Journal of Hazardous Materials, 16 (1987) 457–500
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

A USER'S CRITIQUE OF THE THORNEY ISLAND DATASET

P.W.M. BRIGHTON

UKAEA Safety and Reliability Directorate, Wigshaw Lane, Culcheth, Warrington, Cheshire WA3 4NE (Great Britain)

(Received November 27, 1986; accepted March 18, 1987)

Summary

The primary purpose of the Thorney Island Heavy Gas Dispersion Trials was to obtain reliable data at large scale to test the validity of mathematical and physical models. Secondly, the trials were intended to obtain data for improving the understanding of the physical mechanisms in heavy-gas dispersion and for testing the fundamental hypotheses in mathematical models.

This paper is a review of the extent to which these objectives have been met. Three major issues are addressed. First, the reports of the experiments and the raw data are examined to assess the magnitude of experimental errors. Second, a comparison is made of different estimates of the inherent statistical variability of the concentration field, from wind-tunnel tests and from the trials themselves. Third, the results of several analyses of the experiments to derive overall cloud properties, such as area and speed, and model parameters, such as entrainment coefficients, are compared to assess how far the data permit divergences in interpretation.

Rather than to present conclusions, it is intended to raise questions which users of the data should bear in mind in conducting their own analysis and drawing their own conclusions.

1. Introduction

1.1 The background of the Thorney Island trials

When field trials started at Thorney Island in 1982, mathematical modelling of heavy-gas dispersion was already a sunrise industry. In their comprehensive review of the state of the art Wheatley and Webber [1] described 45 models in existence at that time. Several of these models were used to predict the outcome of the Thorney Island trials and the range of results for concentration was one to two orders of magnitude (McQuaid and Roebuck [2], Chapter 10).

This extraordinary proliferation of mathematical models was presumably the result of the paucity of reliable experimental data. Definitive data was needed to sort out the sheep from the goats and allow a much clearer assessment of the risks of major accidents in the chemical and petrochemical industries.

For the quantification of those risks is, of course, the goal to which heavy-gas dispersion research contributes. It is only one link in a long chain of rea-

0304-3894/87/$03.50

soning — from initiating event such as a process perturbation or equipment failure, through a possible cascade of subsequent failures in a typical event-tree, to the complexities of the release mechanism, through a possible phase of heavy-gas dispersion to safe dilution before ignition or combustion and explosion, or to a subsequent phase of passive dispersion followed by the mechanisms of toxic attack within the human body.

At the first Symposium in a paper with a very similar title to this, Harris [3] reviewed the place of the Thorney Island trials in this wider context. I do not dissent for anything he said and do not wish to add anything from the point of view of the ultimate user, the risk analyst or decision maker. Rather, I shall take the view-point of the immediate user of the data, those involved in validating mathematical or physical models, or in developing new ones, which will be the main form in which the outcome of the Thorney Island trials will reach the end-user.

1.2 The aims of the paper

This paper originated in a request from Dr. Jim McQuaid as Symposium organiser for "a user's guide to the unwary". In pursuit of this aim, I have carried out a review of the experimental errors in the basic data of the trials, the gas-concentration records (see Section 3). The performance of the gas sensors was extensively examined by McQuaid and Roebuck [2] and in papers at the first Symposium [4,5]. The trials records themselves are also relevant so I thought it useful to extract and summarise all this information in one place.

The paper has developed into a review of further aspects relevant to the problem of validating models of heavy-gas dispersion. How precisely can they be validated by this data? Will the proliferation of models mentioned above be brought under control by the verdict from experiment? In the light of analyses of the Thorney Island data presented at this Symposium and earlier, it is apparent that the data may well permit significant divergences in interpretation. So I have tried to identify all the significant sources of uncertainty in the trials. The view-points presented will be largely personal, though I have consulted other users of the data in forming them. The review may well be incomplete, and rather than to present conclusions, I intend to raise questions which users of the data should bear in mind in conducting their analysis and drawing their own conclusions.

1.3 The objectives of the Thorney Island trials

Insofar as this paper is a critique, it will judge the results of the experiments against the original technical objectives. These were twofold (McQuaid and Roebuck [2], chapter 4). The first was:

"to obtain reliable data at large scale with which to test the predictive capability of mathematical and physical models. Such data comprise primarily the

distributions of concentration as a function of time and position for a variety of weather conditions and the meteorological parameters required to specify the weather conditions."

This objective is essentially backward-looking in that it presupposes that the user has already developed models on the basis of earlier information. The second objective is forward-looking in being directed at the improvement and development of models:

"to obtain data with which to improve physical understanding of the mechanisms of heavy-gas dispersion and to test the fundamental hypothesis in mathematical models. Such data comprise measurements of turbulent fluctuating velocity and concentration distributions and photographic records of cloud behaviour, in addition to the data needed for [the first objective]." The objectives were separated because of the different types of instrumentation needed.

2. Requirements for the validation of models

With the large number of mathematical models available and the increasing amount of wind-tunnel and field-trial data, it is becoming increasingly important to decide what can be regarded as adequate validation. The visual comparison of experimental points and theoretical curve is no longer sufficient. A thought-provoking review of the principles of validation has been given by Mercer [6]. Questions that he raises will be echoed throughout this paper. One additional point that I would like to make here is that there is an inverse relation between the complexity of models of heavy-gas dispersion and the complexity of the processing of experimental data needed to validate them.

The most widely used models are the integral or box models, which represent the dispersing cloud by a single volume of gas of uniform concentration. As pointed out by Chatwin [7], these make no pretence at predicting concentrations at individual positions*. Thus the basic data gathered in the trial cannot be used directly: rather complex data-processing must be carried out to determine quantities such as cloud area, mean concentration and translational speed (see Refs. [9] and [10] for example). There is considerable scope for diverging interpretations of the data both because these quantities can be defined in different ways and because they are subject to significant sampling errors since the spatial coverage of the concentration field was limited.

The other type of mathematical model is the three-dimensional time-dependent numerical solver of the Navier–Stokes equations with a suitable turbulence closure scheme. In principle, it is much easier to compare these with

*However, it appears that advanced box-models with generalised concentration profiles can give a fairly convincing reproduction of individual concentration records [8].

the experiments since they can calculate the time-history of concentration at individual points. Physical wind-tunnel modelling also has this feature.

With reference to the technical objectives of the trials described in Section 1.3, I believe it is also the case that the simpler models of gas dispersion need a larger input of physical understanding. Individual fluid-dynamic mechanisms such as gravity-current propagation, entrainment, advection, shearing by the wind and cloud elongation need to be understood and explicitly modelled.

With properly formulated three-dimensional codes, most of these processes should be predicted automatically through solving the fundamental physical equations. The exception for which physical understanding may be required is the formulation of Reynolds-stress closure methods which take proper account of the influence of stable density gradients on turbulence. Box models also use hypotheses on this in the formulation of the entrainment relations.

Finally, wind-tunnels probably give least occasion for requiring physical understanding explicitly (I do not wish to imply that wind-tunnel modellers do not need physical understanding). The main doubt about scaling down to wind-tunnel size is whether molecular viscosity and diffusity introduce any systematic deviations.

In view of the greater difficulties in validating box-models, what follows will be weighted towards that question and there will be relatively little mention of the Phase II trials, in which obstacles were present. However, a factor affecting all model comparisons is the statistical variability affecting instantaneous releases into a turbulent flow, a topic which has received much greater attention at this Symposium than at the first. The result is that in practice overall properties of the gas clouds may be as important for validating the complex models as for the box-models (see Section 4 below).

3. Experimental errors

3.1 Concentration measurements

The gas sensors used were electrochemical oxygen-deficiency cells, chosen after an extensive review of the possibilities. Leck and Lowe [4] have given a detailed account of the operating principles of these instruments and of the laboratory assessment of their performance. Performance in the field and subsequent validation of the data is also described, with additional details, in McQuaid and Roebuck ([2], Chapter 7 and Sections 15.7 and 16.2).

The cells respond to the concentration of oxygen in the gas, giving an output of around 0 to 0.5 V in its absence and around 7 to 8 V in air. These two points define gas concentrations of 1 and 0, respectively. (To avoid confusion in the text of this Section 3, I shall express concentration as a volume fraction and reserve the use of percentages to express relative error and relative humidity.) In the field each instrument was calibrated every month and before each trial, using the reading in air and a reading in pure nitrogen. This gave the individual

factors used to produce concentration values in the magnetic data from the voltages actually recorded, assuming a linear response.

There were several sources of error in this procedure: non-linearities in the response to intermediate concentrations; the finite response time due to diffusion times within the cell; noise in the instruments themselves and in the data transmission system; the digitisation in the analogue-to-digital converter producing round-off errors; and drift in output voltages caused by changes in ambient temperature and humidity. Each of these affected the measurements in a different way and the effects cannot necessarily be summed to give an overall error bar. Each of these is now reviewed in turn.

Non-linearity

The extent of non-linearity varied randomly from cell to cell. Statistical data on these errors at gas concentrations of 0.25 and 0.75 are given by Leck and Lowe [4] and McQuaid and Roebuck [2]. At the lower concentration, the absolute concentration readings of 170 cells had a mean linearity error of -1.5×10^{-3} with a standard deviation of 2×10^{-3}. Thus these errors have 95% confidence limits of about -5×10^{-3} to $+2.5\times10^{-3}$. Absolute errors are of a similar size for concentrations of 0.75. For low concentrations, this non-linearity error can be assumed to tend linearly to zero at zero concentration by definition of the calibration method. (Changes in the output voltage at zero concentration are discussed under "drift".) Thus for concentrations below 0.25, non-linearity produces a *relative* error in concentration with confidence limits of about -2% and $+1\%$. These limits include the effect of errors in measuring the gas concentrations used for preparing mixtures for the non-linearity checks: according to Leck and Lowe, the blending pump had a relative accuracy of 0.05%.

Linearity tests with mixtures including Freon 12, the heavy-gas agent in the trials, showed that this had an adverse effect on linearity. Using a contaminant gas of 33% Freon 12 and 67% nitrogen, over a concentration range of 0 to 0.3, Leck and Lowe [4] found that the relative error in concentration had 95% confidence limits of -9% and $+1\%$, indicating a systematic relative overestimate by 4% on average.

This error band may be rather pessimistic because the influence of Freon 12 in reducing the sensitivity of the electrochemical cell to oxygen was quite slow, taking about 4 min to stabilise. This effect can be seen in records from sensors inside the gas bag (see Fig. 1a) and for this reason these could not be used to measure concentrations after release.

A disturbing doubt remains about whether this effect could have affected measurements outside the gas bag. From Trial 17 onwards, there were sensors as near as 10 or 20 m to the spill point which detected gas concentrations over 0.3, corresponding to Freon levels said by Leck and Lowe to give the maximum non-linearity. The short duration of the exposure might encourage hope that

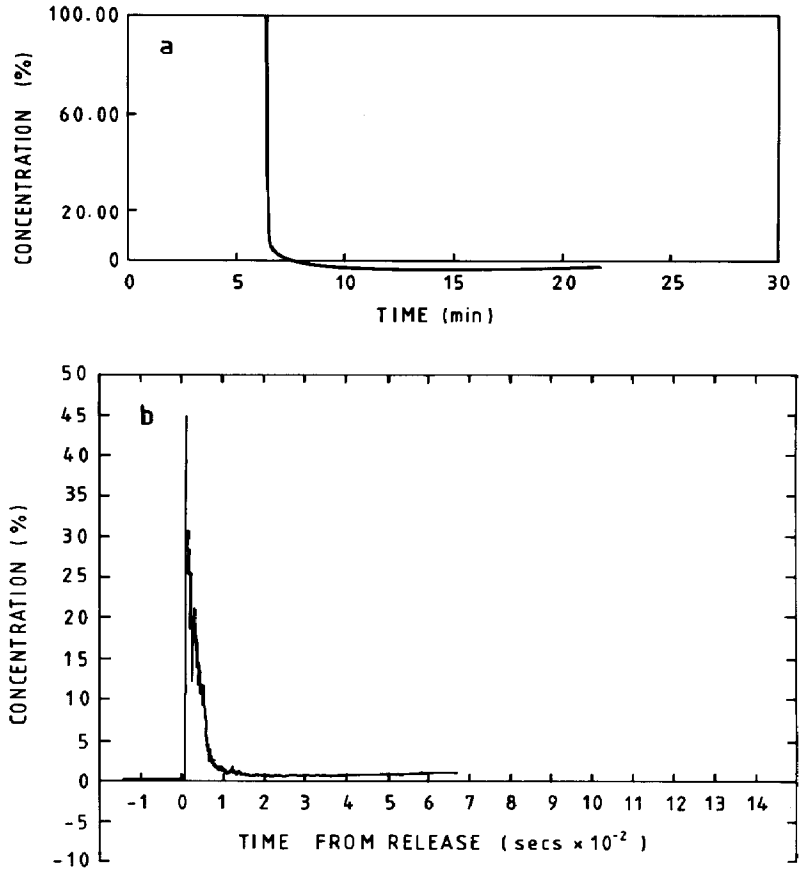

Fig. 1. Effects of high Freon levels on gas sensor measurements: (a) Concentration record from gas sensor at height 4.5m in the gas container in Trial 8; (b) concentration record from x = 400m, y = 230m, z = 0.4m in Trial 19.

the Freon would not have its full effect on sensitivity, but the concentration traces show a very slow return to zero with apparent departure times much later than those from gas sensors further away from the source [11]. The exposure to Freon sometimes seems to cause an absolute error in concentration up to 0.01 when a return to zero would be expected (Fig. 1b).

Finally, McQuaid and Roebuck ([2], Section 7.5.2) give some results from linearity checks carried out on samples of sensors from the field. Using a mixture of Freon 12 with dry air, they found that a gas concentration of 0.095 to 0.096 was measured with relative-error 95% confidence limits of $\pm 9\%$. With a 2:1 nitrogen–Freon 12 mixture as the contaminant gas, measurements of gas concentration of 0.0865 had 95% confidence limits of $\pm 19\%$ on the relative error. They do not comment on the apparent divergence from the laboratory

results in the report, but B. Roebuck (personal communication) points out that, whereas the sensors had been calibrated in the humid ambient air, the linearity checks were made with dry mixtures and therefore were subject to the random drift due to humidity changes, as discussed below. This can account for the errors found in the field linearity checks, which did not therefore actually achieve their purpose.

