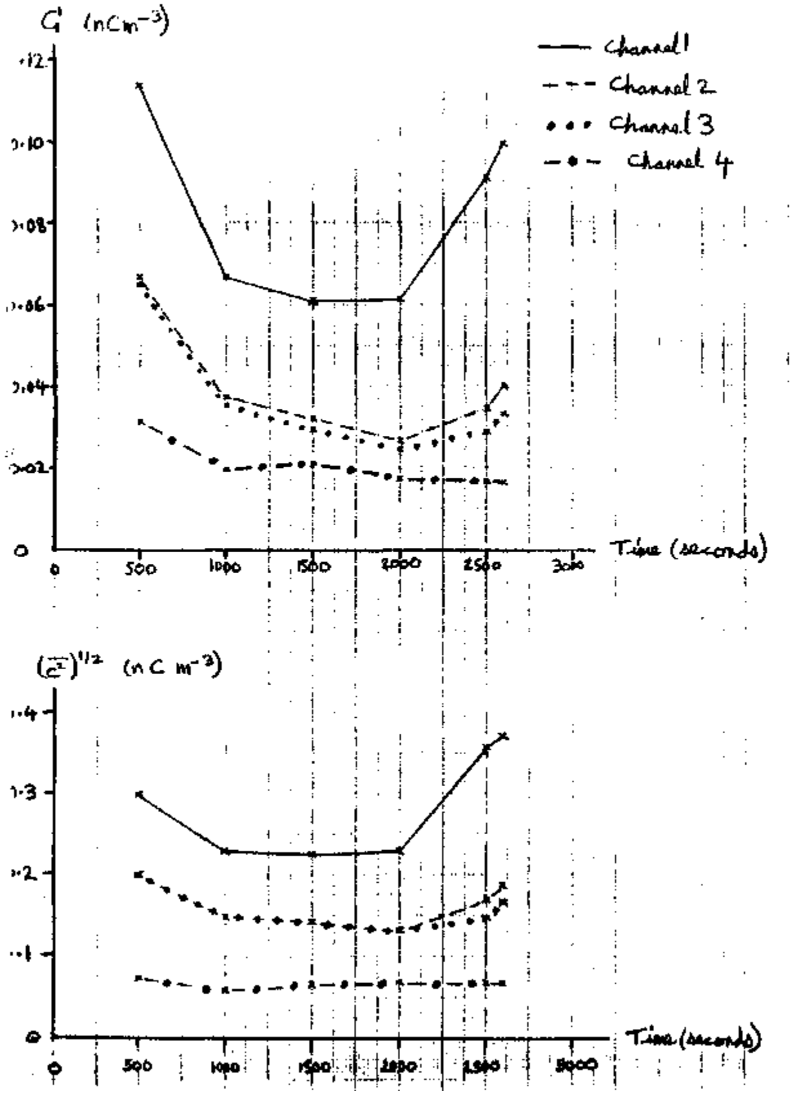
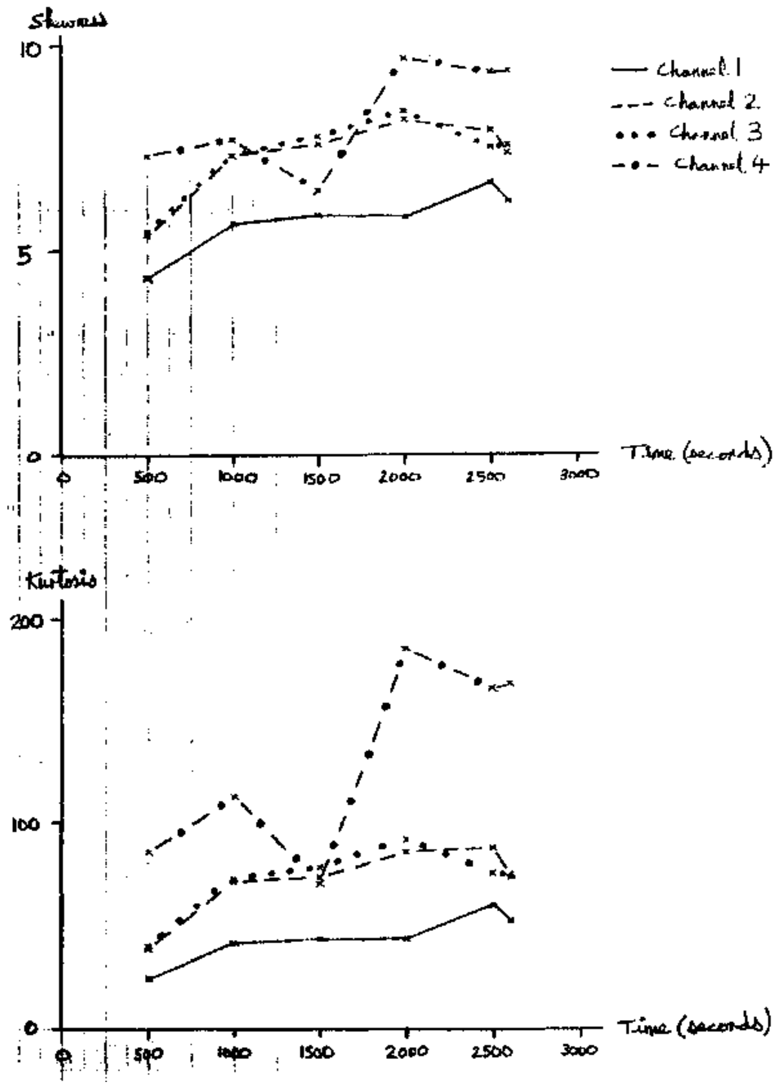


