Journal of Hazardous Materials, 16 (1987) 339–356
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

A COMPARISON BETWEEN THE VARIABILITY EXHIBITED IN SMALL SCALE EXPERIMENTS AND IN THE THORNEY ISLAND PHASE I TRIALS

J.K.W. DAVIES

Safety Engineering Laboratory, Health and Safety Executive, Broad Lane, Sheffield S3 7HQ (Great Britain)

(Received October 3, 1986; accepted February 19, 1987)

Summary

The turbulent nature of atmospheric dispersion gives rise to a large amount of variation in the outcome of nominally similar instantaneous releases of large quantities of heavier-than-air gas. This has repercussions for hazard assessment work, on the nature of dispersion models, and in the estimation of any empirical constants they may contain. Despite this, current box models concern themselves solely with the prediction of ensemble mean concentrations, and do not consider the equally important problem of the prediction of the extent of the variability about the mean. Consequently, such 'mean-only' models fail to provide the hazard analyst with all the distributional information he requires.

An empirical study of the distributional properties of the dispersion process would ideally require data from a large number of replicated releases. Such large datasets would require a prohibitive amount of time and money to produce. However, some limited datasets obtained at large scale and at small scale are available for analysis, and form the basis for the work described in this paper.

The results show, amongst other things, that the Lognormal distribution is well able to describe the between-replicated behaviour of the maximum concentration and of the dosage at a point in the flow field. This finding has important simplifying consequences on the use of 'mean-only' box models in the study of flammable and toxic releases.

1. Introduction

Since 1976 the Health and Safety Executive has been conducting a research programme into the atmospheric dispersion of large quantities of heavier-than-air gas in order to gain a better understanding of the risks involved in the storage, handling and use of large quantities of liquefied toxic or flammable gases. McQuaid and Roebuck [1] discuss the historic background and technical objectives of the programme, and stress its importance in the development of mathematical models for predicting the likely consequences of a hypothetical instantaneous release of heavy gas.

The turbulent nature of atmospheric dispersion gives rise to a large amount of variation in the outcome of nominally similar instantaneous releases. This

in turn has repercussions for hazard assessment work, on the nature of the dispersion models, and in the estimation of any empirical constants they may contain.

The statistical concepts needed in the mathematical description of the dispersion process are discussed in Chatwin [2], who points out that as far as releases of flammable gases are concerned, the fundamental concept is that of the probability distribution of gas concentration as a function of position in the flow field and time from release. For releases of toxic gases, on the other hand, the fundamental concept is that of the probability distribution of the dosage, i.e. the time integral of (some power of) the concentration, as a function of position. If either of these distributions could be modelled for a given hypothetical release then the hazard analyst would be able to make direct probabilistic assessments of the level of risk. In both cases, the effect of 'intermittency' could be handled by allowing the corresponding cumulative distributions to have a discontinuity at the origin. This problem does not arise in the datasets considered here, in view of the fact that meandering is artifically reduced in windtunnels, although meandering could well assume considerable importance in practice.

So far as the author is aware, all current 'box' models concern themselves solely with the prediction of ensemble mean concentrations, and do not consider the equally important problem of the prediction of the extent of variability about the mean. The consequence is that such 'mean-only' models fail to provide the hazard analyst with all the distributional information he requires.

Of course, knowledge of the mean and spread of a probability distribution is not in general sufficient to determine the form of the distribution. It will be seen, however, that for practical purposes a two-parameter probability model is quite well able to describe the distributions encountered.

The preceding remarks make clear the necessity of having available datasets of replicated releases, conducted under a range of initial conditions, on which to study the distributional properties of gas concentration (for flammable gases) and dosage (for toxic gases). Owing to cost constraints, it was not practicable to replicate any of the medium-scale releases at Thorney Island, though some of the latter may be regarded as forming a small set of replicates. However, at smaller scale a release was replicated 20 times [3]. It is this dataset, hereafter referred to as the WSL dataset, which is the subject of most of the work reported here.

Section 2 describes the findings on the time variation of the ensemble mean and standard deviation of gas concentration in the WSL dataset, the implications for the fitting of box models being the subject of Section 3. Section 4 is a study of the distributional properties of gas concentration.