Frequency response

The rapidity with which the electrochemical cells respond to a step change in gas concentration depends on the rate of diffusion across the membranes to the cathode. To achieve the specified frequency response of 1 Hz, it was necessary to use electronic enhancement. Leck and Lowe [4] give full details of this, as well as of the design changes made to produce high-speed gas sensors with a 10 Hz response. A frequency response of 1 Hz corresponds to a rise time of 0.35 s from 10% to 90% of a step change. Data was recorded from each instrument in the field 20 times a second. This was much faster than necessary for the standard gas sensors, but allowed rapid-response instruments to be moved between data channels without restriction.

The effect on concentration measurements of response time and of time averaging in subsequent analysis has been studied in some detail by Nussey et al. [5]. They analysed the attenuation of the peak concentration for all gas sensors in Trials 6–19 as a function of averaging times up to 6.6 s. Compared to an average over 0.6 s, reductions by more than half were found in some cases. Use of averaging over 1.8 s produced reductions by a factor of around 0.8 to 0.9 at the upper quartile.

From an analysis of the spectra of the concentration records Nussey et al. concluded that the frequency response in the field was not significantly different from that found by Leck and Lowe [4] in the laboratory. They also sampled raw data to determine the rms noise level at around 0.03 to 0.11%. They concluded that averaging times of 0.5–1 s would improve the results from the standard gas sensors by suppressing the noise without degrading the response to the real signal.

The real errors due to finite frequency response are highly dependent on the finest structure in the turbulent cloud, on which there is no direct information: Carn and Chatwin [12] give an account of current theoretical approaches and wind-tunnel information.

Noise and digitisation

Leck and Lowe [4] and McQuaid and Roebuck [2] report that all instruments were tested for noise by sampling the output for one minute. The maximum peak-to-peak noise was found to be equivalent to an absolute error of $\pm 2.5 \times 10^{-4}$ in gas concentration.

Analogue signals had to be transmitted over distances up to 60 m before

conversion to digital data. No information is available on noise levels resulting from this process. The digitiser could accept input in the range ± 10.235 V with a resolution of 12 bits. Since the gas sensors operate over a span of 7–8 V, the concentration data is effectively recorded in units of about 6×10^{-4}–10^{-3}, as can be seen by inspection of the 20 Hz records. Thus each individual item of data has a uniformly distributed round-off error up to ± 3–5×10^{-4}. As explained above, it is recommended that this data be averaged in groups of 10 to 20 readings. If the round-off errors are independent, then the error should be reduced by a factor 3 to 4 on average. However, particular signal forms could have round-off errors after time averaging up to the maximum for individual readings.

Drift due to changes in ambient temperature

It was possible to control the sensitivity of the gas sensors to temperatures in the range 5–30 °C by adjusting a compensating thermistor to reduce absolute gas-concentration errors to less than 1×10^{-3}/K (Leck and Lowe [4]). This was achieved at steady temperatures for most sensors but it was found during temperature cycling that additional error was incurred during changes in temperature because of a slight mismatch between the temperature characteristics of the cell membrane and the compensating thermistor. Leck and Lowe give histograms of the static and dynamic temperature drifts for the whole set of instruments used. They express dynamic errors in terms of equivalent gas concentration per K. The actual response of the cell to changing temperatures will depend on how the characteristic time for the temperature change compares with the timescales of response to components of the cell. In Leck and Lowe's [4] Fig. 13, the temperature cycle involved a linear change lasting 200 min, which seems likely to be much longer than the instrument's temperature response time. In this limit dynamic errors should be proportional to the rate of change of temperature. So the abscissa of their Fig. 14a must be multiplied by 200 min to give change in gas concentration reading per K/min. This gives an average dynamic drift of 0.8 K^{-1} min in gas concentration readings, i.e. a temperature rise of 1 K/min corresponds to a drift in reading of 0.8, or 0.01 K/min to a drift of 0.008.

For temperature changes over periods similar to the response time, this approach will give an overestimate of the drift. For rapid temperature jumps, only the size of the temperature change will be relevant, not its rate. For oscillatory or fluctuating ambient temperatures, the response will be more complex. The information presented by Leck and Lowe does not seem appropriate for estimating dynamic temperature drifts in any of these situations.

Drift due to changes in relative humidity

Changes in relative humidity (RH) mean changes in the composition of the air. The resultant changes in oxygen concentration will affect the calibration

of the oxygen deficiency cells, giving an equivalent absolute error in gas concentration of 0.05 with RH changed from 0 to 100%. Leck and Lowe [4] found this expectation borne out in tests on prototype gas sensors, but further tests on production instruments showed considerably greater errors. They were tested in an environmental chamber cycling between 50% and 95% RH. The first cycle produced very large errors, but this stabilised on subsequent cycles to a level of the order of 0.1 absolute change in gas concentration reading for the 45% difference in RH. To limit the effects of this shortcoming, only cells with an absolute error of 0.12 for this RH cycle were selected for use in the field. This is equivalent to an absolute error of about 2.5×10^{-3} in gas concentration reading for a 1% change in RH.

Effects of noise and drift in the field

During each trial, the calibration factor for each gas sensor was determined from the voltage reading in ambient air just before the release. (The instruments were permanently switched on to avoid warm-up effects.) Environmental variations then might produce drift during the main data-collection period of up to 30 min. Analysis of the drifts throughout Trials 4–29 showed that the extent of this drift was less than 5×10^{-3} in gas concentration reading in 75% of cases and less than 0.01 in 90% of cases (McQuaid and Roebuck [2], Section 7.5.2). This shift of the voltage level in ambient air also means an error in the calibration factor.

McQuaid and Roebuck [2] also note that the response of cells to pure nitrogen changed somewhat between calibrations by up to 0.2 V, producing a potential relative error of up to $\pm 3\%$. Calibrations were carried out at monthly intervals or more frequently if a number of trials were carried out in rapid succession [4].

Before issue of data to sponsors, raw gas-sensor output voltages were converted to gas concentrations using the air voltage reading immediately before release and the most recent pure nitrogen reading. A linear drift correction was then applied which generally eliminated drift effects throughout the concentration record (see Ref. [2], Section 15.7.1 and Fig. 15.5). However, in processing all the data, I have found that a significant number of concentration records displayed apparent non-linear patterns of residual drift which could make interpretation of the data difficult, particularly when trying to determine departure times for gas (cf. Ref. [11]).

Figure 2 shows some examples from Trials 9 and 13 to illustrate these difficulties. In Fig. 2a, it is fairly easy to distinguish both an arrival time and a departure time (about 940 s), but in between the non-linear drift produces significant negative concentrations (or possibly this is an example of exceptionally bad noise). In Fig. 2b, there are again negative values in the middle of the record but the broad secondary peak around 700–1000 s could be mainly the effect of non-linear drift. The sensor immediately above this on the same

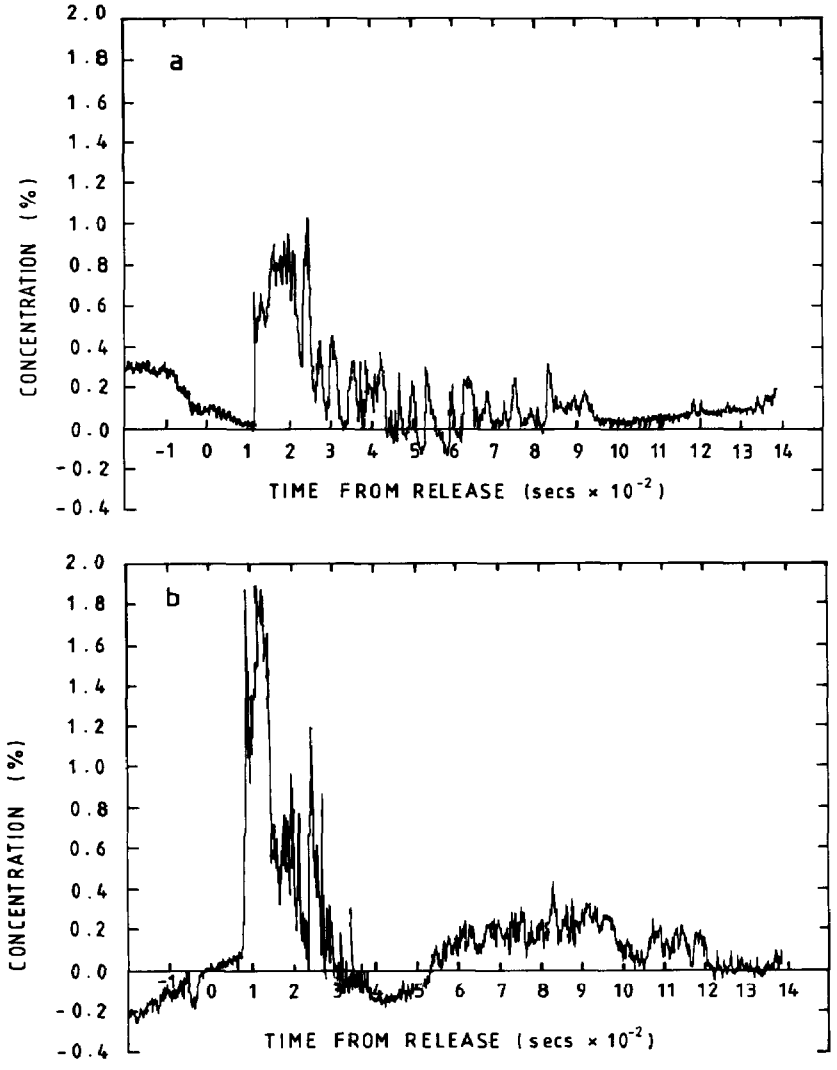

Fig. 2. Examples of standard gas sensor records showing non-linear drift: (a–e) Trial 9; (f–h) Trial 13.

mast also has a broad secondary peak, but centred at about 500–600 s (Fig. 2c). In this case the drift correction seems to have been carried out on the assumption that the true signal ended around 100 s after release. Trial 9 was characterised by a low windspeed and a long duration of significant gas concentrations at many ground-level sensors. In these circumstances the linear drift correction may be indeterminate because recording of data ended before zero concentration was reached (Figs. 2d and 2e).

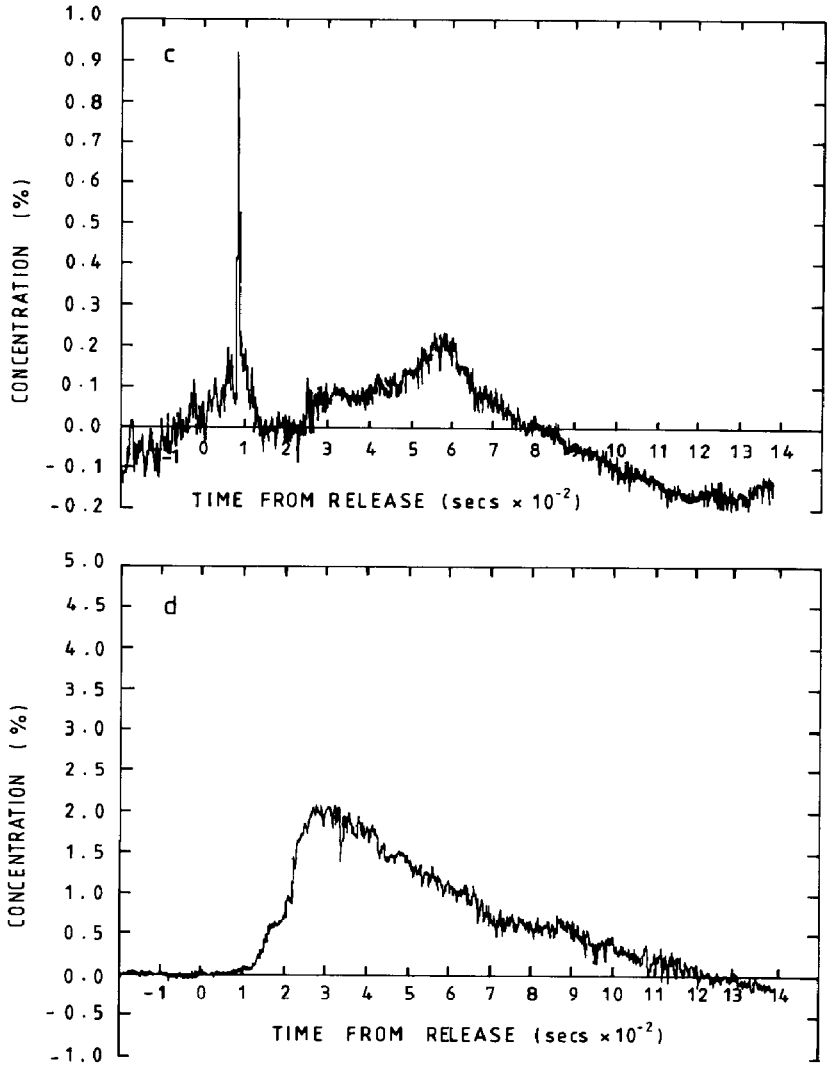

Fig. 2 (continued).

In high wind-speed cases such as Trial 13 non-linear drift tends to be less troublesome because signal durations are shorter, so even quite violent drift does not make it too difficult to distinguish the bulk of the main peak (Figs. 2f and g). However, the relative uncertainty in departure time may sometimes be equally great (Fig. 2h).