Figure C13

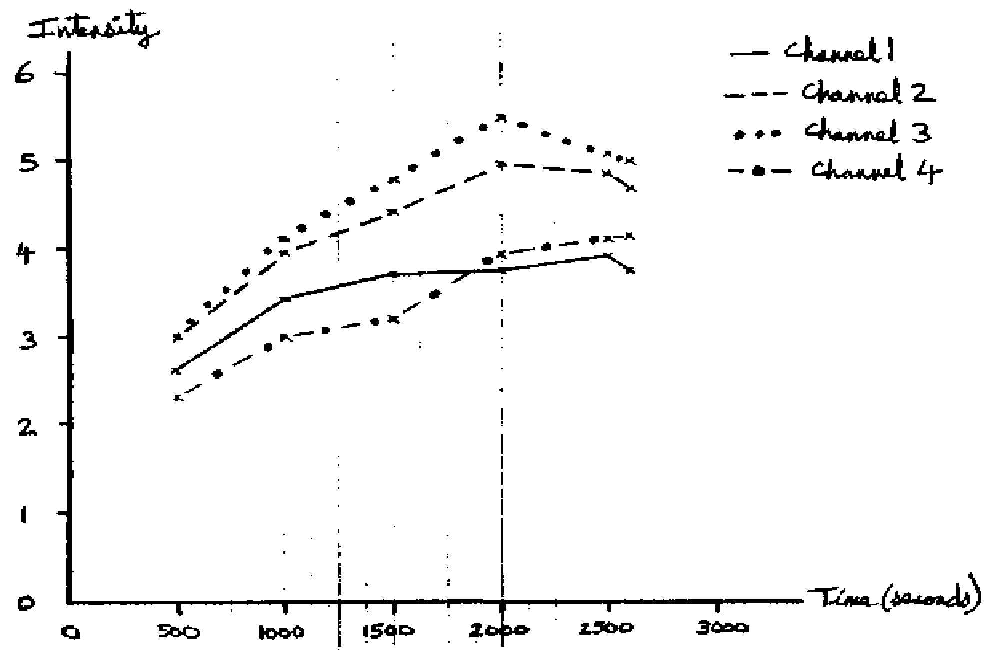
ASL ϕ9 Accumulated values



ASL ϕ'1 Accumulated values

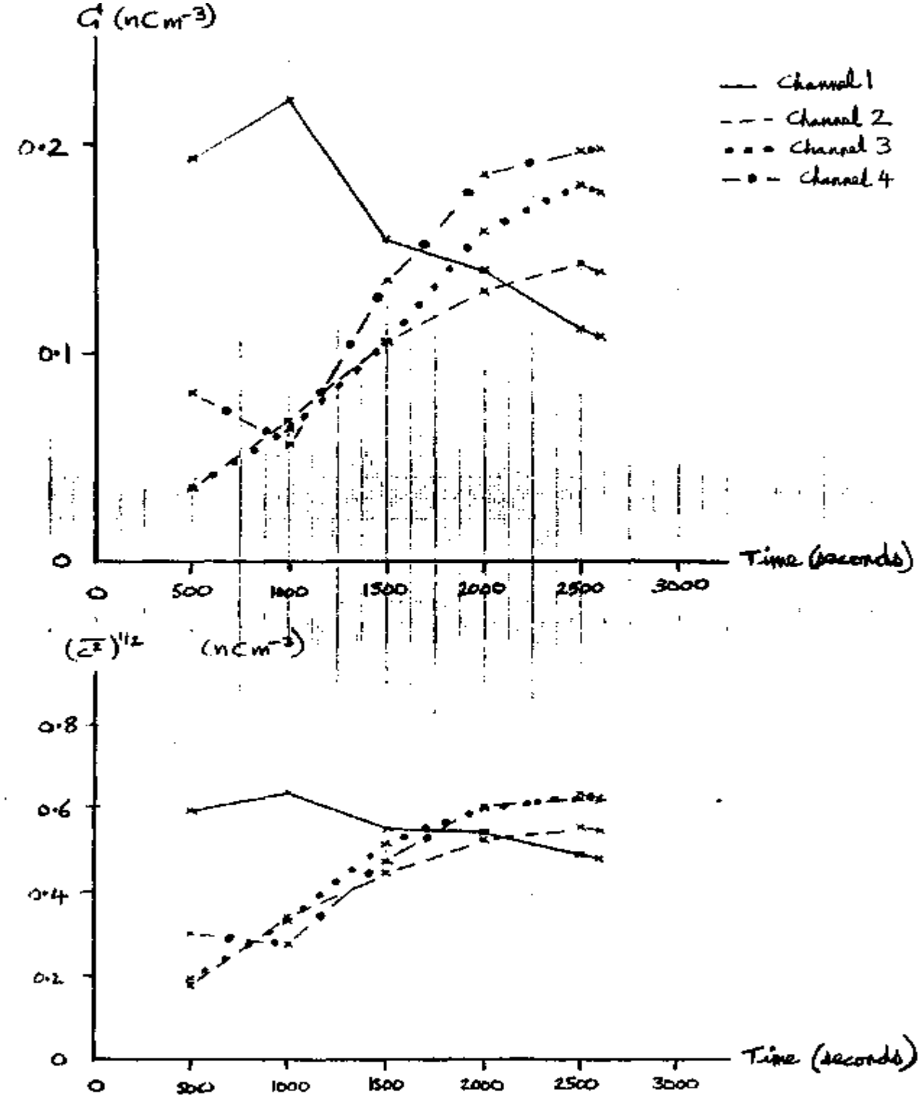


ASL ϕ9 Accumulated values Figure C13



ASL 10 Accumulated values

Figure C14



ASL 10 Accumulated

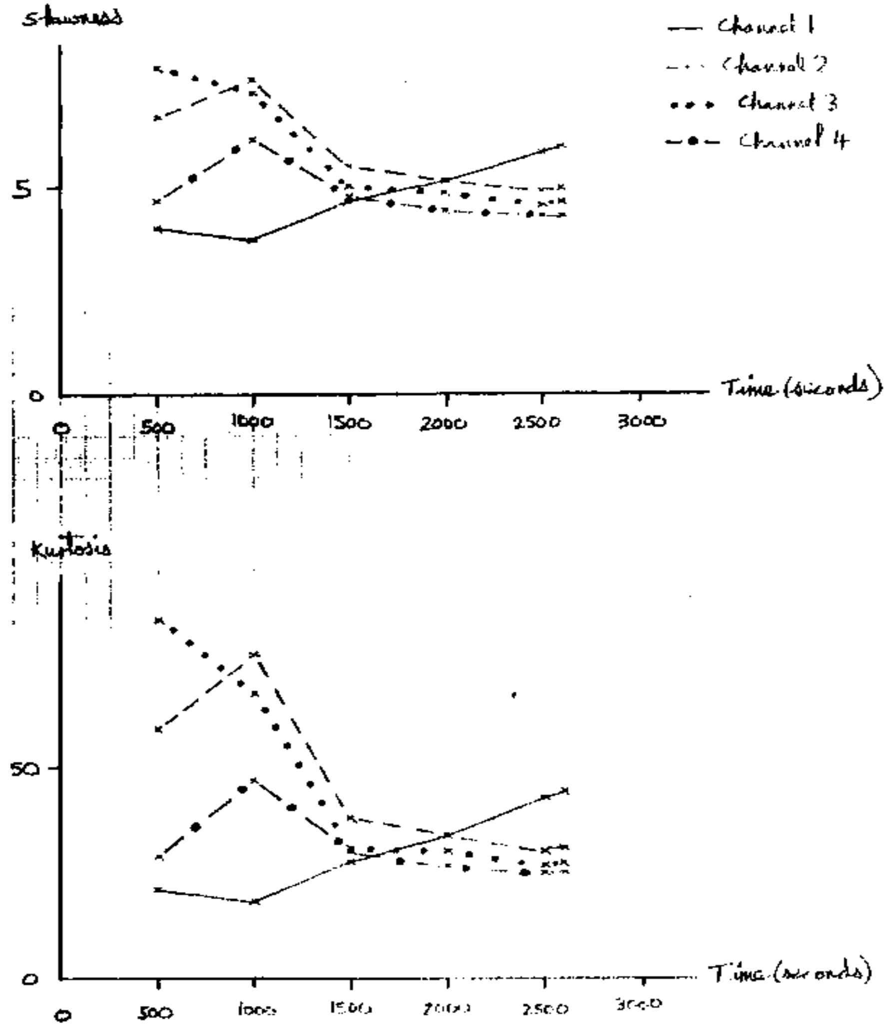
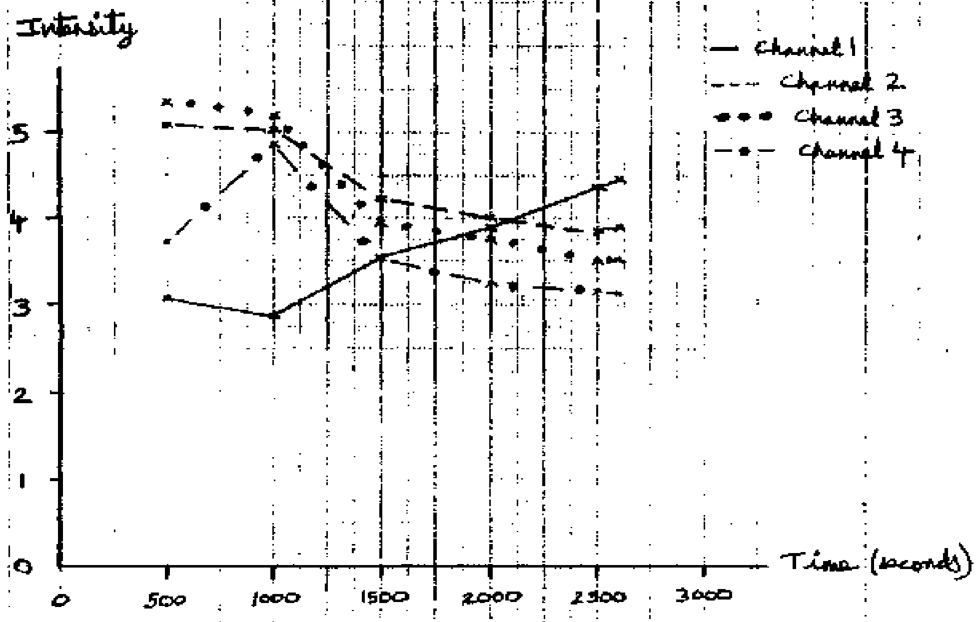
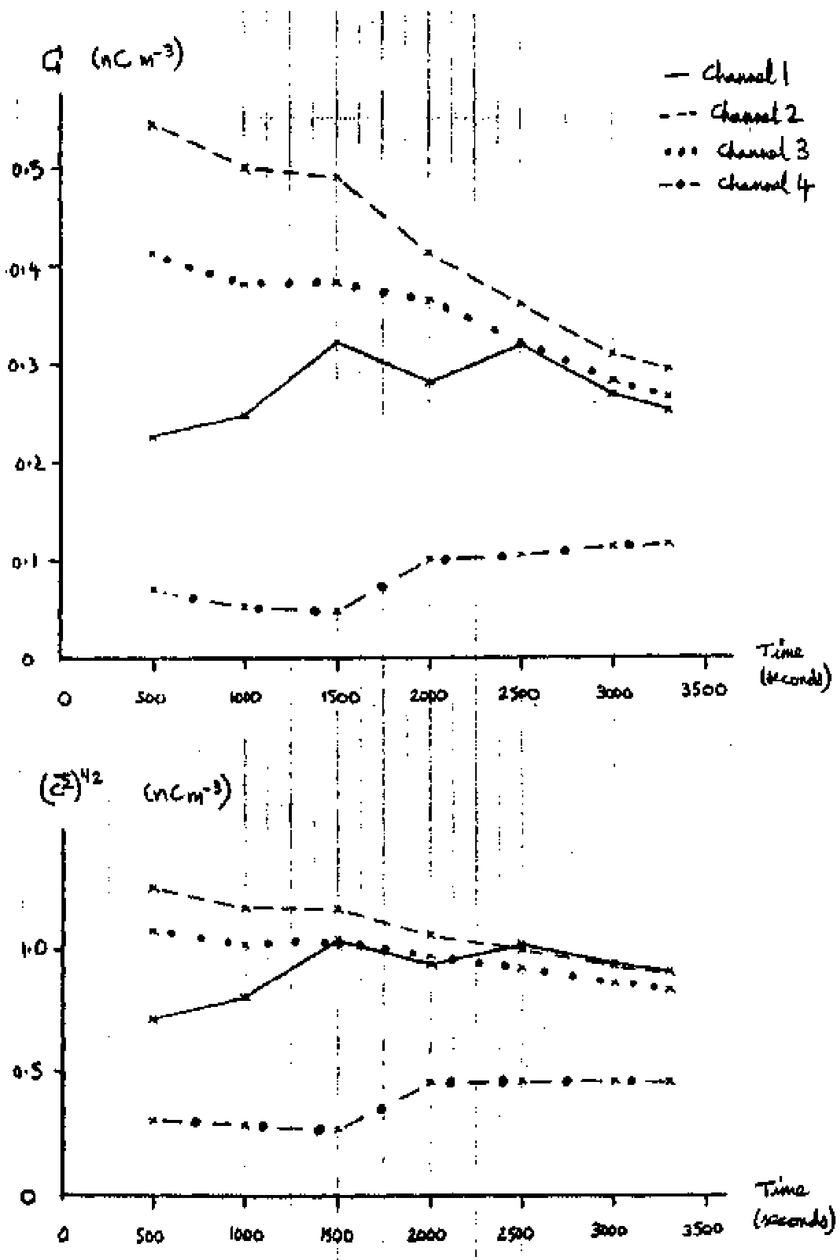