Section 5 is a parallel study of the distributional properties of dosage, and contains a description of a general parametric model of the mean concentration field which could be used in the analysis of data from dense gas dispersion

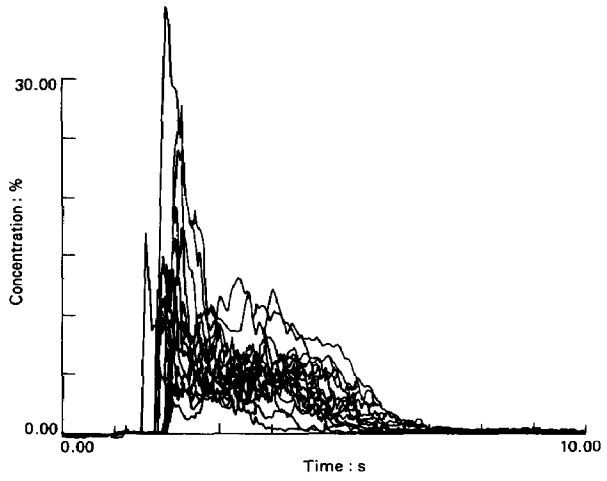

Fig. 1. The 21 concentration–time traces in the WSL dataset.

experiments. Section 6 discusses the matter of the comparison of results derived from the WSL dataset with those derived from the Phase I Thorney Island series and the wind tunnel experiments of Meroney and Lohmeyer [4].

In passing, it should be noted that in the notation adopted here, which is consistent with the usual statistical practice, the ensemble mean of concentration is denoted by $\mu(C)$, and the ensemble standard deviation by $\sigma(C)$. In fluid mechanics literature, $\mu(C)$ is usually written as $\bar{C}$ or $<C>$, and $\sigma(C)$ as $(\overline{c^2})^{0.5}$.

2. Moment properties of the WSL dataset

The 21 concentration–time traces in the WSL dataset are shown in Fig. 1. They were recorded at 20 Hz, the upper frequency limit of the detector, each trace covering a period of 21 s, of which the first second was a 'run-up' before gas release. (NB All the times quoted here are from the start-up of recording and therefore contain the 1 s run-up period.) The replications were conducted with an initial bulk Richardson number of $Ri = 1.9$, based on a windspeed of 0.84 m s^{-1} at a reference height of 110 mm. Figure 1 in fact shows the first 10 s of each replication only, since this was a conveniently short space of time in which the gas concentration was sensibly different from zero. Figure 1 may be compared to Fig. 46 of Hall et al. [3], which was plotted from the same dataset on a scaled time axis.

The level of sampling variability in the WSL dataset was estimated by the method of 'bootstrapping' [5]. The principle of the method may perhaps best be illustrated with reference to an ideal experiment. Suppose we have observed a sample $x_1, x_2, \ldots, x_n$ of size n from a given (not necessarily Normal) parent population, and we wish to study the sampling variation of, say, the sample

mean, $\bar{x}$. To do this we draw further samples of size n from the parent population, compute the mean value of each new sample, and construct a histogram of the mean values thus derived, continuing this process until a clear picture emerges. In practice, of course, it is impossible to proceed in this way because, apart from other considerations, we do not know the cumulative distribution function (CDF) of the parent population, and so are not able to draw further samples from it. We do, however, know the empirical CDF of the observed sample: this is simply a step function with jumps of size n^{-1} at the sample values arranged in increasing order. The method of bootstrapping simulates the ideal experiment described above the using the known empirical CDF in place of the unknown true CDF.

It may be seen that the larger the sample size, n, the closer will the empirical CDF approach the true CDF, and the closer will the histogram of the sample mean derived from bootstrapping the empirical CDF approach the histogram that would have been derived from the true CDF. In cases where the distribution of the bootstrapped quantity may be derived theoretically, it has been found that its sampling variance is inversely proportional to the sample size, n, a property that bootstrapped estimates share with those derived by more conventional methods in which the parent population is assumed to belong, for example, to the Normal family. The sampling behaviour of the bootstrapped mean and standard deviation for the WSL dataset turned out to be in good agreement with those expected on the basis of Normal sampling theory. This is by no means always the case in practice, and it is of some importance to know that our results do not contain errors due to a false assumption of Normality in the parent population.

The first statistical property to be considered is the ensemble mean as a function of time. Figure 2 shows the median and the upper and lower 2.5 percentile curves of the variation in sample mean derived from 1000 bootstrapped samples of size 21. (NB The upper 2.5 percentile of a random variable such as a sample mean is that value which is exceeded 2.5 percent of the time, the lower 2.5 percentile being defined similarly. The interval between the lower and upper 2.5 percentiles thus constitutes a 95 percent confidence interval.) The number of bootstrapped samples required to arrive at stable estimates of the upper and lower percentiles was found by trial and error. The estimates obtained in this way were stable to within a few percentage points over repeated bootstrapping.