Trials 9 and 13 seemed particularly prone to difficulties of this kind — Figure 2 does not illustrate by any means all of the bad apples. This may be associated with particularly variable conditions — Figure 3 shows typical traces of air

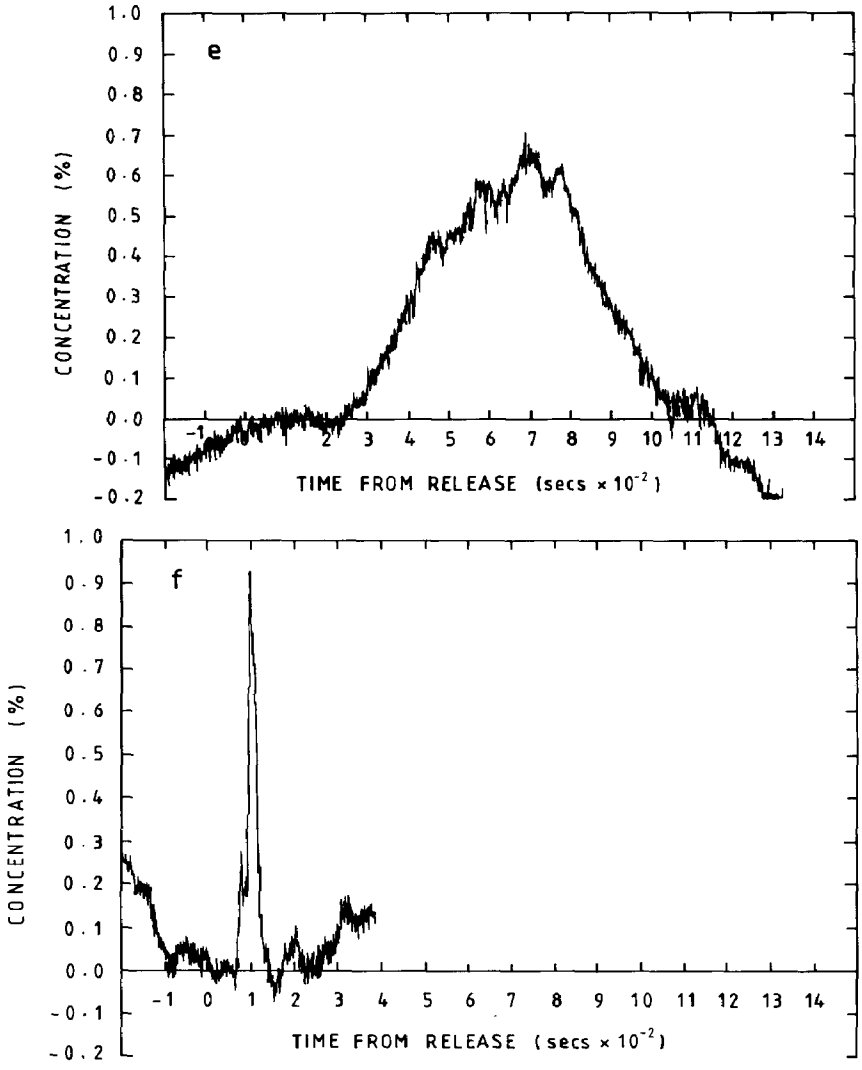

Fig. 2 (continued).

temperature and RH for these trials. In Trial 9, the overall rate of change of temperature was about 0.025 K/min and RH changed by almost 2% during the release. From my interpretation of the laboratory tests on the instruments above, this could give variations of 0.02 in gas readings due to dynamic temperature drift and 0.005 due to RH drift. The non-linear components of these might be smaller. It was noticed in this trial that the orange-marked gas cloud appeared to be transformed to a persistent white low-lying mist. This suggests condensation onto the smoke particles — this may also have affected instrument drift.

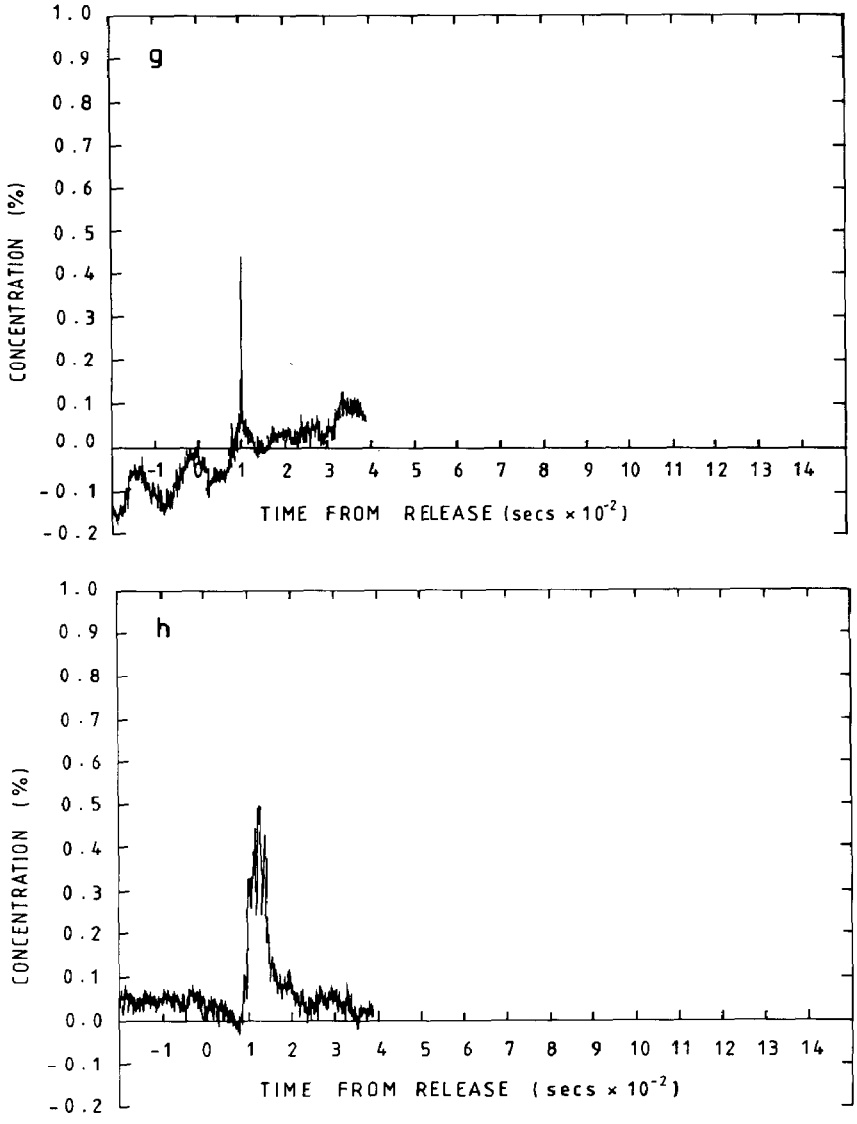

Fig. 2 (continued).

In Trial 13 there were abnormally large oscillations of temperature by ± 0.1 K and of RH by $\pm 0.5\%$: this may account for the particular pattern of drift seen in Figs. 2f and 2g, which also occurred in several other concentration records in that trial.

To get an overall impression of the extent to which these problems may have affected the results, I have made a rough visual survey of the amount of residual

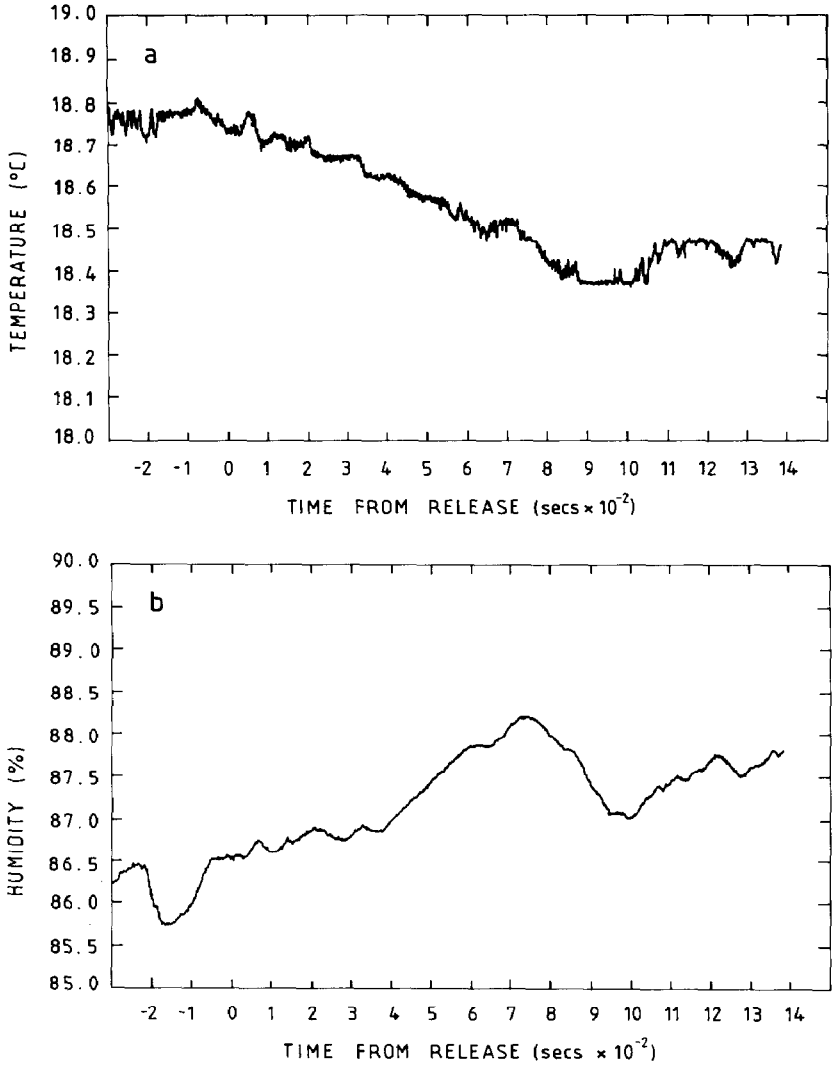

Fig. 3. Ambient temperature and relative humidity in: (a,b) Trial 9.

non-linear drift from five trials. It was defined as the change in apparent zero level over the full record, i.e. maximum minus minimum. At the same time, I estimated the noise level on the traces, defining noise as irregular fluctuations with a period less than 100 s, as seen on all traces before the arrival of gas. The estimates were made using the "hard-copy" volumes of data issued by HSE, so only sensors with peak gas concentrations less than 0.02 had a large enough scale to estimate noise: drift estimates were made for a somewhat larger class of sensors.

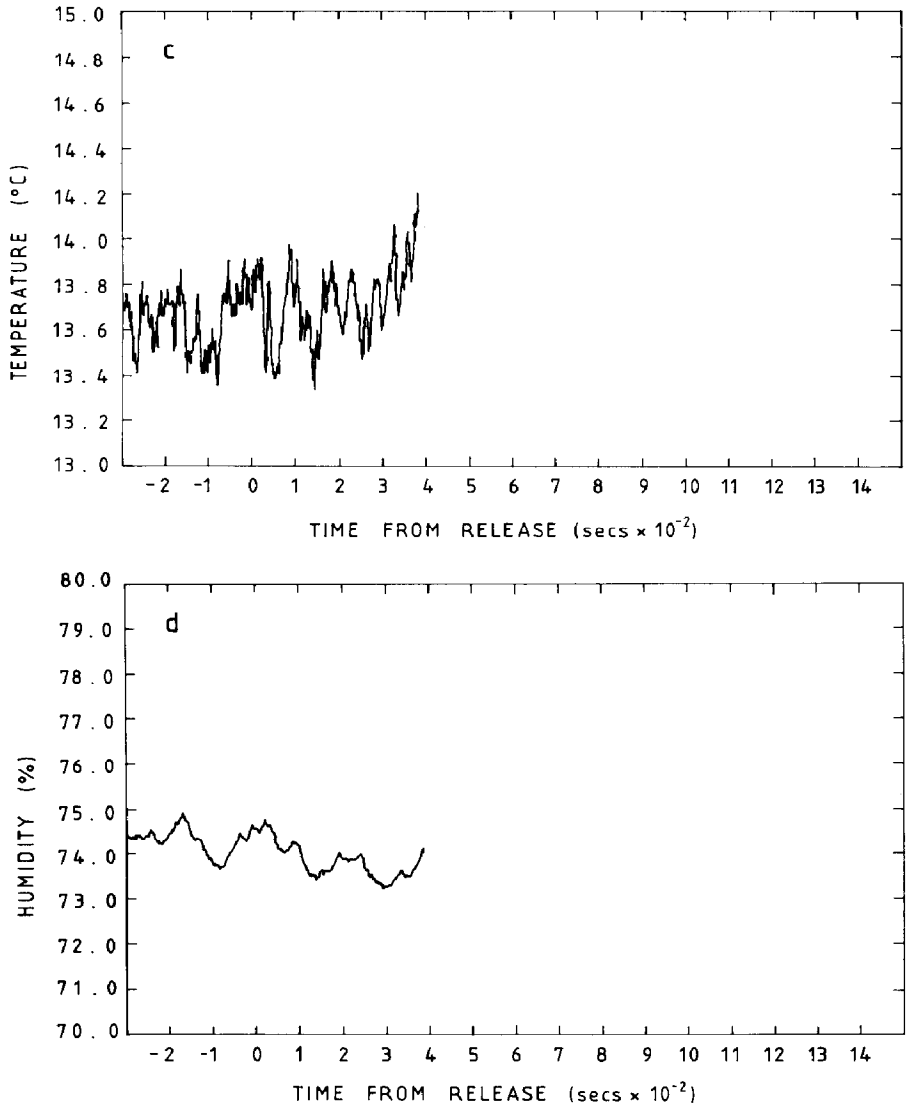

Fig. 3. (continued). (c,d) Trial 13.

The results are listed in Tables 1 and 2. The noise results were generally equal to or better than the noise level of $\pm 2.5 \times 10^{-4}$ measured in the laboratory by Leck and Lowe [4]. These values derived from field results averaged over 0.6 s also include the effect of digitisation.

Residual non-linear drift is quite variable from trial to trial but is often at an average level of $1\text{–}2 \times 10^{-3}$, enough to affect seriously determination of low concentrations. The good performance in Trials 7 and 19 seems to be associ-

TABLE 1

Noise levels in gas concentration records determined by visual inpection

Trial	Date	Frequency of occurrence (% of total) of each noise level (half of peak-to-peak range, units 10^{-4} or 0.01%)							Total examined
		≤1	1.5	2	3	4	5	≥6	
7	8/ 9/82	9	26	47	15	4	–	–	47
9	15/ 9/82	7	20	41	22	2	5	2	41
13	19/10/82	–	3	15	24	21	27	9	33
19	10/ 6/83	2	7	37	41	13	–	–	46
34	6/ 3/84	16	26	32	16	–	–	10	31

TABLE 2

Residual non-linear drift in gas concentration records determined by visual inspection

Trial	Date	Frequency of occurrence (% of total) of each level of non-linear drift (units 10^{-2} or 1%)							Total examined
		0	0.05	0.1	0.2	0.3	0.4	≥0.5	
7	8/ 9/82	74	11	11	4	–	–	–	47
9	15/ 9/82	–	6	21	32	17	9	15	53
13	19/10/82	14	12	36	26	7	–	5	42
19	10/ 6/83	55	20	16	5	2	–	2	56
34	6/ 3/84	35	8	18	18	8	8	6	51

ated with steadier environmental temperatures. In Trial 34, we found during processing the data that several concentration records appeared to have a significant residual *linear* (e.g. Fig. 4). Evidently, our judgement differed from HSE's in this case and for the results of Brighton and Prince [13], a further linear drift correction was applied. These cases are not included in Table 2.