Figure C14

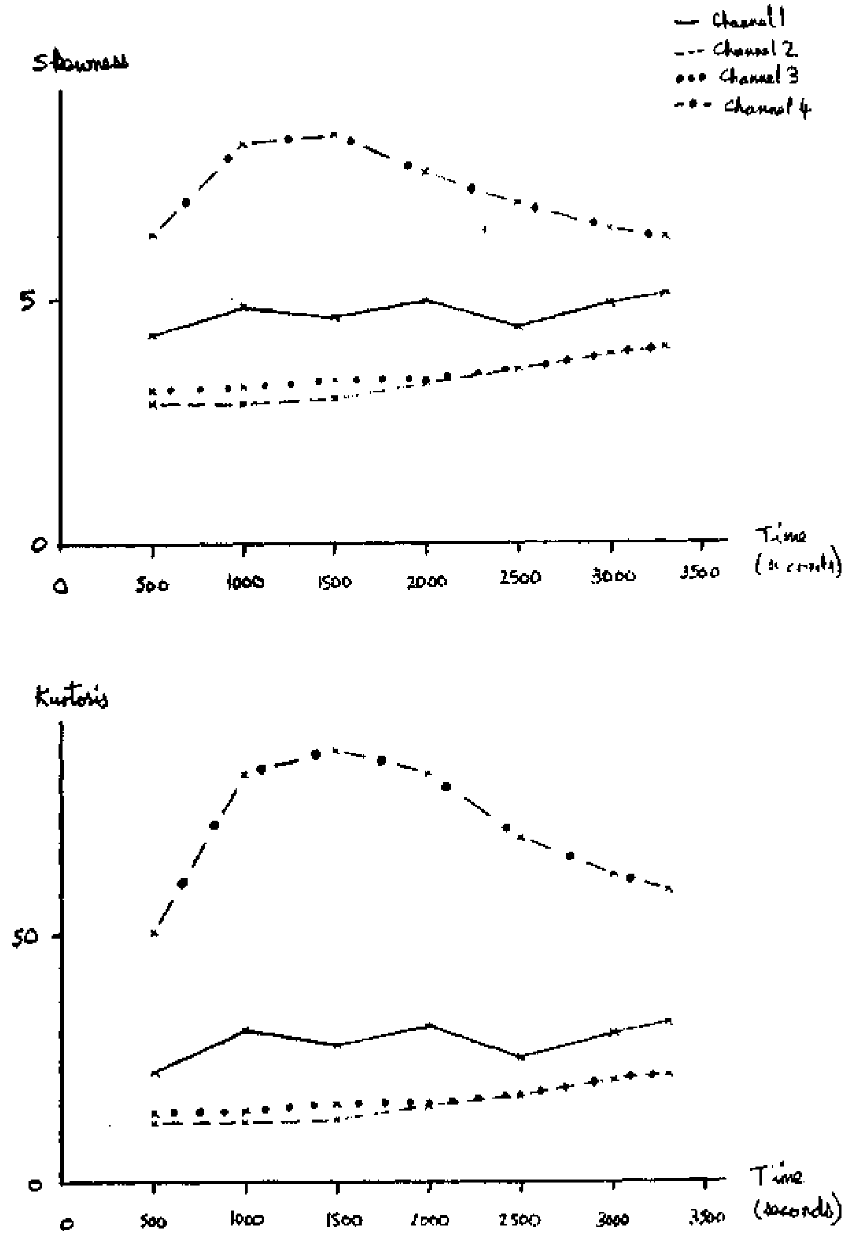
ASL 10 Accumulated values



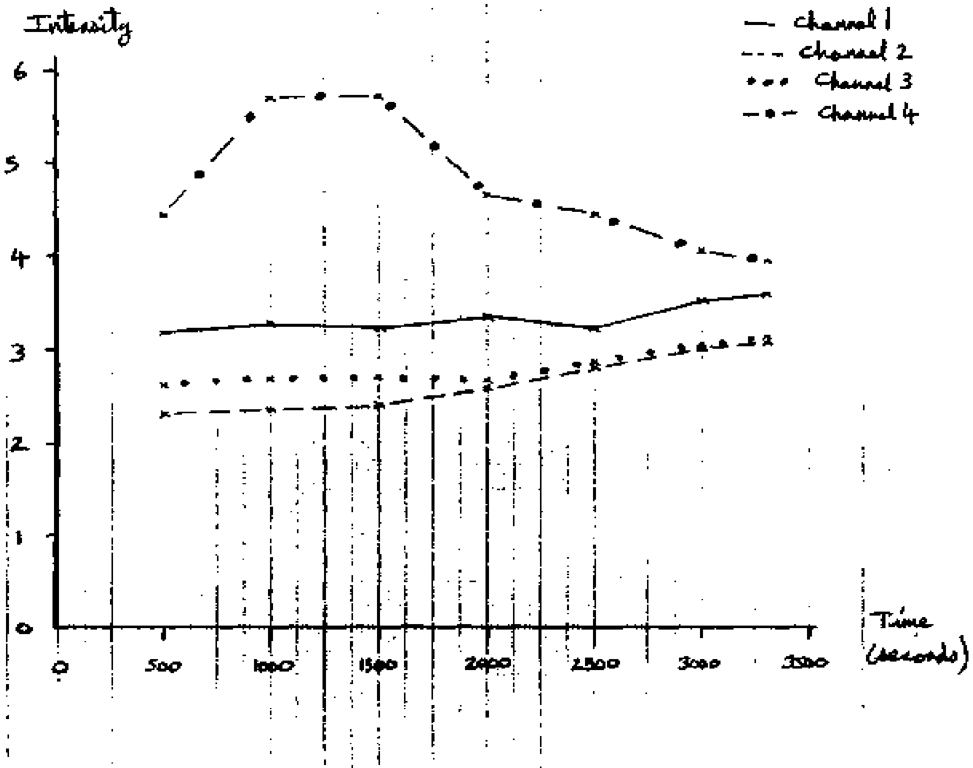
ASL 11 Accumulated values Figure C15



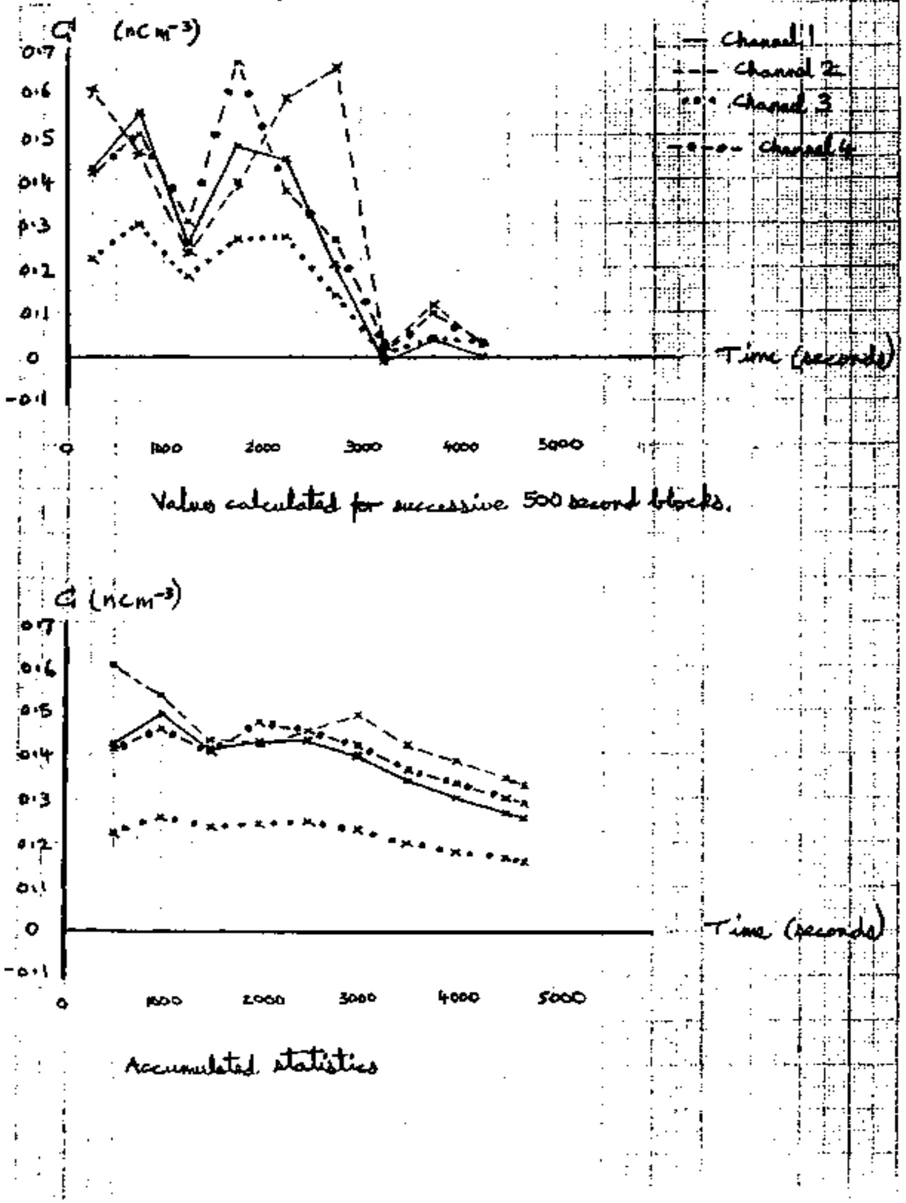
ASL 11 Accumulated values



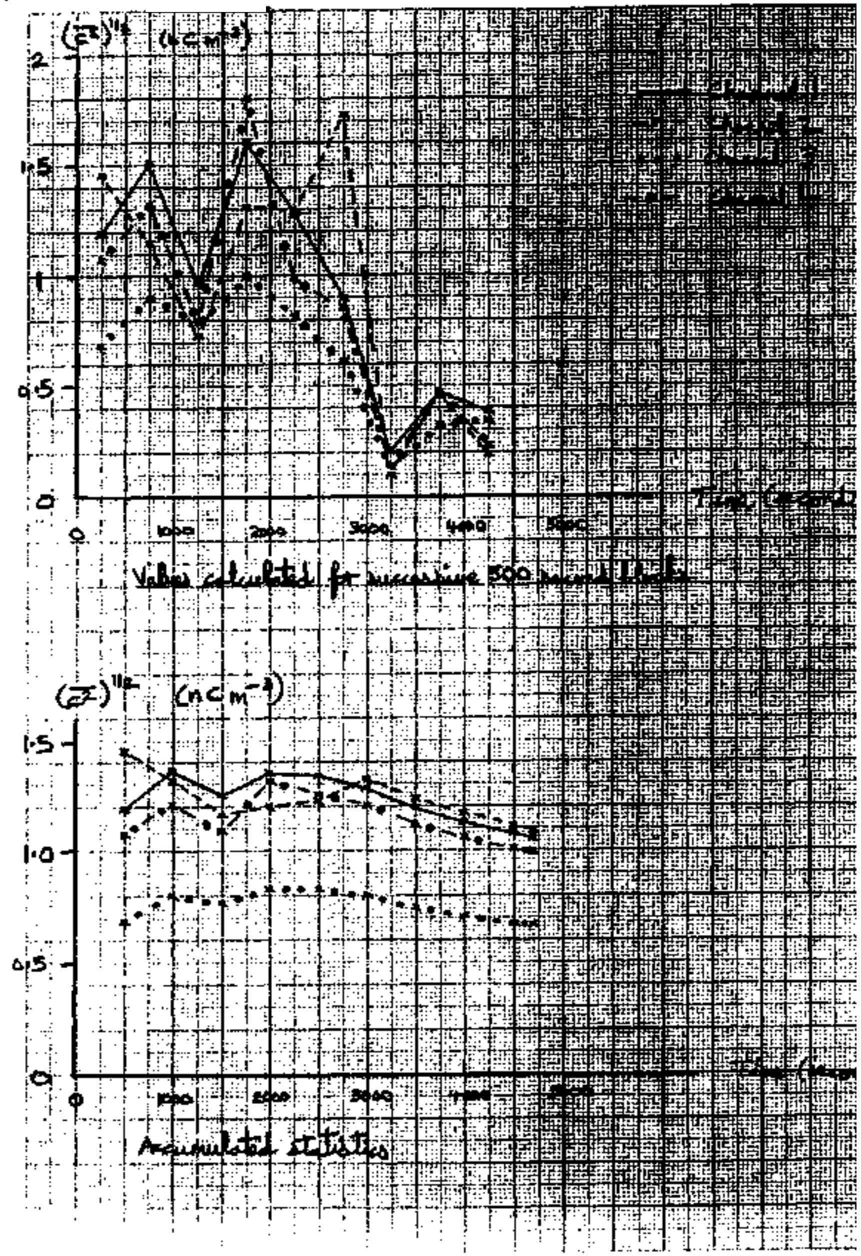
ASL 11 Accumulated values Figure C15



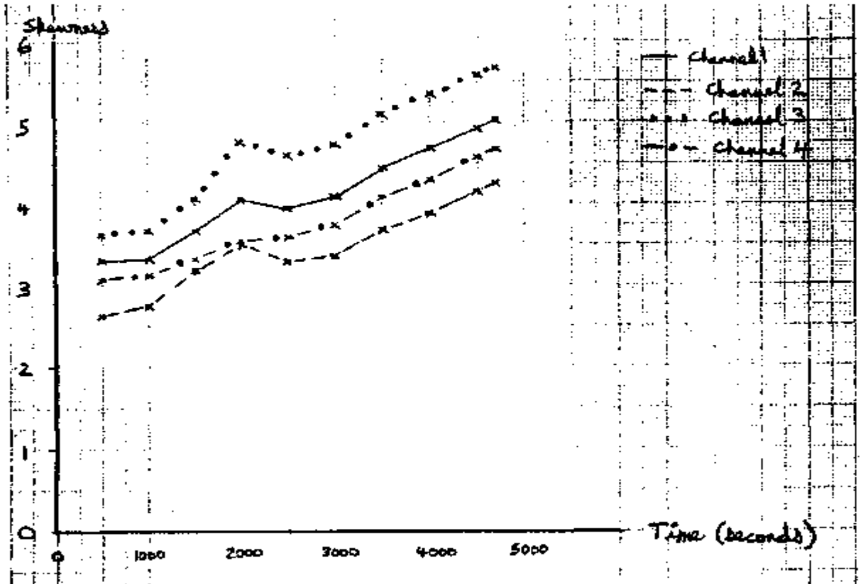
ASL 13 Figure C16

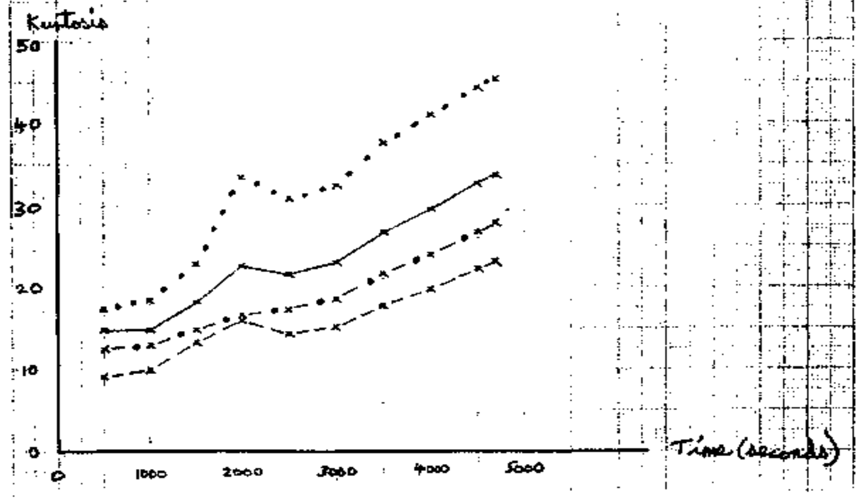


ALS 13



ALS 13 Accumulated values **Figure C16** ALS 13





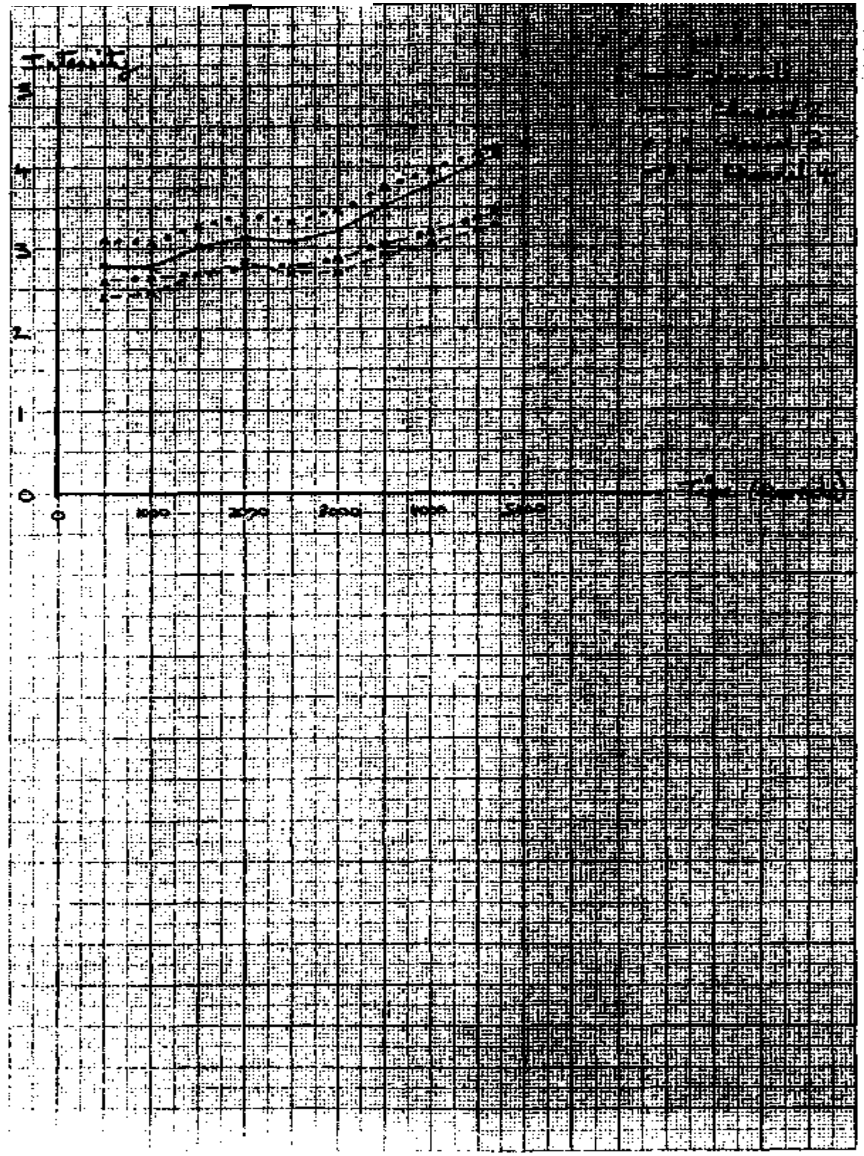
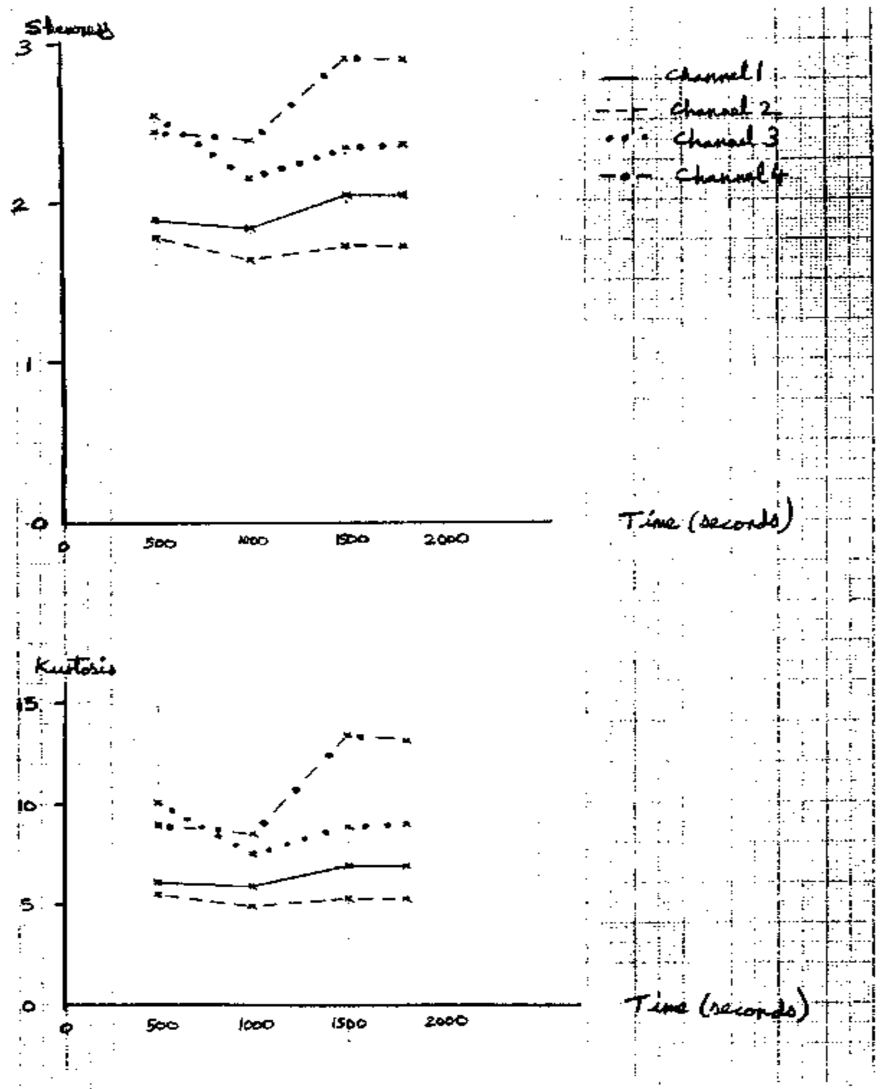