To conclude this discussion of drift effects, I would like to point out that the gas in the container often differed in temperature from ambient by a degree or two — this is clear from a discontinuity at the moment of release in traces from the thermocouples in the container. This could easily lead to additional temperature changes at sensors of 0.1 K/min, which from our discussion above could cause changes in gas concentration reading of 0.08. This would be correlated with passage of the gas and should disappear afterwards so as not to be at all obvious. However, it may well be that this interpretation of the dynamic temperature drift effect is pessimistic and discussed above, the laboratory data do not apply to temperature changes as rapid as those accompanying first arrival of the gas at a sensor.

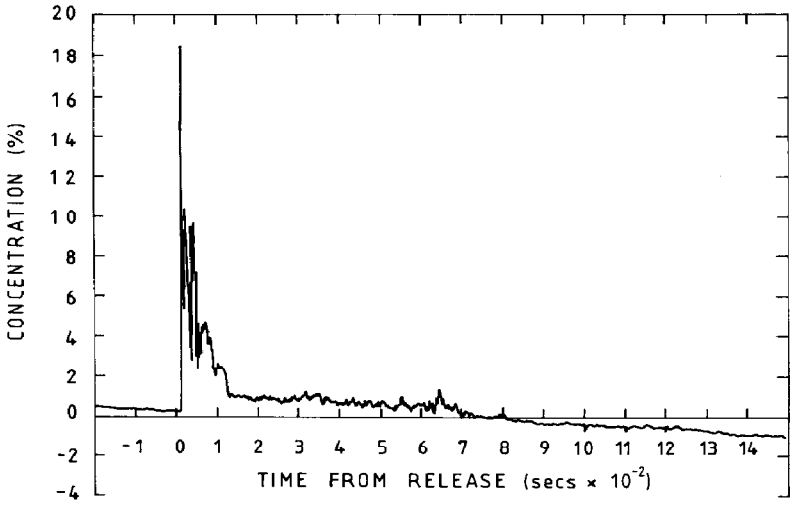

Fig. 4. Example of residual linear drift in Trial 34.

Summary of errors in standard gas sensor measurements

I have attempted above to collect together all the information, in Leck and Lowe [4], McQuaid and Roebuck [2] and in the concentration data themselves, which is relevant to assessing the errors in the concentration measurements. The initial laboratory testing and the field experience appear to give a fairly complete and consistent detailed account of the sources of error and the reasons for them. Because errors arise in different ways and have different effects, it is not possible to lump them together into a single figure. Here I summarise the results of the survey with particular reference to errors in measuring low gas concentrations, and I also note some problems on which the reports on the trials do not perhaps give sufficient information.

- Non-linearity seems to be a systematic error for a particular instrument but varies widely between instruments. Laboratory tests in the presence of Freon give 95% confidence limits on the relative error of -9% and 1%. A problem here is an apparent delayed effect of Freon on the sensitivity of the sensors — there are inconsistent accounts in Leck and Lowe [4] and in the "Notes on validation" in the hard-copy data volumes.
- The frequency response of the instruments does not introduce any error into time averages of gas concentration over more than 0.5 s.
- Noise and digitisation errors give an absolute error level of less than $\pm 2.5 \times 10^{-4}$ in gas concentration in most cases. This means that, in the absence of unfavourable drift behaviour, the resolution of the instruments is normally considerably less than the value of 1×10^{-3} originally specified (cf. McQuaid and Roebuck [2], Figs. 16.3 and 16.4).
- Drift of the instrument calibration occurred because of variations in temper-

ature and relative humidity. A linear component of drift was removed by HSE before issue of data, but this left a residual relative error of $\pm 1\%$ at the 90% confidence error due to the change in calibration. Changes in the response of cells to pure nitrogen, of unspecified cause, gave relative errors in gas concentration up to $\pm 3\%$.

The remaining non-linear component of drift appears to depend on the variability of ambient temperature and so varied considerably from trial to trial. In bad cases, it introduced typical absolute errors of $1\text{–}2 \times 10^{-3}$ in gas concentration, particularly in low-wind speed cases when the signals lasted a long time. This is the most important single source of error and it would be valuable if a complete and systematic survey of these drifts were carried out.

A further, hidden component of drift may be present due to temperature differences between the released gas and the atmosphere. It is difficult to use Leck and Lowe's [4] account of dynamic temperature drift for this situation, but it can be interpreted to imply larger errors than any others listed above.

3.2 Meteorological measurements

The meteorological instruments were mostly standard designs, whose characteristics are described in Section 10.5 of McQuaid and Roebuck [2]. Ten sonic anemometers were also deployed to make detailed turbulence measurements and several corrections are needed to the data from these as described by Puttock and Colenbrander [14] and Puttock [15]. Also these instruments tended to produce certain types of spurious signal easily recognised by visual inspection. McQuaid and Roebuck [2] give more details of meteorological transducer faults discovered in Trials 5–14. Full information on transducer performance has recently been collected together for the full trials programme in a report available from British Maritime Technology Ltd. [65].

Davies and Singh [16] and Puttock [15] have analysed the meteorological data in order to relate the atmospheric conditions to standard categories such as those of Pasquill or to Monin–Obukhov similarity theory. The extent of mismatch is not really an experimental error, but does constitute a deviation between the experimental conditions and the idealisation embodied in mathematical models or indeed wind-tunnel simulations. Also temporal variations of conditions is one source of the variability discussed in Section 4.

3.3 Other experimental errors

An important source of uncertainty in the measurements is the accuracy with which the initial conditions could be specified. Estimates of the uncertainty in volume released and in initial density are given by McQuaid and Roebuck [2] in Table 15.1 for Phase I and in Table 20.1 for Phase II.

Finally, it is worth mentioning a problem affecting interpretation of photographic records of the Trials. In Phase I, the stills photographs were imprinted with a time that changed only every minute. As a result it is impossible to

determine the time since release for a given frame to a precision greater than the interval between frames, unless the container was caught in the act of falling in one frame. In addition it has been found [17] that although directed by a radio system to fire at a steady rate, the ground cameras were in the habit of taking extra shots at indeterminate times.

4. Experimental variability

4.1 The statistical nature of turbulent dispersion

Although governed by deterministic equations, turbulent flows have in practice a high degree of randomness. Immeasurably small differences in initial conditions and boundary conditions cause large differences in individual measurements in the flow-field for different runs of the same experiment. Moreover in experiments like those at Thorney Island the initial conditions were prescribed only by means of a few statistics of the atmospheric velocity field: the detailed eddy structure would be quite different from run to run. Chatwin [18] argued that such statistical fluctuations are very important for assessing hazards. Clouds may be ignitable because of local high concentrations even when the mean cloud concentration is well below the flammable limit. Ride [19] and Griffiths and Harper [20] have discussed how the non-linear dependence of toxic gas effects on concentration implies that a fluctuating concentration may give much higher mortality than the equivalent mean concentration.

Whatever the importance of fluctuations in hazard assessment, Mercer [6] has pointed out that they must be taken into account when comparing any model predictions, whether from the computer or from the wind-tunnel, with releases in field trials. Each field-trial result is just one realisation of the ensemble of possible outcomes, whereas models produce either an ensemble mean or a different member of the ensemble. The magnitude of the concentration fluctuations sets a limit to the level of agreement which can be achieved between model and field trial. In the case of the instantaneous releases at Thorney Island, it appears that variability due to turbulence may be a considerably more important source of uncertainty in using the results than the measurement errors discussed in Section 3 above. In this Section I will review results on statistical fluctuations from wind-tunnel work and from the Thorney Island Trials and assess the resultant uncertainty in using the Trials for comparison with models. Except where otherwise stated, measurements were taken at ground-level or as near as practically possible (0.4 m at Thorney Island).

4.2 Replicated wind-tunnel experiments by Meroney and Lohmeyer

Most of the available information comes from the wind-tunnel experiments of Meroney and Lohmeyer [21,22]. They made instantaneous releases with density 4.17 times that of air both in calm conditions and in wind speeds up to

1 m/s. Concentrations were measured at a single fixed location for each release. Releases in the same conditions were made five times (or thereabouts) with each of several measuring locations. With one particular location 95 repeats were made to obtain details of the probability distribution. Meroney and Lohmeyer [21] tabulated means and standard deviations of maximum concentration for each release and of the times of arrival, maximum concentration and departure. In the subsequent paper [22], they give plots of this information in various forms, including probability distributions from the 95 replications, and a plot of $(\overline{C'^2_m})^{1/2}/\bar{C}_m$ against downwind distance, where C_m denotes the maximum concentration measured in each member of the ensemble.

To apply these results to Thorney Island, some simple scaling can be applied. Denoting conditions in the Thorney Island experiments by a subscript "TI", and in Meroney and Lohmeyer's experiments by "ML", we first note the V_{TI} can be taken as 2000 m^3 and that three values of $V_{ML}=35$ cm^3, 165 cm^3 and 450 cm^3 were used. The downwind distances x are scaled using $x_{TI}/x_{ML}=(V_{TI}/V_{ML})^{1/3}$. We scale velocities by maintaining equality of the Richardson number [23] giving

$$\frac{\Delta'_{ML}\, L_{ML}}{U^2_{ML}}=\frac{\Delta'_{TI}\, L_{TI}}{U^2_{TI}}$$

where Δ' denotes relative density difference, L length scale and U wind speed. Hence

$$\frac{U_{TI}}{U_{ML}}=\left(\frac{V_{TI}}{V_{ML}}\right)^{1/6}\left(\frac{\Delta'_{TI}}{\Delta'_{ML}}\right)^{1/2}$$

Since $\Delta'_{TI}\approx 1$ and $\Delta'_{ML}\approx 3$, this gives $U_{TI}/U_{ML}=11.3$, 8.75 and 7.40 for the three small-scale release volumes. Since U_{ML} was varied from 0.2 to 1.0 m/s, this means that Meroney and Lohmeyer's results cover a very similar Richardson number range to Thorney Island.

The fluctuations in maximum concentration at nominal ground level (i.e. height 2 mm, which scales up to be similar to the height of the "ground-level" sensors at Thorney Island) are plotted against the scaled-up downwind distance for two selected small ranges of windspeed in Fig. 5. The results are widely scattered, but seem to indicate that $I_m=(\overline{C'^2_m})^{1/2}/\bar{C}_m$ reaches a fairly level value of around 0.3 very soon after release, with little dependence on Richardson number: here C_m denotes maximum concentrations in individual realisations at a particular point and the overbar denotes the ensemble mean. However, there does seem to be a significant dependence of I_m on release volume V_{ML}, with larger values at the smallest size of 35 cm^3. This was noted by Meroney and Lohmeyer [22] and Carn and Chatwin [12] and ascribed to differences in Richardson number. Our analysis based on differentiating between different windspeeds indicates that this is not so. Perhaps the smaller

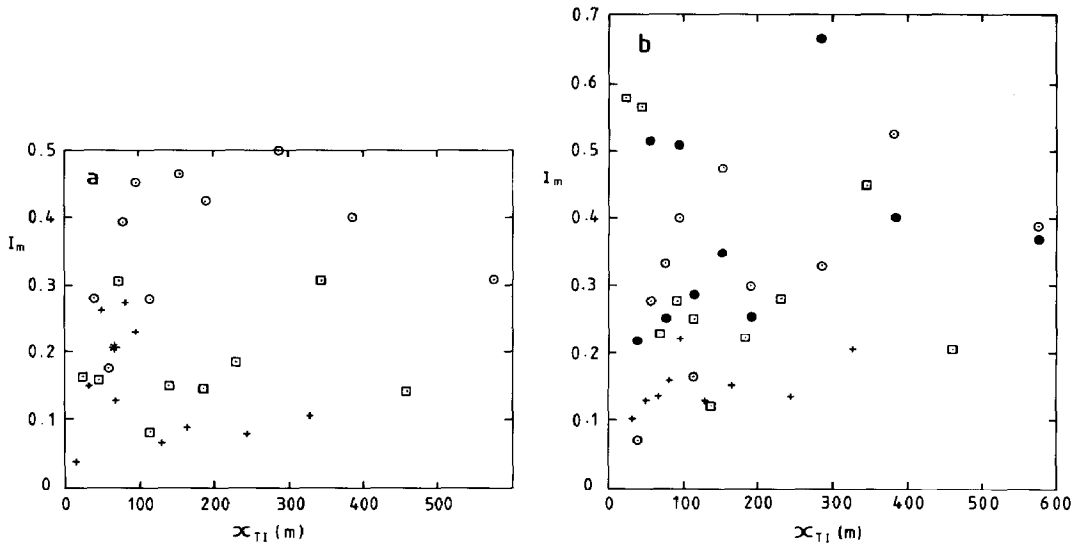

Fig. 5. Ratio of standard deviation of maximum concentration to mean values for replicated releases in Meroney and Lohmeyer's [21] wind-tunnel experiments for selected wind-speed ranges. Wind speeds and downwind distances are scaled to conditions typical of Thorney Island. (a) Wind speeds between 2 and 4 m/s: ⊙, $U_{TI}=2.3$ m/s ($V_{ML}=35$ cm^3); +, $U_{TI}=3.0$ m/s ($V_{ML}=450$ cm^3); ⊡, $U_{TI}=3.5$ m/s ($V_{ML}=165$ cm^3); *, as + but 95 replications. (b) Wind speeds between 6 and 12 m/s: ⊙, $U_{TI}=6.8$ m/s ($V_{ML}=35$ cm^3); +, $U_{TI}=7.4$ m/s ($V_{ML}=450$ cm^3); ⊡, $U_{TI}=8.7$ m/s ($V_{ML}=165$ cm^3); ●, $U_{TI}=11.3$ m/s ($V_{ML}=35$ cm^3).

scale of the cloud relative to the length-scale of ambient turbulence caused greater meandering. Indeed the measurements of I_m at zero windspeed show that there is no significant difference for the three values of V_{ML}.

4.3 Effect of sample size

The large degree of scatter in Fig. 5 illustrates well Carn and Chatwin's [12] remarks on sample sizes needed for estimating means and variances. Since Meroney and Lohmeyer used separate releases to obtain concentration values, each of their data points results from a statistically independent trial. Since C_m is found to obey a near-Gaussian distribution [22,24] Carn and Chatwin's equations (24–28) can be used. We use their notation here. If m and s^2 are the mean and variance of samples of N values of C_m, then $N=5$ and $\alpha=0.05$ gives $t_\alpha=2.13$, $\chi^2{}_{1-\alpha}=0.71$ and $\chi^2{}_\alpha=9.49$, so that the 90% confidence limits on $\bar{C}_m$ are

$(m-0.95\,s,\ m+0.95\,s)$

and on $(\overline{C'^2_m})^{1/2}$ are

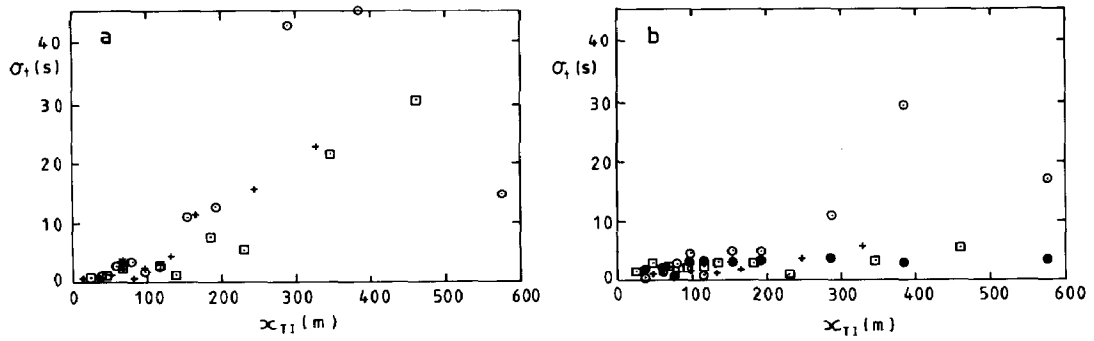

Fig. 6. Standard deviations of arrival times versus downwind dinstance. Results scaled to Thorney Island conditions from Meroney and Lohmeyer's (1982) wind-tunnel experiments. (a), (b): selected wind-speed ranges as in Fig. 5.

$(0.65\ s, 2.37\ s)$

With $I_m \approx 0.3$, the first implies that 5 trials give $\bar{C}_m$ only to within $\pm 40\%$. The confidence limits on the fluctuations seem consistent with the results in Fig. 5, after allowing for the systematic dependence on V_{ML}.

For Meroney and Lohmeyer's data with $N=95$ replications, $\alpha=0.05$ gives $t_\alpha=1.66$, $\chi^2{}_{1-\alpha}=73$ and $\chi^2{}_\alpha=118$ and so the 90% confidence limits for $\bar{C}_m$ are

$(m-0.17\ s, m+0.17\ s)$

and for $(\overline{C'^2_m})^{1/2}$ are

$(0.89\ s, 1.14\ s)$

The experimental result is shown in Fig. 5(a) and does indeed lie in the middle of the general scatter.

These values of $(\overline{C'^2_m})^{1/2}/\bar{C}_m$ are not quite the same as values of $I=(\overline{C'^2})^{1/2}/\bar{C}$ at a specific time and place, because the time that the maximum is reached is itself a random variable. However, I presumably has similar or slightly higher values in the region of the maxima. At times and places near the edge of the cloud I is likely to be considerably larger.

Some of Meroney and Lohmeyer's results on the statistics of arrival times are shown in Fig. 6, again scaled for comparison with Thorney Island conditions. At the lower windspeed there seems to be a sudden increase in the standard deviation σ_t around 100–150 m downwind, a feature seen in the data at other windspeeds, not shown in Fig. 6. The small values of σ_t are perhaps associated with the well-defined near-circular vortex formed in the early stages of the releases, which at Thorney Island was so evident in the aerial photographs. The subsequent behaviour may be associated with break-up of the cloud edge by ambient turbulence.

The variability in position of the front of the cloud can be estimated by multiplying σ_t by the speed of the cloud. Taking this as 2/3 the wind speed (cf. Ref. [9]) gives values of 30–50 m as the standard deviation of the front of the cloud at 400 m downwind.

4.4 Replicated wind-tunnel experiments by Hall, Hollis and Ishaq

Hall, Hollis and Ishaq [23] also conducted a few wind-tunnel experiments with replicated releases, simulating conditions in the Porton trials. They did 21 replications of a measurement at a point equivalent to 37 m downwind from the release point in the Thorney Island experiments at a windspeed of 8.0 m/s. The smallest peak concentration recorded was 4% and the largest 36% with a mean of about 15%. Davies [24] has analysed the statistics of this dataset in some detail and from his results on the concentration records after adjustment to give coincident arrival times, one can deduce that $(\overline{C'^2_m})^{1/2}/\bar{C}_m \approx 0.68$, considerably larger than typical values from Ref. [22] shown here in Fig. 4.1b. Davies also found that in a later time interval, with near-constant mean concentration, I had the value 0.45 ± 0.08. Hall et al. also made seven repeat runs with the same measurement position in calm air and found about a 30% variation between largest and smallest concentration.

4.5 Variability in the instantaneous releases at Thorney Island

There is thus rather conflicting evidence from wind-tunnel trials about levels of variability in the instantaneous releases. Carn and Chatwin [25] have made a noteworthy effort to estimate ensemble means and concentration fluctuations from the Thorney Island data themselves. They took 5 trials at similar windspeeds as an ensemble and then used a kernel-type estimator to derive a multivariate regression for concentration as a function of time and position. So far only limited results are available from comparing individual concentration traces with the estimate of the ensemble mean, but the results in their Figs. 2 and 3 are consistent with values of $I \approx 0.3$ near the time of peak concentration. At points off the path of the centre of the cloud, the passage of the gas may be rather brief and considerably higher ratios of fluctuation to mean are found because of the variability in position of the cloud edge (Fig. 4 of Ref. [25]).

A totally different way of estimating variability in the Thorney Island trials has been used by Davies [24]. His analysis concerns only the statistics of peak concentration as a function of position. For each of ten flat-ground trials, he fitted a quadratic function of position to the distribution of $\log C_m$. He then interpreted deviations of individual values from the fitted function as representing statistical variability. This procedure does not allow for statistical fluctuations that are correlated over a significant proportion of the cloud. The quadratic model is the simplest capable of describing the main features of the behaviour of $\log C_m$ and probably nothing more sophisticated is justified in view

of the sparseness of the data. Nevertheless, the results for the standard deviation of $\log C_m$, which is approximately equal to the intensity, are in reasonable agreement with the estimates discussed above. Values range from 0.45 to 0.79. As in our analysis of the Meroney and Lohmeyer experiments there is no systematic dependence of the results on Richardson number or windspeed.

4.6 Variability in continuous releases

For truly steady-state plumes, the statistics of the concentration field can be derived from time-averaging over a sufficiently long period instead of ensemble-averaging over a sufficient number of replicated releases. In the Thorney Island continuous release trials, the actual release time was limited to about six or seven minutes. Near the source this gave effectively steady-state conditions at the gas sensors for time approaching six minutes but at 250 m downwind the concentration records indicate completely transient conditions [26]. Mercer and Davies [27] have studied the statistics of turbulent velocities, concentrations and correlations between them in order to analyse the effect of density gradients on turbulent mixing. Their results also have a bearing on the question of variability.

At a location 86 m downwind the concentration at height 1 m was not particularly steady and had two main peaks well separated in time at about 3 and 6 min. Mercer and Davies used an averaging time of one minute to calculate turbulence quantities. From the concentration record the typical time between successive peaks and troughs of the concentration record was considerably less than this, so the results should be good estimates of the fine structure of the concentration distribution. The intensity of concentration fluctuations I varied between 0.2 and 0.5 during the whole of the 9-min period of passage of gas. Because of their interest in the properties of the turbulence, Mercer and Davies only examined concentration records from sensors near sonic anemometers, but clearly their analysis could be extended to many more gas sensors in the continuous trials.

The statistics assembled from wind-tunnel simulation of the Thorney Island trials by Davies and Inman [28] provide another means of assessing the magnitude of variability. They have carried out a large number of simulations of different Thorney Islands trials at various scales and measured peak concentrations at individual points. Because of statistical variability both at full scale and in the model individual pairs of readings cannot be expected to coincide. Figure 10 of Ref. [28] gives log–log plots of model against full-scale concentration for the continuous release trials. The points lie around the 45° line of perfect match within fairly large error bands. By counting points on the graphs I estimate that 95% confidence limits for the ratio of model-scale concentration C_{DI} to full-scale concentration C_{TI} are 1/4 and 5. Assuming that $\log_e(C_{DI}/C_{TI})$ is normally distributed this gives the variance σ as about 1/4 $\log_e$ 20. Since C_{DI} and C_{TI} are statistically independent, the variances of

$\log_e C_{DI}/\bar{C}_{DI}$ and $\log_e C_{TI}/\bar{C}_{TI}$ are $\sigma_1 = \theta/\sqrt{2} = 0.53$. Using Meroney and Lohmeyer's [22] relation

$$\sigma_1 = \log_e \left[\overline{C'^2_m}/\bar{C}^2_m + 1\right]$$

for the log-normal distribution leads to the result $(\overline{C'^2_m})^{1/2}/\bar{C}_m = 0.84$, a rather high value compared to other estimates given above.

Davies and Inman also give results from continuous releases performed at Thorney Island for the U.S. Coast Guard and Department of Transportation. In these the gas was released from the container into a rectangular area surrounded by a 2.5 m high fence. Table 5 of Ref. [28] shows that the difference $C_{DI} - C_{TI}$ had a mean near zero but a standard deviation of between 0.57% to 1.14% (depending on the scale of the model) at full-scale peak concentration levels between 1 and 3%. Taking $\bar{C}_m$ as 2% this corresponds to variability in individual measurements $(\overline{C'^2_m})^{1/2}/\bar{C}_m$ between 0.20 and 0.40.

4.7 Implications of variability for model validation

From the above discussion, a reasonable estimate of the intensity of fluctuations in the conditions of the Thorney Island heavy-gas releases is about 0.5 near the centre of the cloud at ground level where the peak values tend to occur. So individual concentration measurements are liable to deviate from the ensemble mean by more than 50% on roughly one third of occasions. Larger variability is likely near the edges of the cloud.

Individual estimates of the intensity of fluctuations I range from 0.2 to around 0.8 depending on the dataset used and the method of analysis. There appears to be no detectable systematic variation with Richardson number or with the mode of release. (In the limit of infinite Richardson number, however, turbulence will be suppressed and there will be no fluctuations.) Indeed this level of fluctuation is very similar to that found in a thorough wind-tunnel study of passive plumes from a ground-level point source in a turbulent boundary layer by Fackrell and Robins [29]. Their Fig. 5 gives $I \approx 0.5$–0.8 for a ground-level source, though elevated sources give considerably higher values.

The implications of this variability for model validation depends on the physical nature of the phenomenon. If it is all caused by small-scale turbulent mixing within the cloud, then concentration fluctuations should be correlated over relatively short distances. Overall cloud properties, such as area, centroid position and higher moments of the spatial ensemble-mean concentration distribution, should then be subject to much less inherent variability. It would then be possible to make a meaningful comparison between predictions of an ensemble-mean model and a single release, because over the cloud as a whole deviations between model and reality should average out. If individual trials do deviate from the mean on larger scales, this may be caused by large-scale internal processes, such as instabilities of the vortex-ring formed in the early stages, or by large-scale variations of the wind-field. The magnitude of varia-

tions due to the second cause could be assessed by using the measured wind-field values in more detail. Instead of a single steady wind speed, box models could easily be adapted to use a time-varying wind speed while for three-dimensional numerical models spatial variations in upwind velocity could also be accomodated. Some variability may be caused by other variations in meteorological conditions such as the apparent variation of roughness with wind direction [15,30].

If the statistical variability does significantly affect the overall cloud properties, then the statistical sampling theory discussed above in Section 4.3 can be used to assess the uncertainty inherent in using the 16 flat-ground instantaneous releases at Thorney Island as a database for parameterising gas-dispersion models. With $N=16$, the 90% confidence limits are given by $t_\alpha=1.75$, $\chi^2_{1-\alpha}=7.26$ and $\chi^2_\alpha=25.0$. Thus the confidence limits on $\bar{C}$ are

$(m-0.44\,s,\ m+0.44\,s)$

and on $(\overline{C'^2})^{1/2}$ are

$(0.77\,s,\ 1.44\,s)$

Here m and s are regarded as estimates derived by fitting some model to the data which is in principle capable of correctly representing the variation of concentration with time, Richardson number etc. With $I=0.5$, the trials should be sufficient to determine $\bar{C}$ to within $\pm 22\%$ with 90% confidence.

4.8 Evidence on the variability of overall cloud properties

Some limited guidance on overall variability is available from published analysis of the Thorney Island data.

Brighton et al. [9] found that the non-dimensional area-increase rate, or frontal Froude number K, determined from the overhead photographs of the early phase of gravity spreading had a mean value of 1.05 with a standard deviation of 0.12. However, most of this variability seems to be caused by uncertainty in the initial conditions of the trials (see Section 3.3 above).

In their Fig. 16.7, McQuaid and Roebuck [2] give a plot against distance of peak concentrations measured at individual ground-level sensors close to the path of the centroid for the Phase I releases. The concentrations at a particular distance vary only by an overall factor of about three, despite the range of release conditions. However, there seems to be a very significant correlation between the individual values for particular trials. For instance values for Trials 14 and 15 are the two highest at virtually all distances. Values for Trials 18 and 19 are amongst the three lowest at all distances. Yet these four trials were conducted at similar windspeeds and Richardson numbers. This is consistent with the hypothesis that overall cloud concentrations have a variability of much the same size as that found for spot values.