Figure C17

ALS 16



ALS 16

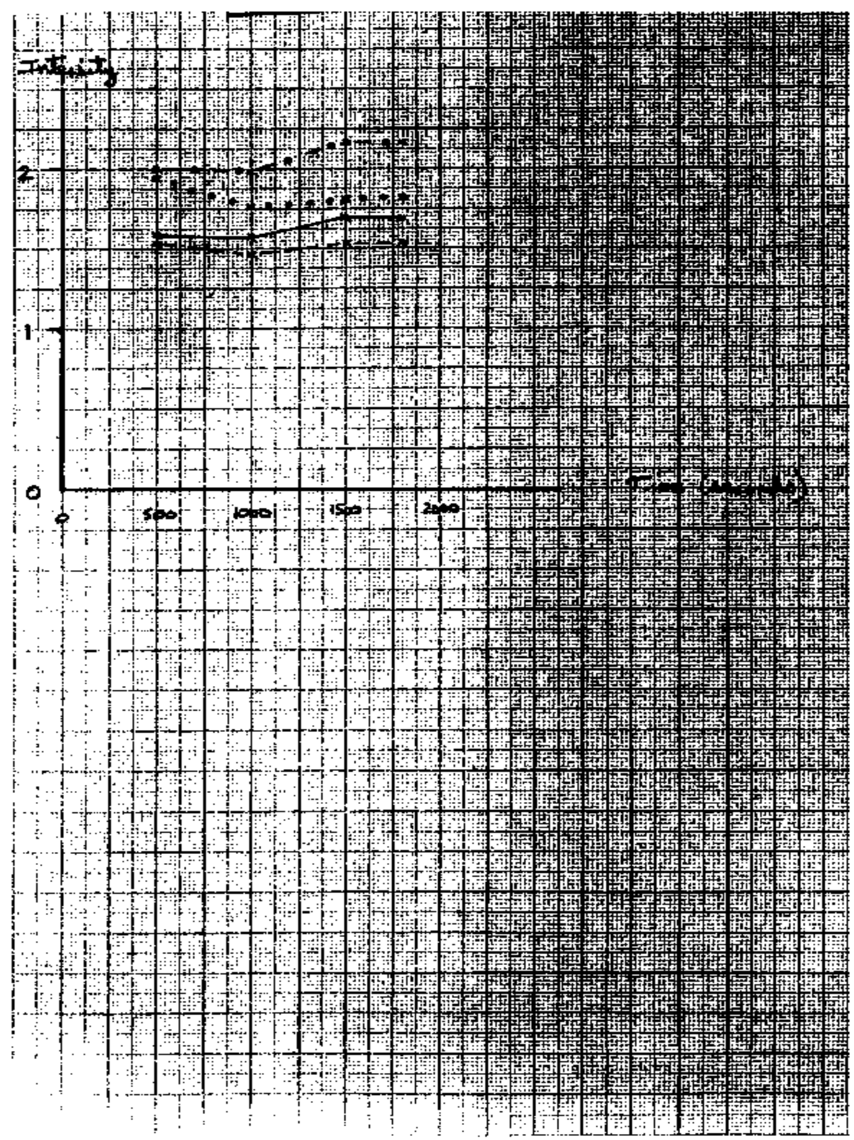
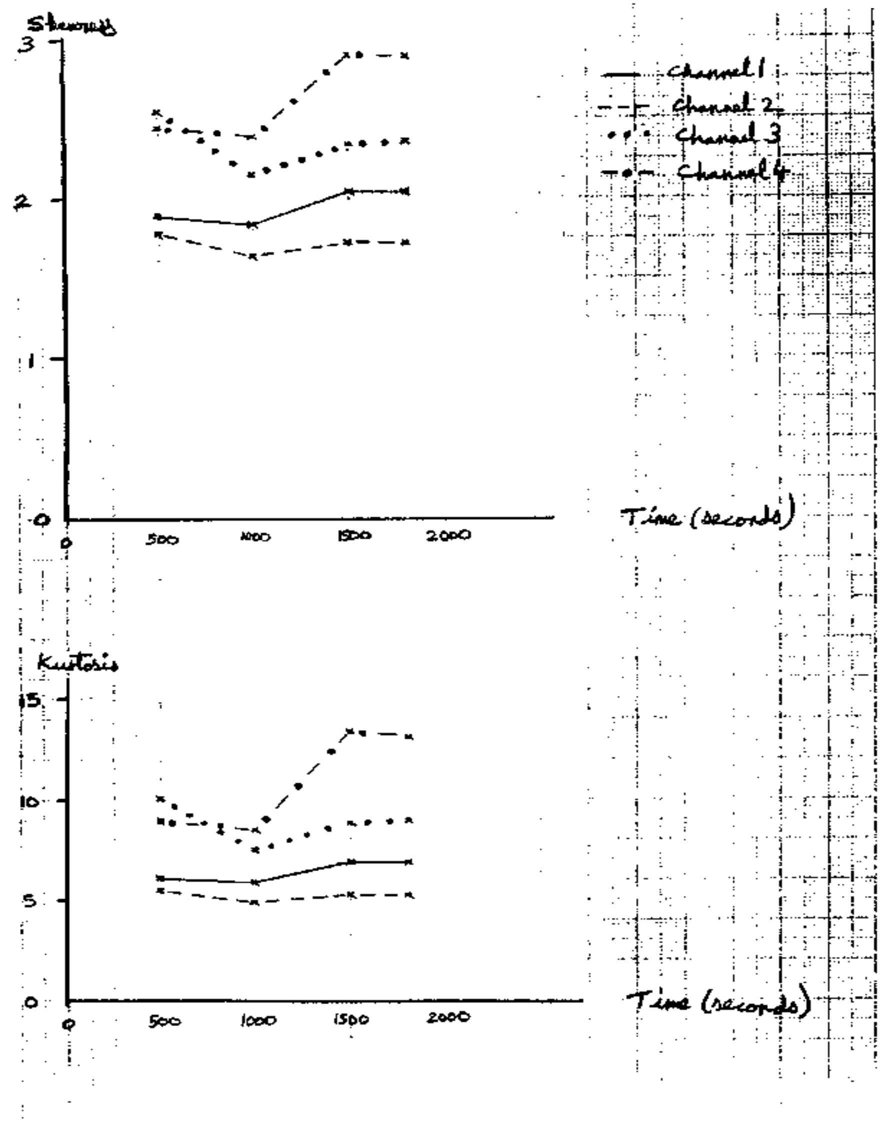