Results of the optimisation of simple models can also provide estimates of

TABLE 3

Root mean-square relative concentration deviations between the area-averaged data [10,13] and the generalised Picknett model with $\alpha_E = 0.7$ and $\mu = 1$. α_T is optimised individually for each trial [31,13]

Trial	Maximum no. of masts in cloud simultaneously	$S_4^{1/2} = [\overline{(\Delta C)^2/C^2}]^{1/2}$
7	9	0.175
8	11	0.155
9	22	0.177
11	7	0.261
12	23	0.128
13	7	0.216
14	9	0.284
15	5	0.373
16	7	0.448
17	9	0.121
18	8	0.253
19	10	0.382
26	32	0.518
28	17	0.425
29	16	0.299
34	26	0.279

the variability of overall properties. Wheatley et al. [31] fitted a generalised Picknett model to the area-averaged ground-level concentration data [10] for the flat-ground trials. The best results using various goodness-of-fit measures (GFMs) were obtained by fixing the edge entrainment coefficient α_E as 0.7, the Richardson number exponent μ as 1 and the velocity scale as the cloud speed, and then varying the top entrainment coefficient α_T. The GFM denoted S_4 is a mean value of the squared relative deviations of the predicted concentrations from the area-averaged values:

$$S_4 = \frac{\sum_i n_i \, (\bar{C}_i - C_i^{\lambda})/C_i^{\lambda 2}}{\sum_i n_i}$$

Here i denotes successive values of time, $\bar{C}$ is the measured area-average concentration, and C^{λ} the predicted concentration; n_i is the number of measuring points in the cloud at each time interval. Thus $S_4^{1/2}$ is a measure of the intensity of fluctuations of $\bar{C}$ — the fitted model is probably sufficiently flexible to make systematic deviations negligible. The optimised values of $S_4^{1/2}$ for various trials (not previously published) are listed in Table 3. Interpretation is complicated by the fact that estimates of $\bar{C}$ are subject to sampling errors because of the

TABLE 4

Variability in distances to given peak concentration levels on the path of the cloud centroid in the Phase I Thorney Island trials, from Fig. 16.7 of Ref. [2]. It is *assumed* that the relative rms deviation of concentration is 0.5

Concentration	Mean distance from source, x	Estimated standard deviation σ_x/x
5 %	147	0.316
2.5%	247	0.269
1 %	420	0.194

limited number of masts employed. Local values of concentration are subject both to statistical variability as discussed above and to systematic variation with position in the cloud. The most reliable results come from the low-Richardson number trials when large numbers of masts were in the cloud simultaneously, indicating that fluctuations in the true area-averaged concentration may not exceed a relative value of around 0.15. However, this procedure does exclude any fluctuations in mean concentration which last throughout the measuring period as illustrated by the concentration distance plot in Ref. [2].

The results in Table 3 are also of interest in indicating the maximum precision with which the model can be made to agree with this data. Because the top-entrainment coefficient α_T has been optimised individually for each trial, there is no deviation because of error in determining the mean meteorological conditions. Because a single value of α_T was then chosen to give the overall model [31], root-mean square deviations of concentration for this overall prediction will be larger than the values in Table 3. A visual comparison is available in Ref. [31], but a quantitative comparison has not been done.

Fitpatrick and Nussey [32] have also optimised a model and given numerical values of the residual deviation. They used a revised version of the DENZ box-model and varied various sets of parameters to fit the peak concentration data from seven of the Phase I trials. Depending on the set of parameters varied, the relative root mean-square concentration deviation was between 0.28 and 0.62; by comparison with the results of Table 3 it would appear that only a minor proportion of this variance is caused by the failure of the model's top-hat concentration to match the actual spatial concentration distribution.

Finally, we consider variability in measured and predicted hazard ranges. To estimate this from the data, we have taken a visual mean curve from the plot in Fig. 16.7 of Ref. [2] of peak concentration against distance. Then we have drawn two curves corresponding to concentrations greater or less by a factor of 1.5, corresponding to the typical level of variability of peak concentration values. As discussed earlier in this section, the data suggest that fluctuations

of this size may persist throughout the measurement period. The two curves representing the standard deviation concentration can be read off at a fixed concentration value to give an estimate of the standard deviation of the distance to that level of peak concentration: this distance is frequently used as the main indicator of the extent of the hazard from gas dispersion. The results are listed in Table 4. The relative standard deviation in hazard distance estimated in this way has much the same value as the standard deviation of the ratio of observed to predicted hazard distance from the DEGADIS model given by Spicer and Havens [33] for six of the flat-ground Thorney Island releases. Similar results were obtained by Puttock [8]. The magnitude of this uncertainty in hazard distance is also similar to the variability in the position of the leading edge of the cloud deduced from Meroney and Lohmeyer's statistics of arrival time in Section 4.3. above.

5. Use of the data for validating box-models

5.1 The influence of the release mechanism

The objective in designing the release mechanism was to form a heavy-gas cloud instantaneously as a well-defined initial condition. This has been very beneficial in allowing analysis of the heavy-gas dispersion without the uncertainties inherent in real release conditions. These were used at China Lake [63] and Maplin Sands [64] and tend to introduce considerable uncertainty in defining release rates and additional physical effects such as heat transfer.

The system of a tent-like container with retractable lid and rapidly drawn-down sides gave a visually very clean release — it operated rapidly enough to leave a well-defined standing column of gas for a few moments before gravity slumping got under way. However, the release method has introduced artificialities of its own which have provoked quite a lot of analysis. Since the gas is not released "isokinetically", i.e. at the same speed as the air, it has to be accelerated by pressure forces. Rottman et al. [34] have shown how in the absence of gravity slumping this leads to quite a complex horse-shoe shape.

The vertical slumping, in the absence of ambient wind, or when its effect is small, causes the doughnut-shaped vortex which has become the hall-mark of the Thorney Island trials.

These phenomena, which can so easily be seen in the excellent photographs of many of the trials, have perhaps attracted excessive attention. Numerous elaborate theoretical and experimental studies have been made on the mechanics of wave formation, vortex stretching and turbulent kinetic energy generation in this phase [35–42]. The overhead photos also indicate that the vortex has greatly diminished in intensity by the time the cloud radius has reached 50 m. In Phase I the nearest concentration sensors were only at this distance. Since the cloud often subsequently becomes several hundred metres across, these phenomena really occur only during the very early stages of each release.

The over-emphasis on use of the photographic data has led modellers to apply uncritically results from this special early phase of dispersion [9] to cloud evolution on the much longer time-scale over which concentration data were gathered. Conversely, physical phenomena characteristic of the later phase of dispersion have been almost completely neglected. I refer to the tendency of clouds to become elongated in the wind direction [10] and the way in which the sharp gravity-current front evolves to give a smoother concentration profile at the cloud edges. I known of only two studies relevant to the physics involved in these processes [43,44].

In practice, in order to predict concentrations downwind with a box-model, it has been found that the peculiarities of the doughnut-vortex flow can be ignored. A simple edge-entrainment hypothesis with the standard gravity-spreading law gives very good results, after possibly including a time delay to account for the initial conditions [8,31,42,46,47].

5.2 The spatial resolution of the sensor array

Quantities such as the area of the cloud and its mean speed of advection are needed for direct comparison with box models, which use them as fundamental variables. They are also of possible relevance for testing any kind of heavy-gas dispersion model, as they may be more stable as estimates of ensemble mean properties than individual concentration measurements, as discussed in Section 4 above.

In the early stages of cloud motion covered by the overhead photography, these quantities can be determined with a fairly high degree of precision [9]. However, the cloud outline remains visible only up to about 50 m from the source, while the sensor array extends about ten times as far. Away from the source, the masts were located on a square grid of spacing 100 m and so determination of cloud geometry must involve a strong element of interpolation. The accuracy which can be achieved is limited both by the spatial resolution of the array and by the fact that individual measurements are subject to statistical variability, a significant part of which is probably independent from mast to mast.

Different investigators have used different methods of determining cloud outlines, and so the degree of uncertainty resulting from the limited resolution of the sensor array can be assessed by comparing the results. Gotaas [48] and Riou [49] have drawn contours or cloud outlines at specific times by manual analysis of the concentration data. Hartwig [50] and Pfenning and Cornwell [51,52] have produced contours at specified times using computerised interpolation methods. Brighton et al. [9] used a rather different approach of fitting a geometrical model of the cloud outline to the arrival and departure time data, which were determined by visual inspection of concentration records. This method also has the advantage of easily generating the geometrical estimates at arbitrary times. Therefore it has been used as a common standard by

TABLE 5

The uncertainty in determination of the cloud outline from concentration measurements

Author and method	Sample size	Time range (s)	Statistics of differences from results of Brighton et al. [9]			
			Width (m)		Length (m)	
			Mean	Standard deviation	Mean	Standard deviation
Gotaas [48]	10	40–400	−60	44	+17	46
Instantaneous 0.1% contours — manual						
Riou [49]	16	15– 30	− 6	19	+ 6	17
Interpolation of arrival and	22	60– 90	−33	19	+13	42
departure times — manual	16[a]	120–240	−34	39	+38	73
Hartwig [50]	17	50–200	− 6[b]	34[b]	—	—
Computer — instantaneous 0.1% contours						
Pfenning and Cornwell [51,52]	4	30– 90	+ 6	14	+21	38
Computer — instantaneous 1% contours						

[a]Sample size for lengths is 15.
[b]Comparison made of areas A — results quoted for difference in equivalent diameter, $(4A/\pi)^{1/2}$.

which to gauge the differences of interpretation of other authors, who have all chosen different times at which to determine cloud outlines. This is not to imply that the method of Ref. [9] is any better than the others and the aim of the comparison is not to identify any particular method as superior. The results are assumed to reflect the intrinsic uncertainties in the data.

Table 5 summarises the comparison of cloud dimensions. In most cases, the data has been obtained by measuring length and width from the printed contours — they are all usually fairly elliptical in form. Most of Hartwig's [50]

TABLE 6

Ratios of cloud speeds determined by different authors for individual Phase I trials. Results averaged over the set of trials analysed by each pair of authors. (B = Brighton et al. [9]; G = Gotaas [48]; R = Riou [49])

	Ratio R/G	Ratio B/R	Ratio G/B
Number of trials analysed by both authors	10	12	11
Mean ratio	1.09	1.12	0.84
Standard deviation	0.36	0.34	0.15

data is in the form of graphs of area against time and the comparison has been made by calculating an equivalent diameter. (Part of the data in Ref. [50] for Trial 11 has been omitted since it shows area decreasing to zero and is clearly not representative of the real area.) The time range over which contours were obtained is given because errors tend to become greater away from the source where the mast spacing is greater, the cloud edges more diffuse, and concentrations generally lower. The table gives the statistics of differences in the dimensions, with the result of Brighton et al. [9] always being subtracted from the other. Thus on average Gotaas obtains a cloud width 60 m less than we do, though this difference has a standard deviation of 44 m. This is the largest bias found, and in many other cases the alternative methods do not differ significantly on average. The implications of the results may be gauged as follows. Suppose that in determining where the cloud outline intersects the 100 m line between two masts, that two investigators each made independent uniformly distributed *random* choices. The difference in results would then have a symmetric triangular distribution between -100 m and $+100$ m: the mean of this is 0 m and the standard deviation is easily evaluated as $100/\sqrt{6}=41$ m. Comparing this to the results in Table 5, I am inclined to infer that the data allow only a determination of which masts the cloud outline passes between, and that little more can be said. The uncertainty in cloud length tends to be somewhat greater because it is sensitive to different evaluations of departure time. Brighton et al. [9] and Riou [49] used a judgement based on the character of the concentration record, rather than a fixed concentration level (cf. comments in Section 3.1 about gas sensor performance in this respect).

In Refs. [9], [48] and [49], estimates of cloud speed are given based on the movement of the centroid of the cloud outline over the period covered by the analysis. Again the limited spatial resolution of the sensor array gives scope for differing interpretation, so Table 6 gives the statistics of the ratios of cloud speeds determined by the three methods in individual Phase I trials. Again the bias in the different methods relative to one another is overwhelmed by the random scatter, except for a somewhat better degree of correlation between the results in Refs. [9] and [48].

An additional perspective on the importance of the spatial resolution of the sensor array is afforded by the calculations of mass balance in Ref. [10]. For most of the Phase I trials, there is a significant period when the estimate of the amount of gas detected remains well within 50% of the amount released. These calculations involve estimation of the vertical structure of the cloud as well as its horizontal extent and in Ref. [10] the various potential sources of error are discussed. For the continuous release trials, estimation of the mass-flux balance is presented in Ref. [53].

5.3 The determination of entrainment rates

The great majority of box-models of heavy-gas dispersion calculate the concentration in the cloud by means of an entrainment equation of the form

$$\frac{dV}{dt}=2\pi\alpha_E RH\frac{dR}{dt}+\pi R^2\frac{\alpha_T u'}{Ri^{\mu}} \tag{5.1}$$

or a very similar one [1]. Here t is time, V is cloud volume, R and H the radius and height of the cloud; u' is a velocity scale determining the rate of top entrainment and Ri is the associated Richardson number

$$Ri=g\Delta' L/u'^2 \tag{5.2}$$

where g is the acceleration due to gravity, Δ' the relative density difference and L a lengthscale, usually H. α_E and α_T are the edge- and top-entrainment coefficients and μ is a further empirical parameter, the Richardson-number exponent.

The analysis of the Thorney Island data required to validate box-models has mainly involved determination of values of α_E and α_T which optimise the fit according to various criteria; μ is generally taken as unity on the basis of results from laboratory experiments [1] but in fact a range of values seems to fit the data reasonably well [31].

The edge-entrainment coefficient determines cloud concentrations in the early stages of each release and top entrainment becomes dominant later [10]. Authors appear to agree that this number is well determined by the Thorney Island data and have recommended similar values: 0.65 [47], 0.7 [31,46] or 0.85 [8]. On the other hand, in optimising the parameters of the DENZ model under various sets of constraints, Fitzpatrick and Nussey [32] obtain various results for α_E between 0.46 and 0.7. However, they were using only the peak concentration data as a function of position for a selection of seven trials. Other authors have generally also included the time variation of concentration and arrival and departure time information in their considerations.

Turning to the top-entrainment coefficient, we cannot directly compare values from different authors because definitions of u' and L in eqns. (5.1) and (5.2) vary. However, several authors report that this parameter is somewhat ill-determined. Carpenter et al. [47] found that there was a range of values of α_T giving near-minimal values of an overall objective function applied to the whole dataset from Trials 7 to 19. Crabol et al. [46] fitted the parameters in their model individually for each trial and obtained results for α_T ranging over more than an order of magnitude (see their Fig. 20). The values show no dependence on initial Richardson number and so the variation does not seem due to any systematic deficiency in the model. Wheatley et al. [31] obtained a very similar range of values of α_T across Trials 7–19 and they also show very shallow ill-defined minima for certain trials (see their Fig. 21). There is no correlation between the two sets of values of α_T obtained in Refs. [31] and [46]. Fitzpatrick and Nussey [32] obtained values of α_T ranging over more than a factor of ten from their global fit to peak concentration values for a set of seven trials.

The origins of this uncertainty can be seen in logarithmic plots of concentration against time. At relatively small times, these tend to follow a line of slope $-\alpha_E$ and at larger times one of slope $-(\mu+2)$ (see Fig. 2 of Ref. [10]). This feature is expected from the solution of eqn. (5.1) with H as the length scale L:

$$\frac{V}{V_0}=\frac{C_0}{C}=(1-\gamma)\ \tau^{\alpha_E}+\gamma\ \tau^{2+\mu} \tag{5.3}$$

where C represents concentration and subscript zero denotes initial values. τ is a dimensionless time equal to $(R/R_0)^2$ because of the gravity-spreading relation. γ is given by

$$\gamma=\frac{\alpha_T}{2K\ (2+\mu-\alpha_E)\ Ri_0^{\mu+1/2}}\frac{R_0}{H_0} \tag{5.4}$$

where K is the frontal Froude number [9].

Inspection of Fig. 2 of Ref. [10] shows that the concentration values over which the second term of eqn. (5.3) becomes dominant in practice are at most 0.3% in high-Richardson number trials. At these concentration levels, the data is subject to potentially large relative errors, particularly from residual nonlinear drift, as discussed in Section 3. Also for several of these trials, a significant part of the cloud was outside the sensor array by the time top entrainment became important. In lower Richardson number trials, these problems are less severe and the top entrainment is both more determinate for individual trials, and more consistent between trials according to Wheatley et al. [31].

From eqns. (5.3) and (5.4), the concentration predicted in the top entrainment phase is inversely proportional to the top-entrainment coefficient, so uncertainty in α_T may affect concentration predictions in direct proportion. To put this uncertainty into perspective, I have made a limited comparison of the entrainment rates derived from the Thorney Island experiments with relevant experimental data. I have also considered how far the Richardson numbers encountered in full-scale accident studies may exceed the maximum achieved in the Thorney Island experiments. Clearly uncertainties will be magnified if a substantial extrapolation is required, particularly if the value of the exponent μ in eqn. (5.1) is in doubt.

Comparison with laboratory entrainment relations

Britter [54] has argued that the range of Richardson number for which the top entrainment rate is needed for practical heavy-gas dispersion calculations is far more limited than often assumed. Many of the laboratory experiments used to parameterise box-models in the past were at inappropriate values of Ri and moreover involved different mechanisms of turbulent mixing. The most relevant experiments are emission of dense fluid from a line source in a wind-

tunnel into an ambient turbulent flow. This involves the same mechanism of turbulence generation, by shear at a lower boundary, as in a gas cloud, and also has the very convenient feature that the Richardson number remains constant downstream allowing determination of the entrainment rate for fixed Ri by measuring the rate of deepening of the layer with distance.

A set of experiments of this type by McQuaid [55] was analysed by Britter [56] to give the entrainment relation

$$w_e/U = 2\times10^{-4}\ Ri_B^{-1} \text{ for } 0.02 < Ri_B < 0.15 \tag{5.5}$$

where w_e is the entrainment velocity, U the mean velocity in the tunnel and $Ri_B = g\Delta'_0 q_0/U^3$ with q_0 the volume release rate per unit width. Below the lower end of the range the dispersion soon becomes passive for $Ri_B < 0.01$ with $w_e/U \approx 0.018$. The upper end of the range of Ri_B is the maximum reached in the experiments.

This result can be directly compared with the entrainment relation

$$w_e/u' = 0.14\ Ri^{-1} \tag{5.6}$$

extracted from the Thorney Island experiments by Wheatley et al. [31], since their Richardson number is also based on the vertical integral of concentration, $Ri = g\Delta'_0 \int c\,dz/u'^2$. The velocity scale u' is $0.16U_c$ so taking Ri_B as the Richardson number based on the cloud speed U_c, eqn. (5.6) may be rewritten as

$$w_e/U_c = 5.7\times10^{-4}\ Ri_B^{-1} \tag{5.7}$$

The numerical coefficient here scales as U^3 and if we take $U_c = 0.7\ U$, eqn. (5.7) becomes identical to eqn. (5.5). This seems quite reasonable since in the wind-tunnel experiments U was the mean speed in the whole wind-tunnel, not just in the dense-fluid layer.

This consistency is also confirmed by the agreement of the DEGADIS model with the Thorney Island experiments [57], for the entrainment relation deduced by Britter was directly incorporated into the formulation of this model.

Range of Ri *in the top-entrainment relation*

The simple eqn. (5.3) for the concentration has been shown to give a very good description of the Thorney Island area-averaged data [31] and can therefore be used to determine the relative importance of top and edge entrainment in the experiments. Let E denote the value of the first term on the right of eqn. (5.3) and T the value of the second. The time at which top entrainment starts being significant is defined by $T/E = 0.1$. Since the Richardson number evolves according to

$$Ri = Ri_0/\tau \tag{5.8}$$

(assuming the standard gravity-spreading law and dropping the suffix B), this

TABLE 7

Values of Richardson number and concentration at various phases in cloud development. T is the volume increase of the cloud due to top entrainment and E that due to edge entrainment

		$T/E=0.1$	$T/E=1$	Cloud becomes passive
Trial 12				
$Ri_0=422$	Ri	0.35	0.13	0.01
$H_0/R_0=1.86$	C/C_0	6.7×10^{-3}	1.7×10^{-3}	1.8×10^{-6}
Release volume 1000 times larger	Ri	0.78	0.29	0.01
$Ri_0=4220$				
$H_0/R_0=1.86$	C/C_0	2.3×10^{-3}	5.9×10^{-4}	5.9×10^{-8}
The same, with low aspect ratio	Ri	1.66	0.60	0.01
$Ri_0=325$				
$H_0/R_0=0.04$	C/C_0	2.4×10^{-2}	6.1×10^{-3}	5.9×10^{-8}

defines a Richardson number $Ri_{0.1}$ for top entrainment becoming important. Since $\gamma \ll 1$, the factor $1-\gamma$ in E may be ignored in order to derive the result

$$Ri_{0.1}=0.055\ Ri_0^{0.35}\ (H_0/R_0)^{-0.43} \tag{5.9}$$

by taking $\mu=1$, $\alpha_E=0.7$ and $\alpha_T=5.7\times10^{-4}$. The concentration can also readily be evaluated at this time.

Results of these calculations are shown in Table 7. This also gives values of the Richardson number when top entrainment has made the same contribution to cloud dilution as edge entrainment, i.e. $T=E$. From Britter's [56] entrainment relation the dispersion becomes effectively passive when $Ri=0.01$ and Table 7 also shows at what concentration this stage is reached.

Three release conditions are considered in Table 7. Trial 12 had the highest Richardson number in the Thorney Island trials and thus displayed top entrainment at the highest value of Ri. This value happens to be similar to the maximum reached in McQuaid's [55] wind-tunnel experiments, and so the agreement in entrainment relations discussed above is indeed based on similar ranges of Ri.

A release 1000 times larger in initial volume with the same aspect ratio has an initial Richardson number merely 10 times greater. Furthermore, eqn. (5.9) shows that the maximum Richardson number at which the entrainment relation needs to be known increases even more slowly with the scale of release as shown in Table 7. For this size of release the initial aspect ratio is much more

likely to be small as a result of pool boil-off and then the large area of the cloud means that top entrainment becomes important sooner. Nevertheless, even in these conditions the top-entrainment relation needs to be extrapolated only an order of magnitude in Ri beyond the experimental database. In most practical problems the extrapolation will be rather less.

5.4 The cloud speed

Analysis of the Thorney Island results [31,46] has shown conclusively that advection of the cloud by the wind cannot be described by taking the speed U_c equal to the wind speed U_w at some fixed fraction of the cloud height, as was assumed in many box-models. Several authors [8,30,45–47] have proposed to describe the translational motion of the cloud by means of an overall momentum equation. This is of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho V U_c) = M - G + A \tag{5.10}$$

where M is the rate of inflow of momentum in entrained air, usually multiplied by a loss factor f, G is the drag from the ground and A is the acceleration due to air flow past the cloud. There are considerable differences in which of these terms to include and how to parameterise them.

Carpenter et al. [47] use all three terms. In M a loss factor of 0.55 is applied to air entrained through the edge while $f=1$ for top entrainment. G is given by the standard friction factor for non-stratified turbulent flow and A is represented by the pressure force estimated to act on a solid obstacle of the same size as the cloud.

Crabol et al. [46] and Deaves [30] use only the term M with f taking the values 0.58 and 1, respectively. Puttock [8] recommends the first term alone with $f=0.7$ in the early stages of dispersion, but at later times the terms G and A are taken into account implicitly by making f dependent on the Richardson number so that it becomes 1 at low concentrations. Wheatly et al. [31] use all three terms, but take M proportional to $1/2\ (fU_w + U_c)$ with $f=0.8$. A is assumed to be given by a shear stress at the cloud top dependent on $U_w - U_c$ and on the Richardson number. As in Ref. [47], the drag term G is given by a neutral-flow relation.

This diversity of fews stems partly from the fact that the full form of eqn. (5.10) contains a large number of potentially adjustable empirical constants, typically two for each of the three terms. Each author has made *a priori* assumptions about some of these constants and then fitted the rest to the data by various means. Fitzpatrick and Nussey [32] show that the constants in the model of Wheatley et al. [31] may assume a wide range of values if different optimisation criteria are used.

The other source of variation in interpretation is the rather large uncertainty in cloud speed revealed in Section 5.2. The graphs in Ref. [31] display

a large degree of scatter between trials. Perhaps one way in which the results from the Thorney Island trials could have been substantially improved would have been the deployment of velocity sensors at the location of every concentration sensor, so that mean cloud speed could be estimated in the same way as mean cloud concentration.

6. Physical understanding of heavy-gas dispersion

In discussing the magnitudes of errors and uncertainties in Sections 3–5 above, I have been addressing the question of whether the first of the objectives of the Thorney Island trials has been achieved: that is, whether that data are reliable and can be used for validating models. The second objective was more nebulous (see Section 1.3) and it is harder to provide quantitative evidence as to whether physical understanding has been improved.

During the planning of the trials, this quest for physical understanding seems mainly to have been directed at questions of the influence of density gradients on turbulence structure, which has to be parameterised in order to close the set of partial differential equations for turbulent flow. However, I have argued above in Section 2 that a greater degree of physical understanding is needed in order to develop simple integral or box models.

From this view-point, the flat-ground trials seem largely to have confirmed pre-existing physical understanding of the importance of gravity spreading, and edge and top entrainment. All the models presented at the Symposium have embodied much the same formulation of these processes, and the main questions have been about the best ways to optimise the fit of models to the data. There has, on the other hand, been more diversity of opinion on the mechanisms of cloud advection, and here the trials do seem to have stimulated some new thinking (see Section 5.4).

In the Phase II trials, the presence of obstacles has led to several insights into novel fluid-dynamic processes [2,13,58,59]. There is the splashing of a gravity current meeting a solid obstacle and the reflection of a wave through the gas cloud. There is the persistence of gas in a blocked region upwind of a solid barrier and the mechanics of gravity propagation through a permeable barrier. The cubical obstacle has given indications of a curious effect of a horseshoe vortex in lowering concentrations on the side facing the gas source.

As regards the type of physical understanding originally sought, success has been very limited. Nussey et al. [5] sought correlations between fluctuations in vertical velocity and in concentration, using one-minute averages in six Phase I trials. They could not find any significant differences between turbulence intensities in the cloud and in the ambient flow. Furthermore, because the averaging times required were similar to the time of passage of the gas, they felt that "any attempt to compute fluxes from the data would need careful consideration and interpretation".

Nevertheless, Hartwig and co-workers [50,60,61] have succeeded in determining vertical concentration gradients and concentration fluxes using 20-s time averages. The results show a clear lowering of the eddy diffusivity at high Richardson number. However, the scope of the investigation was limited by the rarity of occasions on which the high-frequency instruments on the "mobile" masts successfully gathered data well within the gas cloud.

In the continuous trials, difficulties caused by the non-stationarity of the flow are less severe, and Mercer and Davies [27] have demonstrated significant effects of the presence of gas on turbulence intensity. There is also an interesting reduction in the region above the top of the gas cloud, presumably reflecting a reduction of shear stress at the ground. However. the results again seem to be based on too small a sample to give much definite information about the dependence of turbulent mixing on the Richardson number.

These difficulties should not be serious for the development of turbulence modelling schemes applicable to heavy-gas dispersion. The question of parameterising turbulent fluxes in stratified shear flows arises in meteorology and oceanography and in other technological contexts such as mine ventilation, fires in enclosed spaces and sodium flows in fast nuclear reactors. Any complex numerical scheme should be fairly universal and therefore can be developed on the basis of information from all these fields, as well as from fundamental fluid physics research in the laboratory. Indeed, since the main justification of developing complex heavy-gas models is to apply them to dispersion over variable terrain and past obstacles, even the fullest "physical understanding" based purely on flat-ground experiments would be insufficient.

7. The achievements of the Thorney Island Heavy Gas Dispersion Trials

In this paper I have sought to review all the evidence bearing on the level of uncertainty in the data from the Thorney Island trials. The task of drawing conclusions is left to the reader. The uncertainties are of many different types and their importance will depend on the end use.

In seeking to expose sources of uncertainty, I have taken for granted the many technical achievements of the trials. A vast amount of data has been collected, from 16 instantaneous heavy-gas releases over flat ground, and 10 with obstacles, from 3 continuous releases and from at least 6 trials with a dyke surrounding the source, performed for the U.S. Department of Transportation. Up to 100 concentration records were gathered in each trial. It is particularly noteworthy that the sensor array was designed with the ideal size. With the given resolution of about 0.1% for the standard gas sensors, concentrations at this level were just reached at the edge of the array of masts. A larger array would have been wasteful, a smaller one would have lost valuable data. It is difficult to see how the experiments could have been improved without extra money or new technology. The only possible change in experimental design

that I might have liked with hindsight is the deployment of velocity sensors in association with the gas sensors, perhaps at the expense of less coverage of the gas field at greater heights.

The trials have certainly decreased the uncertainty associated with making predictions of heavy-gas dispersion. McQuaid and Roebuck [2] describe how before the trials predictions of the results were obtained from eleven organisations. Comparison of the results revealed a scatter of concentration values spanning more than an order of magnitude at any given distance. Despite the statistical variability and the uncertainties in fitting box-models to the data, it is fairly safe to assume that a new comparison of models explicitly tuned to the Thorney Island dataset, or at least validated against it, would display considerably less scatter.