Figure C17

ASL 16 Accumulated values



ASL 16 Accumulated values

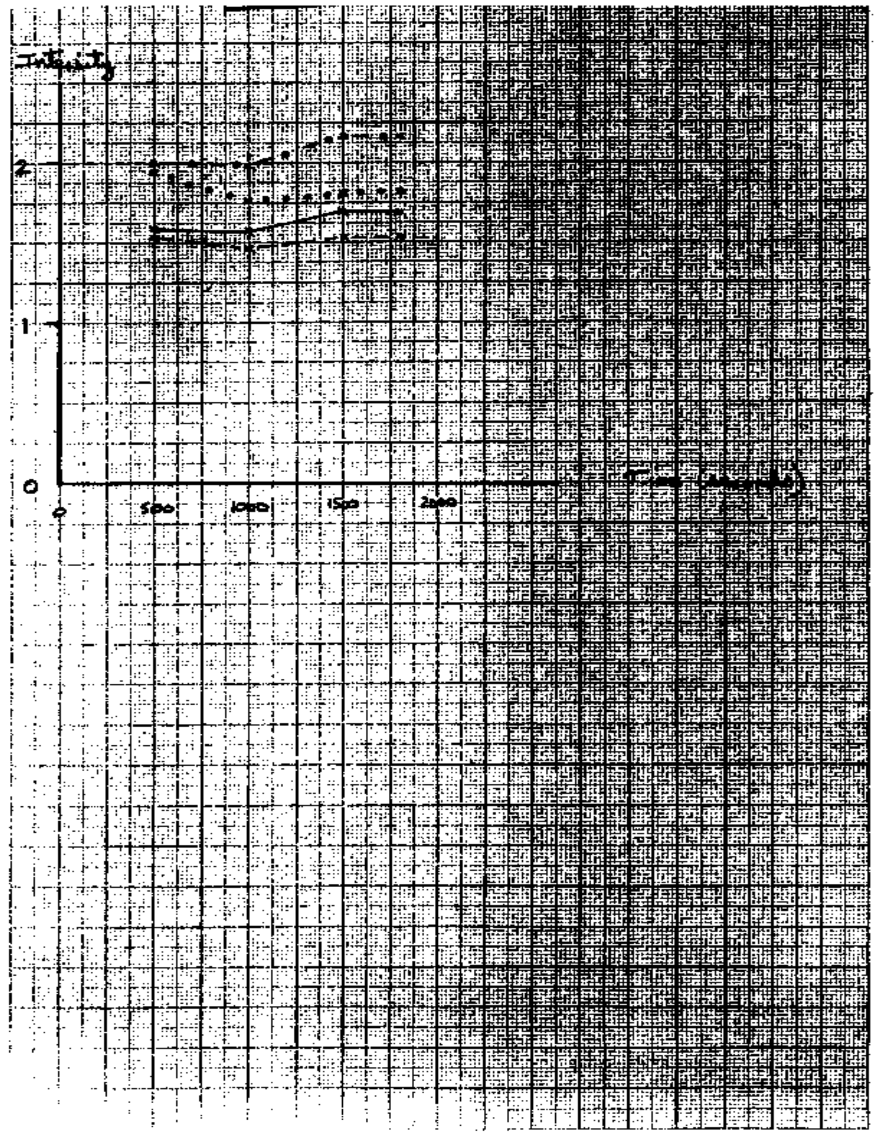
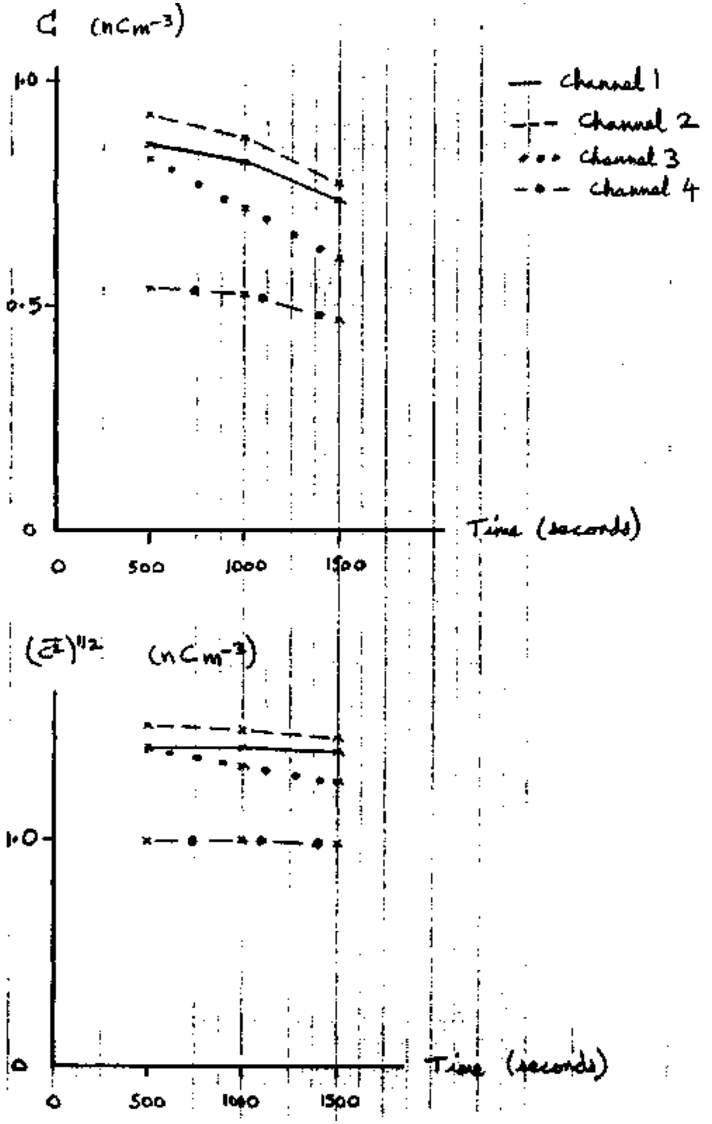
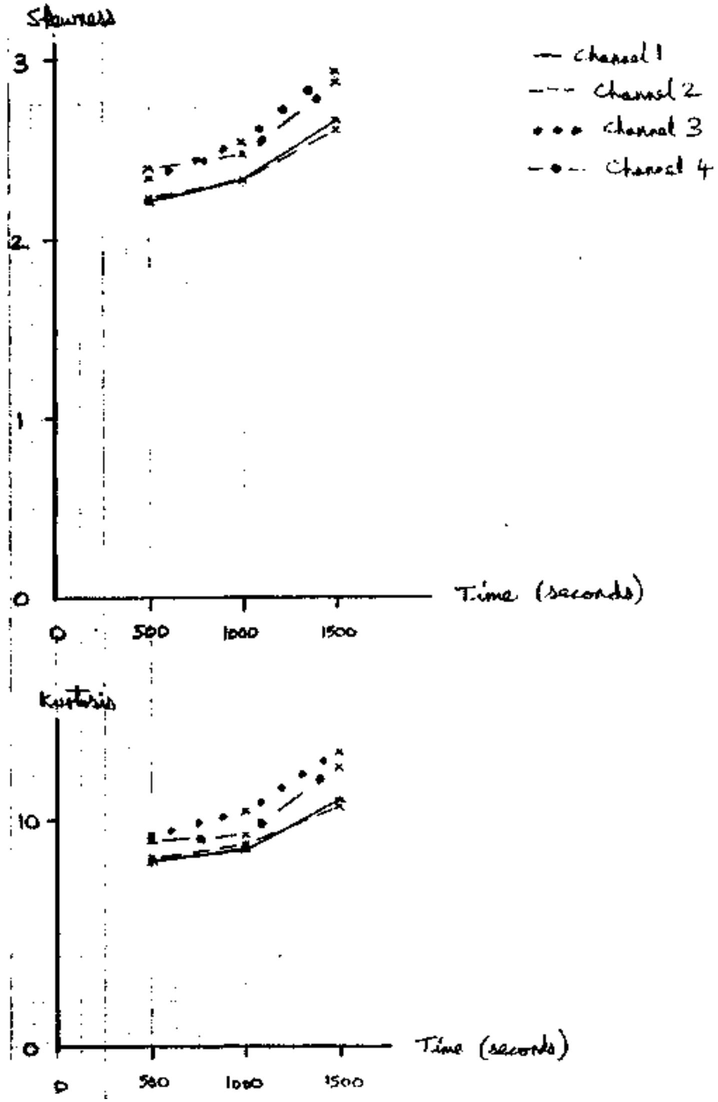


Figure C18

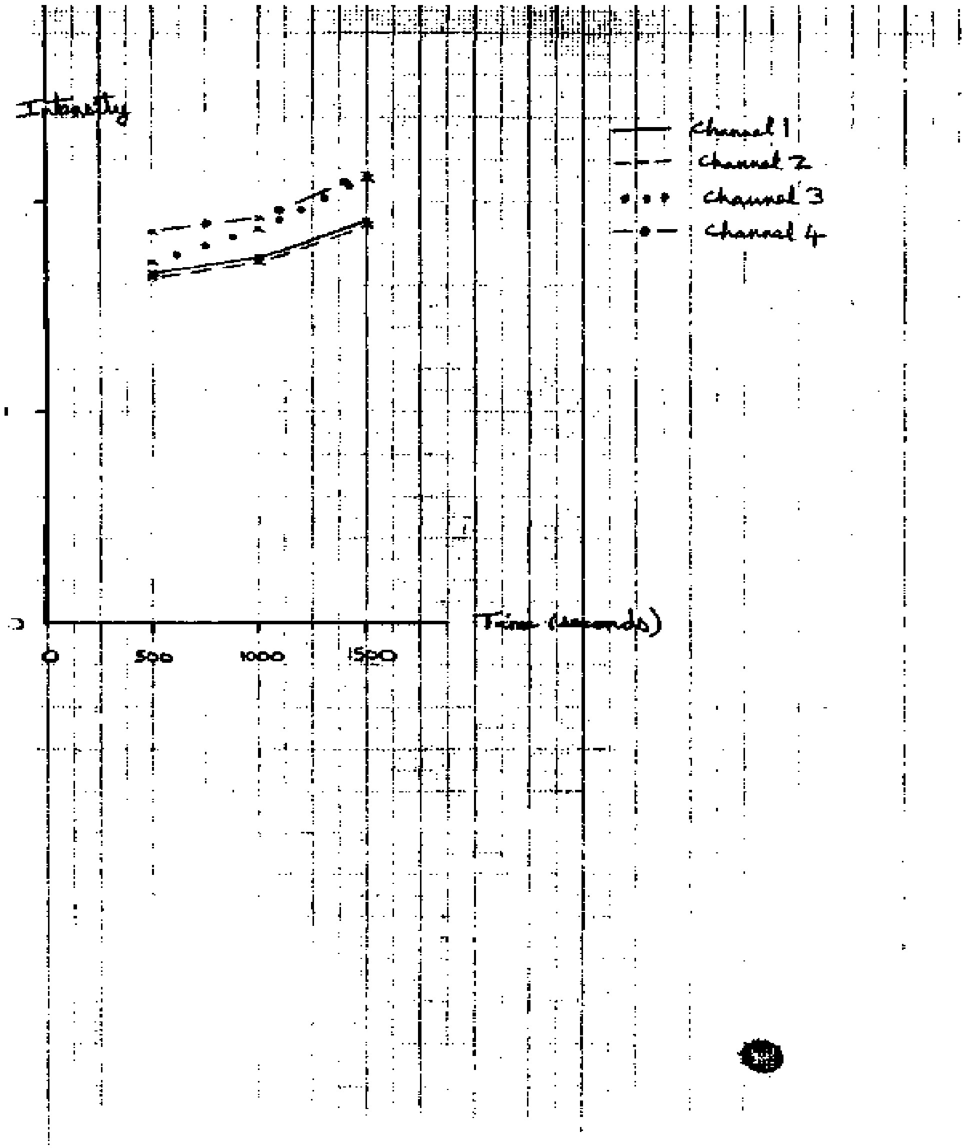
ASL 17 Accumulated values



ASL 17 Accumulated values



ASL 17 Accumulated values <u>Figure C18</u>



Intermlttency and conditional p.d.f. based on marked fluid

Chatwin and Sullivan (1989) define an intermittency π_0 by

$$\pi_0 = \frac{C}{\theta_0}$$

where θ_0 is the source concentration. The p.d.f. of concentration is

$$p(\theta) = (1-\pi_0)g(\theta) + \theta_0 f(\theta),$$

where f(θ) is the p.d.f. conditional on being in marked (source) fluid, and g(θ) is the p.d.f. conditional on being in unmarked fluid. In the present experiments it is always true that $\pi_0 \ll 1$, so that p(θ) ≈ g(θ) for all θ, except possibly at some large values of θ where f(θ) ≫ g(θ). In other words, even at these short downstream distances, marked fluid occupies a very small proportion of time. A question of interest is whether the marked fluid can still have a significant effect on concentration statistics. If not, then the value of π_0 and f as practical, rather than conceptual, tools would be rather limited.

The first two moments are given by:

$$\begin{cases} C = (1-\pi_0)C_g + \pi_0 C_f \\ \sigma^2 = (1-\pi_0)\sigma_g^2 + \pi_0(1-\pi_0)(C_f - C_g)^2. \end{cases}$$

The contribution to the mean from f(θ) is negligible if

$$\pi_0 C_f \ll C \qquad \text{i.e. if } \frac{C_f}{\theta_0} \ll 1$$

Also when $\pi_0 \ll 1$

$$\sigma_2\sigma_g^2 + \pi_0\sigma_f^2 + \pi_0(C_f - C_g)^2 = \sigma_g^2 + \pi_0\,\sigma_f^2 + \pi_0 C_f^2(1 - C_g/C_f)^2$$

It is safe to assume that $C_g/C_f \le 1$. A sufficient and necessary condition to ensure that $\sigma^2 \approx \sigma_g^2$ is then

$$\sigma^2 \gg \pi_0\,\sigma_f^2 \text{ and } \sigma^2 \gg \pi_0\,C_f^2.$$

So when $\pi_0 \ll 1$ there is a negligible contribution from f(θ) to the mean and variance iff the following three conditions are satisfied:

1) $C_f/\theta_0 \ll 1$

2) $\pi_0 \sigma_f^2 \ll \sigma^2$

3) $\pi_0 C_f^2 \ll \sigma^2$

In practice one could estimate C_f, σ_f^2 etc. by assuming that the top π_0 of concentration values represent $f(\theta)$. This would give (assuming sampling errors are negligible) an upper bound for C_f and a lower bound for σ_f^2 Thus, if 1) and 3) are satisfied by the estimated C_f then they are also satisfied by the true value. The satisfaction of 2) cannot be determined from these estimated values however.

If one assumes that $p(\theta)$ is either exponential or beta (of the form $s(1-\theta)^{s-1}$ for $0 \le \theta \le 1$ with $s \ge 2$), and that $f(\theta)$ is given by the top proportion π_0 of concentration values, then $\sigma_f = \sigma$ So it seems plausible that 2) might be satisfied. The effect of an overlap between f and g would be to increase both σ_f and σ_g, while decreasing $(C_f - C_g)^2$, with σ fixed. Because the number of points involved in the overlap is very small compared with the number of points contributing to g, σ_g would not be expected to change significantly. It also seems reasonable that σ_f might not be increased enough to invalidate 2).

Turning to 1) and 3), if they are satisfied by using the maximum occurring concentration value, θ_{max}, to estimate C_f then they will be satisfied by the true value. Results based on θ_{max} are presented below for the experiments discussed above. All units are nCm^{-3}.

Experiment	Channel	C	π_0	θ_{max}	$\frac{\theta_{max}}{\theta_0}$	σ^2	$\frac{C\theta_{max}^2}{\theta_0\sigma^2}$
16	2	2.04	0.0036	15.5	0.027	9.77	0.088
16	4	0.872	0.0015	24.6	0.043	3.59	0.258
17	2	0.773	0.0014	11.6	0.020	2.13	0.186
8	1	0.0779	0.00014	4.07	0.0071	0.141	0.016
10	4	0.198	0.00035	7.68	0.013	0.383	0.053
11	2	0.294	0.00052	9.70	0.017	0.815	0.059
13	2	0.331	0.00058	11.1	0.019	1.19	0.060

Thus, subject to the reservations about the satisfaction of condition 2), it seems that the first two moments are dominated by unmarked fluid statistics, except possibly for channel 4 of experiment 16.

N.B. For these data, estimating $f(\theta)$ from the upper fraction π_0 of concentration values would mean using roughly 2-60 datapoints, depending on the experiment and channel.

Appendix D Guide to Datasets and Software

Datasets

Three main datasets have been utilised on this project. These are the ones referred to as 1), 3) and 6) in the 6 monthly and annual reports. Briefly, they are:

1) Experiments conducted under Agreement 2044/0129 between CDE and UMIST. These took place at Cardington in July 1986 using the ion generator technique of Dr. CD. Jones. The data suffer from the defect of having been digitised with too coarse a resolution (only 8 bits, i.e. 256 bins).

The experiments are listed by Sanders (1987), and data tapes are held at CDE.

3) Wind tunnel experiments of Dr. J.E. Fackrell and Dr. A.G. Robins. These were experiments with limited duration releases from a point source at ground level, with 200-300 replications for each. The data tapes are held by Dr. N. Mole.

6) Experiments carried out by Dr. C.D. Jones in November 1989 in New Mexico. These are the experiments discussed in Appendix C. The original analog tapes are held at CDE, and the l0Hz digitised data are held at CDE and by Dr. N. Mole. Wind data are held on tape at CDE, but there have been problems with reading it. These experiments have now been supplemented by further experiments carried out in October 1990, at the same site. The analog tapes are held at CDE, but have not yet been digitised. These 2 sets of experiments are the only ones described here in which periods of noise were deliberately recorded.

Other available datasets which may prove useful were also described in the earlier reports. These are:

2) Cardington experiments of Dr. C.D. Jones from May 1988, described by Chatwin and Hajian (1990). The digitised data are held in short (of the order of 5 minutes) sequences on floppy disc. Thus analysis of temporal structure is less easily accomplished from this digitised data. The digitised and analog data are both held at CDE.

4) Wind tunnel experiments of Dr. DJ. Hall. These were instantaneous releases of gas contained in a cylinder. Most of these releases were of heavy gas, but some were of neutrally buoyant gas. There were 50-100 replications. These experiments were carried out under a CEC research programme, under which other experiments, both in the wind tunnel and in the field, were carried out. Enquiries about data availability should be made to Professor P.C. Chatwin.

5) Some of the continuous releases in the field described by Mylne and Mason (1990) are also available.

Software

Software has been produced for the analysis of datasets such as these, and will be described briefly below. It consists of programs which perform particular analyses for particular datasets. These programs call subroutines (kept in the file library.f) to perform standard tasks. The only subroutines specific to a particular dataset are those that read the dataset (i.e. one subroutine for each dataset). The program names refer to the datasets they apply to through the following components of the names: 1) umist, 3) ar, 6) asl.

The types of analyses carried out are the following: plotting a time series, calculating (and, for ar data, plotting a time series of) statistics, calculating and plotting pdfs, calculating autocorrelations, calculating and plotting spectra, deconvolution of concentration time series to remove instrument smoothing and noise effects, smoothing of a concentration time series in order to test deconvolution.

All of these are in Fortran 77 (which ought to be standard except for name lengths and open statements for print files); in addition the programs which convert binary data files into ASCII are written in Pascal. A brief description of the programs and subroutines is given below. Fuller details are contained in comments within the programs. All of the plotting programs here call subroutines from the SIMPLEPLOT graphicspackage,andmany programs call NAG subroutines. Obviously if these packages are not available, or not desired, modifications will have to be made in the relevant places.

Programs:

1) Creating a raw (ASCII) data file from a binary data file.

Programs:	asciias.p, asciiumist.p
Inputs:	binary data file
Outputs:	ASCII data file

2) Plotting time series of concentration.

Programs:	pltasl.f, plotanf, pltumist.f
Inputs:	raw data file
Outputs:	plot of time series of concentration (with several channels on same graph for asl)
Subroutines called:	rdasl, readar, rdumist graph devno, initsp, page, group, chset, endplt (SIMPLEPLOT)

3) Calculating statistics.

Programs: aslstats.f, arstats.f, umiststats.f

Inputs: raw data file

Outputs: first 4 central moments, skewness, kurtosis, intensity, histogram, intermittency in a print file

Subroutines

called: rdasl, readar, rdumist
baseln, condit, stats, histo, threshold

Functions
called: dsum.

4) Calculating pdf.

Programs: pdfasl.f, pdfar.f, pdfumist.f

Inputs: raw data file

Outputs: pdf as a function of concentration, in print file and in unformatted file for storage

Subroutines
called: rdasl, readar, rdumist
baseln, pdfunc

5) Calculating autocorrelation.

Programs: aslauto.f, arauto.f, umistauto.f

Inputs: raw data file

Outputs: autocorrelation as a function of time lag, in print file and in unformatted file for storage

Subroutines
called: rdasl, readar, rdumist
outputauto
gl3abf (NAG)

6) Calculating spectrum.

Programs:	specasl.f, specar.f, specumist.f
Inputs:	raw data file
Outputs:	spectrum as a function of frequency, as a print file and as an unformatted file for storage
Subroutines called:	rdasl, readar, rdumist outputspec, taper gl3cbf (NAG)

7) Plotting time series of statistics.

Programs:	plotarstats.f
Inputs:	raw data file
Outputs:	plot of time series of mean, standard deviation, skewness, kurtosis or intensity
Subroutines called:	readar graph devno, initsp, page, chset, group, endplt (SIMPLEPLOT) g0laaf (NAG)

8) Plotting pdfs.

Programs:	pltpdfasLf, pltpdfar.f, pltpdfumist.f
Inputs:	unformatted pdf files produced by pdf calculation programs
Outputs:	plotted pdfs (several channels or times on each graph for asl and ar respectively)
Subroutines called:	graph devno, initsp, page, chset, group, endplt (SIMPLEPLOT)

9) Plotting spectra.

Programs:	pltaslspec.f, pltarspec.f, pltspecumist.f
Inputs:	unformatted spectrum file produced by spectrum calculating programs
Outputs:	plotted spectrum
Subroutines called:	graph devno, initsp, page, group, chset, endplt (SIMPLEPLOT)

10) Deconvolution to remove effects of instrument smoothing and noise.

Programs: decaslf, decarf, decumist.f.

Inputs: raw data file; unformatted files of spectra of measured concentration and of noise, produced by spectrum calculating programs

Outputs: formatted file of estimated true concentration

Subroutines called: rdasl, readar, rdumist

weight, herrec, herprd

c06eaf, c06gbf, c06ebf (NAG)

11) Smoothing data for testing deconvolution.

Programs: smoothumist.f

Inputs: raw data files

Outputs: smoothed time series

Subroutines called: rdumist

gl3bbf (NAG)

Functions called: dsum

12) Averaging several spectra (e.g. when have short periods of noise).

Program: specave.f

Inputs: unformatted files of spectra to be averaged

Outputs: unformatted file containing average spectrum

Subroutines called: outputspec

Subroutines and functions in library. f:

baseln Shifts an array by a constant amount.

condit Rearranges an array so the values which are greater than a prescribed threshold occur at the start of the array, and returns the positions in the array originally occupied by the values which are not greater than the threshold.

dsum Sums elements of a double precision array.