Extrapolation to full-scale accident conditions is likely to give increased divergences between models, for reasons such as those discussed in Sections 5.3 and 5.4. The work of Fitzpatrick and Nussey [32] and Sherrell and Chatwin [62] provides a further warning that widely differing parameter sets, and indeed physically different models, may give equally good fits to the data. A good way of quantifying the uncertainty remaining in dispersion predictions in the wake of Thorney Island will be a new code-comparison or benchmark exercise, involving as many as possible of the models validated by the trials. Cases to be compared should include conditions representative of the trials themselves, to identify differences arising from different methods of analysing the data. Richardson numbers, different initial aspect ratios, and increased surface roughness (relative to cloud height) are amongst the most important variations relevant to comparing how the models will extrapolate.

Perhaps the most significant achievement of the Thorney Island trials has been the way they were organised. The benefits of forming a consortium of almost forty organisations are not only financial, in making possible a programme of this scope, but also scientific in bringing together many different minds in both planning the experiments and analysing the results. Other field trials have generally been planned, executed and analysed by a single organisation. Only the Thorney Island trials have provoked enough scientific activity to justify the mounting of two full-scale Symposia. In particular, this has made possible the intercomparison of different methods of analysis which I have attempted in the latter part of this paper.

This second Symposium has been noteworthy for a much greater emphasis than the first on the fundamental statistical variability of gas dispersion. This has very important implications for the planning of future experiments, the extent to which mathematical models can or should be refined, and the practical use made of the predictions of mathematical and physical models.

Acknowledgements

My special thanks go to Dr. Jim McQuaid for asking me to prepare a paper with this title and encouraging me along the way. At first this seemed an impos-

sible task: in the event, I have found the exercise a healthy corrective to overconfidence in my own interpretation of the dataset. I also thank the many authors of papers at the Symposium who supplied me with preprints, unpublished reports and written and verbal comments. The work was funded by the Health and Safety Executive.

References

1 C.J. Wheatley and D.M. Webber, Aspects of the dispersion of denser-than-air vapours relevant to gas cloud explosions. Report EUR9592en, Commission of the European Communities, Brussels, 1985.
2 J. McQuaid and B. Roebuck, Large scale field trials on dense vapour dispersion, Report EUR10029en, Commission of the European Communities, Brussels, 1985.
3 N.C. Harris, A user's view on the data obtained and those which may still be required, J. Hazardous Materials, 11 (1985) 425-434.
4 M.J. Leck and D.J. Lowe, Development and performance of the gas sensor system, J. Hazardous Materials, 11 (1985) 65-89.
5 C. Nussey, J.K.W. Davies and A. Mercer, The effect of averaging time on the statistical properties of sensor records, J. Hazardous Materials, 11 (1985) 125-153.
6 A. Mercer, Methods of validation of heavy gas dispersion models: a review, In: J.S. Puttock (Ed.), Proc. IMA Conference on Stably Stratified Flow and Dense Gas Dispersion, Chester, U.K., April 9-10, 1986, Oxford University Press, in press.
7 P.C. Chatwin, Towards a box model of all stages of heavy gas cloud dispersion, In: J.C.R. Hunt (Ed.), Turbulence and Diffusion in Stable Environments, Oxford University Press, 1985, pp. 259-291.
8 J.S. Puttock, Comparison of Thorney Island data with predictions of HEGABOX/HEGADAS, J. Hazardous Materials, 16 (1987) 439-455.
9 P.W.M. Brighton, A.J. Prince and D.M. Webber, Determination of cloud area and path from visual and concentration records, J. Hazardous Materials, 11 (1985) 155-178.
10 P.W.M. Brighton, Area-averaged concentrations, height-scales and mass balances, J. Hazardous Materials, 11 (1985) 189-208.
11 P.W.M. Brighton, Using concentration data to track clouds in the Thorney Island experiments, Report SRD R319, UKAEA, Culcheth, U.K., 1986.
12 K.K. Carn and P.C. Chatwin, Variability and heavy gas dispersion, J. Hazardous Materials, 11 (1985) 281-300.
13 P.W.M. Brighton and A.J. Prince, Overall properties of the heavy gas clouds in the Thorney Island Phase II trials, J. Hazardous Materials, 16 (1987) 103-138.
14 J.S. Puttock and G.W. Colenbrander, Thorney Island data and dispersion modelling, J. Hazardous Materials, 11 (1985) 381-397.
15 J.S. Puttock, Analysis of meteorological data for the Thorney Island Phase I trials, J. Hazardous Materials, 16 (1987) 43-74.
16 M.E. Davies and S. Singh, Thorney Island: its geography and meteorology, J. Hazardous Materials, 11 (1985) 91-124.
17 D.G. Beesley, Unpublished memoranda, UKAEA, Culcheth, U.K.
18 P.C. Chatwin, The use of statistics in describing and predicting the effects of dispersing gas clouds, J. Hazardous Materials, 6 (1982) 213-230.
19 D.J. Ride, An assessment of the effects of fluctuations on the severity of poisoning by toxic vapours, J. Hazardous Materials, 9 (1984) 235-240.

20 R.F. Griffiths and A.S. Harper, A speculation on the importance of concentration fluctuations in the estimation of toxic response to irritant gases, J. Hazardous Materials, 11 (1985) 369–372.
21 R.N. Meroney and A. Lohmeyer, Gravity spreading and dispersion of dense gas clouds released suddenly into a turbulent boundary layer, Report GRI-81/0025, Gas Research Institute, Chicago, 1982.
22 R.N. Meroney and A. Lohmeyer, Statistical characteristics of instantaneous dense gas clouds released in an atmospheric boundary-layer wind tunnel, Boundary-Layer Meteorol., 28 (1984) 1–22.
23 D.J. Hall, E.J. Hollis and H. Ishaq, A wind tunnel model of the Porton dense gas spill field trials, Report LR394(AP), Warren Spring Laboratory, Stevenage, U.K., 1982.
24 J.K.W. Davies, A comparison between the variability exhibited in small scale experiments and in the Thorney Island Phase I trials, J. Hazardous Materials, 16 (1987) 339–356.
25 K.K. Carn, Estimates of the mean concentration and variance for the Thorney Island Phase I dense gas dispersion experiments, J. Hazardous Materials, 16 (1987) 75–101.
26 J. McQuaid, Design of the Thorney Island continuous release trials, J. Hazardous Materials, 16 (1987) 1–8.
27 A. Mercer and J.K.W. Davies, An analysis of the turbulence records from the Thorney Island continuous release trials, J. Hazardous Materials, 16 (1987) 21–42.
28 M.E. Davies and P.N. Inman, A statistical examination of wind tunnel modelling of the Thorney Island trials, J. Hazardous Materials, 16 (1987) 149–172.
29 J.E. Fackrell and A.G. Robins, Concentration fluctuations and fluxes in plumes from point sources in a turbulent boundary layer, J. Fluid Mech., 117 (1982) 1–26.
30 D.M. Deaves, Development and application of heavy gas dispersion models of varying complexity, J. Hazardous Materials, 16 (1987) 427–438.
31 C.J. Wheatley, A.J. Prince and P.W.M. Brighton, Comparison between data from the Thorney Island heavy gas trials and predictions of simple dispersion models, Report SRD R355, UKAEA Culcheth, U.K., 1986.
32 R.D. Fitzpatrick and C. Nussey, Unpublished Report, Health and Safety Executive, Broad Lane, Sheffield S3 7HQ, U.K.
33 T.O. Spicer and J.A. Havens, Field test validation of the DEGADIS model, J. Hazardous Materials, 16 (1987) 231–245.
34 J.W. Rottman, J.C.R. Hunt and A. Mercer, The initial and gravity-spreading phases of heavy gas dispersion: comparison of models with Phase I data, J. Hazardous Materials, 11 (1985) 261–279.
35 J.W. Rottman and J.E. Simpson, The initial development of gravity currents from fixed-volume releases of heavy fluids, In: G. Ooms and H. Tennekes (Eds.), Proc. IUTAM Symposium on Atmospheric Dispersion of Heavy Gases and Small Particles, Delft, The Netherlands, August 29–September 2, 1983, Springer Verlag, Berlin, 1984, pp. 347–359.
36 T.K. Fanneløp and F. Zumsteg, Special problems in heavy gas dispersion, In: S. Hartwig (Ed.), Heavy Gas and Risk Assessment III, D. Reidel 1986, pp. 123–135.
37 T.O. Spicer and J.A. Havens, Modelling the Phase I Thorney Island experiments, J. Hazardous Materials, 11 (1985) 237–260.
38 J.A. Havens, P.J. Schreurs and T.O. Spicer, Analysis and simulation of Thorney Island Trial 34, J. Hazardous Materials, 16 (1987) 139–148.
39 M.L. Riethmuller, Small-scale laboratory simulation of the Thorney Island trials, presented at the Symposium on Heavy Gas Dispersion Trials at Thorney Island — 2, Sheffield, Great Britain, September 1986.
40 A.P. van Ulden, A new bulk model for dense gas dispersion: two-dimensional spread in still air, In: G. Ooms and H. Tennekes (Eds.), Proc. IUTAM Symposium on Atmospheric Dispersion and Heavy Gases and Small Particles, Delft, The Netherlands, August 29–September 2, 1983, Springer Verlag, Berlin, 1984, pp. 419–440.

41 A.P. van Ulden, Heavy gas dispersion in still air, In: J.S. Puttock (Ed.), Proc. IMA Conference on Stably Stratified Flow and Dense Gas Dispersion, Chester, U.K., April 9–10, 1986, Oxford University Press, in press; also published under the title "The heavy gas mixing process in still air at Thorney Island and in the laboratory" in J. Hazardous Materials, 16 (1987) 411–425.
42 D.M. Webber and C.J. Wheatley, The effect of initial potential energy on the dilution of a heavy gas cloud, J. Hazardous Materials, 16 (1987) 357–380.
43 P.C. Chatwin, Towards a box model of all stages of heavy gas cloud dispersion, In: J.C.P. Hunt (Ed.), Turbulence and Diffusion in Stable Environments, Oxford University Press, 1985, pp. 259–291.
44 P.F. Linden and J.E. Simpson, Development of density discontinuities in a stratified fluid, In: J.S. Puttock (Ed.), Proc. IMA Conference on Stably Stratified Flow and Dense Gas Dispersion, Oxford University Press, in press.
45 C.J. Wheatley and A.J. Prince, Translational cloud speeds in the Thorney Island trials: mathematical modelling and data analysis, J. Hazardous Materials, 16 (1987) 185–199.
46 B. Crabol, A. Roux and V. Lhomme, Interpretation of the Thorney Island Phase I trials with the box model CIGALE2, J. Hazardous Materials, 16 (1987) 201–214.
47 R.J. Carpenter, R.P. Cleaver, P.J. Waite and M.A. English, The calibration of a simple model for dense gas dispersion using the Thorney Island Phase I trials data, J. Hazardous Materials, 16 (1987) 293–313.
48 Y. Gotaas, Heavy gas dispersion and environmental conditions as revealed by the Thorney Island experiments, J. Hazardous Materials, 11 (1985) 399–408.
49 Y. Riou, Dispersion de gaz lourds, essais de Thorney Island, analyse de la campagne 1982/1983 — essais sans obstacles, Electricité de France, unpublished report, 1984.
50 S. Hartwig, Improved understanding of heavy gas dispersion due to the analysis of the Thorney Island trials data, J. Hazardous Materials, 11 (1985) 417–423.
51 D.B. Pfenning and J.B. Cornwell, Computerised processing of Thorney Island trial data for comparison with model predictions, J. Hazardous Materials, 11 (1985) 347–368.
52 J.B. Cornwell and D.B. Pfenning, Comparison of Thorney Island data with heavy gas dispersion models, J. Hazardous Materials, 16 (1987) 315–337.
53 A. Mercer and C. Nussey, The Thorney Island continuous release trials: mass flux balances, J. Hazardous Materials, 16 (1987) 9–20.
54 R.E. Britter, A review of mixing experiments relevant to dense gas dispersion, In: J.S. Puttock (Ed.), Proc. IMA Conference on Stably Stratified Flow and Dense Gas Dispersion, Chester, U.K., April 9–10, 1986, Oxford University Press, in press.
55 J. McQuaid, Some experiments on the structure of stably-stratified shear flows, Technical Paper P21, Safety in Mines Research Establishment (U.K. Health and Safety Executive), Sheffield, U.K., 1976.
56 R.E. Britter, A note on stably-stratified shear flows, Unpublished, 1980.
57 J.A. Havens and T.O. Spicer, Development of an atmospheric dispersion model for heavier-than-air gas mixtures, Vol. I, Report CG-D-23-85, U.S. Coast Guard, Washington, D.C., 1985.
58 M.E. Davies and S. Singh, The Phase II trials: a data set on the effect of obstructions, J. Hazardous Materials, 11 (1985) 301–323.
59 J.W. Rottman, J.E. Simpson, J.C.R. Hunt and R.E. Britter, Unsteady gravity current flows over obstacles: some observations and analysis related to the Phase II trials, J. Hazardous Materials, 11 (1985) 325–340.
60 S. Hartwig, G. Schnatz and W. Heudorfer, Improved understanding of heavy gas dispersion and its modelling, In: G. Ooms and H. Tennekes (Eds.), Proc. IUTAM Symposium Atmospheric Dispersion and Small Particles, Delft, The Netherlands, August 29–September 2, 1983, Springer Verlag, Berlin, 1984, pp. 139–156.

61 W. Heudorfer, Heavy gas cloud and turbulence, In: S. Hartwig (Ed.), Heavy Gas and Risk Assessment III, D. Reidel, Dordrecht, 1986, pp. 53–72.
62 S.J. Sherrell, Fitting box models to the Thorney Island Phase I dataset, J. Hazardous Materials, 16 (1987) 395–410.
63 R.P. Koopman, R.T. Cederwall, D.L. Ermak, H.C. Goldwire, Jr., W.J. Hogan, J.W. McClure, T.G. McRae, D.L. Morgan, H.C. Rodean and J.H. Shinn, Analysis of Burro series 40-m^3 LNG spill experiments, J. Hazardous Materials, 6 (1982) 43–83.
64 J.S. Puttock, D.R. Blackmore and G.W. Colenbrander, Field experiments on dense gas dispersion, J. Hazardous Materials, 6 (1982) 13–41.
65 D.R. Johnson, V.H. Owen and J.C. Shipway, The performance of environmental sensors during heavy gas dispersion trials at Thorney Island, British Maritime Technology Ltd., Teddington, U.K., 1986.