Graph Plots a selected number of curves on a graph.

Subroutines called: limexc, pen, cvtype, scales, textsz, axes, brkncv, title, setky, line (all SMPLEPLOT)

herprd Forms product of 2 Hermitian sequence stored in form required by NAG subroutines *herrec* Forms reciprocal of a Hermitian sequence stored in form required by NAG subroutines.

histo Calculates and prints a histogram.

outputauto Prints autocorrelation function.

outputspec Prints spectrum.

pdfunc Estimates pdf using a Gaussian kernel with smoothing scale set by equation (3.31) of Silverman (1986), multiplied by a specified factor.

Subroutines called: herprd

M0lanf, c06eaf, c06gbf, c06ebf (NAG)

rdasl Reads an ASCII file of asl data produced by asciiasl.p, and produces an array of concentration values (in nCm^{-3}) containing the time series for several channels.

rdumist Reads an ASCII file of umist data produced by asciiumist.p, and produces an array of concentration values (in $O.lnCm^{-3}$) containing the time series for one channel.

readar Reads an ASCII file of ar data and produces an array of concentration values (non-dimensionalised) containing the time series for all releases in the experiment.

stats Calculates mean, variance, skewness, kurtosis, maximum, minimum and intensity for an array.

Subroutines called: g0laaf (NAG)

taper Tapers (linearly) a specified number of points at the beginning and end of an array towards the mean of the first and last points.

threshold Rewrites an array so any values less than or equal to a specified threshold value become zero, and returns the number of values greater than the threshold.

weight Calculates the weights for a specified weight function, for use in deconvolution.

Other miscellaneous subroutines contained in library.f but not called by the above programs are: contour, error, explik, fncpdf, histog, interm, isum, loglhd, remove_baseline with subsidiary subroutines and functions (these were written at the Meteorological Office - any use should acknowledge them), thrhld.

General input parameters:

Logical
- ifdef = .true, if reading a defiltered (deconvolved) data file
- ifsmoo = true, if reading a smoothed data file
- ifout = .true, if want to print experimental details
- ifpltf = .true, if want to store experimental details

Integer
- nskip = number of seconds skipped at start of file (number of releases skipped for ar)
- nread = number of seconds analysed (number of releases analysed for ar)
- nfreq = sampling frequency of data in Hz.

Format of "raw" data files:

These are the ASCII files containing the unprocessed concentration data. In the case of umist and asl data they are produced from the binary files by asciiumist.p and asciiasl.p respectively.

	Record number	Contents	Fortran format
asl	1	Title of file	a20
	2.	Number of points in time series (including padding zeros)	i8
	3	Digitised concentration values for channels 1-4 at first sampling time	4i5
	4	•• at second sampling time	4i5
		etc.	
ar	1	Experimental parameters	1lel2.6
	2	Experimental parameters	i2,i3,i4,il
	3	Digitised concentrations for first release (at nsam sampling times)	nsam.i5
	4	•• for second release	nsam.i5
		etc.	
umist	1	Experimental parameters and comments	8al,7i4,2il,i4
	2	Digitised concentrations for first second	1000i3
	3	" for second second	1000i3
		etc.	

Fuller details are available from Dr. N. Mole

N.B. The programs as supplied here have undergone a number of alterations to tidy them up and rationalise them for the benefit of other potential users, but have not all been tested in these modified forms. It is possible, therefore, that some minor errors may have been introduced.

The software for certain analyses is still under development, and has not been included (except for some of the miscellaneous subroutines in library.f). This applies, in particular, to baseline removal, maximum likelihood estimation of pdfs, and deconvolution of pdfs to remove the effects of noise.

Appendix E: Proposed new work

Statistical Models of Atmospheric Dispersion

A proposal submitted to the Chemical Defence Establishment, Porton Down - December 1990

SUMMARY

An important part of CDE's work is to further its understanding of the role of concentration fluctuations in atmospheric dispersion. These fluctuations arise from natural variability due to atmospheric turbulence. Previous work, both in association with CDE staff and its contractors, and in other research projects, has led to significant advances in knowledge, and in the quality and quantity of relevant datasets. The present proposal is directed towards the further work that is now needed and its principal aims are:

(a) to advance work already in progress on (i) the intermittency factor; (ii) analysis of data collected by Dr. Jones of CDE in New Mexico and the UK; (iii) new statistical techniques;

(b) to understand the effects of source type and size with particular emphasis on multiple and instantaneous sources (rather than a single continuous source);

(c) the development of sound - but practical - statistical models of atmospheric dispersion that incorporate the results of (a) and (b), and, especially, models of the probability density function (pdf) of the concentration of dispersing gases;

(d) testing the new models against data from a variety of experimental situations and consequent model improvement.

These aims complement well other work being undertaken at CDE and by some of its contractors; the present liaison will continue at or above existing levels. The work throughout will be directed towards the provision of models that can be readily programmed and implemented in the practical tasks of hazard prediction relevant to CDE.

BACKGROUND

An Agreement entitled Fluctuations in Atmospheric Contaminants between CDE and Brunel University began on 1 February 1989. This Agreement was for three years but had to be terminated prematurely because both of the Brunei staff involved in the work left to take up permanent positions at other UK universities. Professor Chatwin will be Professor of Applied Mathematics at the University of Sheffield from 1 January 1991, and Dr. Mole began as a Lecturer at the University of Essex on 14 September 1990. It is now known that Dr. Mole will become a Lecturer at the University of Sheffield (in the same Department as Professor Chatwin), from 1 July 1991.

Work under the Brunei Agreement was progressing very well and has been described in the first (and only) Annual Report in February 1990, and in the Final Report now being typed. The aims of the present new proposal take full account of the success achieved in the work at Brunei. While it is unfortunate that this work had to be curtailed prematurely, it is clearly to the advantage of the new proposal that, if approved, two of the staff involved will be those who carried out the work at Brunel. The proposed starting date of the new Agreement with the University of Sheffield is 1 April 1991 (or as soon as it is possible thereafter to recruit an able and appropriately qualified Research Assistant).

SCIENTIFIC BASIS

The proposal made in October 1987 by Professor Chatwin (which resulted in the Brunel Agreement) emphasized the importance for CDE of much sounder knowledge of the probability density function (pdf) of a contaminant dispersing in the atmosphere. This importance was emphasized in the work description of the Brunei Agreement.

New techniques for estimating the pdf have been developed by Dr. Mole. These emphasise the importance of the intermittency (the probability that the concentration is zero or - more accurately - the probability that the concentration is below a small positive threshold concentration), and of proper consideration being given to the inevitable noise that is present in all datasets and, if ignored, can have severe and misleading effects on models of the pdf derived from data, particularly at small values of concentration. (Note that, in general, the pdfs relevant to CDE's work have the property that "low" concentrations are much more probable than "high" concentrations, thus highlighting the potential that untreated noise has of biasing estimates of parameters like intermittency.

The techniques involve basic theory, both physical and statistical, and, of course, testing against data so that the models can be refined and, ultimately, validated. This work was in progress when the Agreement with Brunel terminated and, in particular, substantial analysis of data collected by Dr. Jones in New Mexico in November 1989 had been performed. However the next stage, which would have been an application of Dr. Mole's methods and subsequent model refinement, was not completed and this is essentially point (a) in the principal aims in the Summary above. It is now relevant that Dr. Jones performed many more experiments in New Mexico in October 1990 (and will perform more in December 1990) under atmospheric conditions characteristic of desert terrains and therefore rarely met with in the UK. The incorporation of the results of these new experiments into Dr. Mole's models and techniques will be a major part of the new Agreement.

Most datasets, including nearly all of those obtained by Dr. Jones, are for the "continuous" releases of contaminant (which, in practice, means releases lasting tens of minutes - up to one hour). However there is strong evidence, both theoretical and experimental, that results obtained for sudden "instantaneous" releases (which, in practice, means releases of a finite quantity of contaminant over a very short time period) are substantially different, particularly in regard to the rms magnitudes of the concentration fluctuations. The relative dearth of datasets for instantaneous releases is due to the cost of obtaining them since, unlike continuous releases, many repetitions are needed to obtain reliable estimates of pdfs. (Some high quality datasets for instantaneous releases are available to Professor Chatwin, but these were taken in wind tunnels using different instrumental techniques from Dr. Jones so that preliminary investigations will be required before their use under the proposed new Agreement can be recommended). Another important factor in practice is the use of multiple (continuous or instantaneous) sources for which, a fortiori, little data are available. Therefore, under (b) of the principal aims in the Summary, it is proposed that the effects on the model development of different types of source (and source sizes) be investigated theoretically (in the first instance) and that, if - as expected - presently available datasets are inadequate for model validation, recommendations for experiments will be made.

The overall aim of the programme is summarised in (c) of the Summary and will require sound scientific and practical judgements to be made in conjunction with CDE staff (Drs. Ride and Jones), based on the results obtained under (a) and (b). As already noted, the whole work programme is linked intimately with data analysis, but it will eventually be crucial for the models to be tested against as wide a variety of data as possible; confidence needs to be established about the applicability of the models. Thus (d) is included separately as an aim of the work.

LIAISON WITH OTHER WORK

From a scientific point of view it is gratifying that the last ten, and especially five, years have seen a rapidly increasing research effort on both pdf modelling and appropriate data collection. CDE, who pioneered some of this work, and Professor Chatwin may feel that their efforts are at last bearing fruit! Principal aim (d) has drawn attention to one aspect of liaison with other workers; more generally it is clearly professionally desirable to take proper account of the results, both theoretical and experimental, that are now increasingly available. Emphasis will of course be placed on work being done by CDE itself, and under its other extra-mural Agreements.

CONCLUSIONS

The programme of work outlined above makes ambitious technical and scientific demands. It is therefore necessary that a well qualified post-doctoral Research Assistant, working under Professor Chatwin's direction and in collaboration with Dr. Mole, be employed for a period of three years, and the provisional costings (attached) take account of this. It should also be emphasized that, in this field, personal contact (especially at conferences in the UK and abroad) with other researchers worldwide results in significant improvements to what would otherwise have been achieved in any single research programme; this explains the items under travel in the costings.

P.C. Chatwin

14 November 1990