Following the bootstrapping method, we are able to say that there is a 95% chance that the $\mu(C)$ curve itself lies between the upper and lower bootstrapped curves. The extent of the spread between the upper and lower curves is consistent with the Normal Theory value of $\pm 1.96\sigma(C)/21^{0.5} = \pm 0.43\sigma(C)$, where the magnitude of $\sigma(C)$ may be ascertained from Fig. 4.

By comparing the time variation in the median concentration with the sampling variation as indicated by the upper and lower 2.5 percentile curves in Fig. 2, we can say that the WSL dataset is consistent with the claim that $\mu(C)$ is

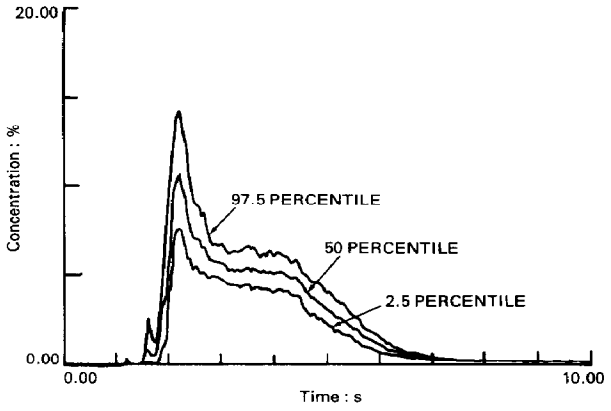

Fig. 2. Bootstrapped 95% confidence bounds for the ensemble mean concentration: original dataset.

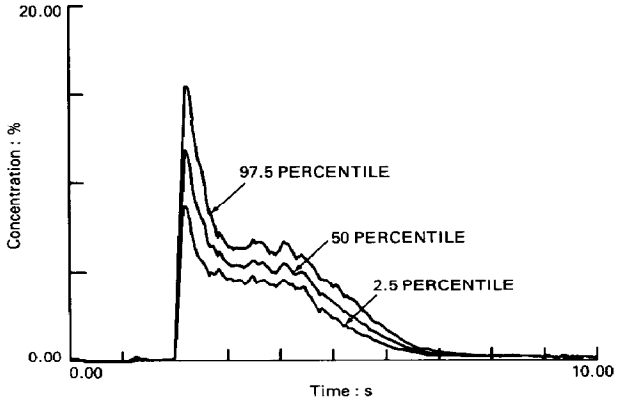

Fig. 3. Bootstrapped 95% confidence bounds for the ensemble mean concentration: re-aligned dataset.

constant between 2.4 s and 4.4 s. These times are the end points of the longest interval such that a horizontal line (at a concentration of roughly 6%) lies entirely within the confidence region included between the upper and lower curves. The same conclusion follows from Fig. 3, which is entirely analogous to Fig. 2 except that it was derived from the original dataset realigned so that all the arrival times for the different replicates coincided with the arrival time of the first replicate.

In the WSL dataset it appears likely that the ensemble mean curve is visually similar to the individual realisations: little smoothing of discontinuities has taken place, because of the relatively low dispersion in cloud arrival time, cloud passage time and maximum concentration. If, however, the dispersion in these three quantities had been much greater, the resemblance would have been more

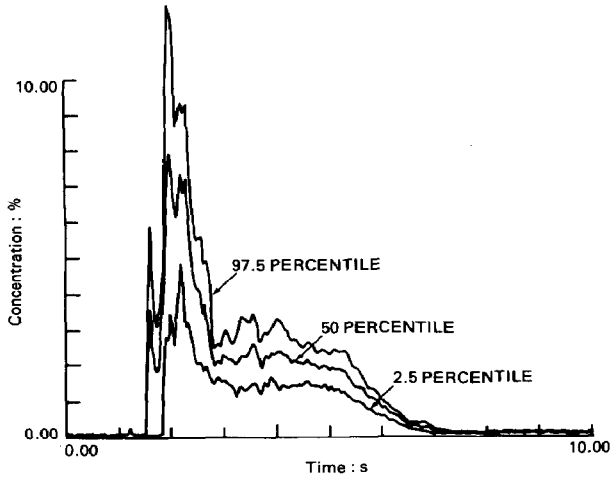

Fig. 4. Bootstrapped 95% confidence bounds for the ensemble standard deviation: original dataset.

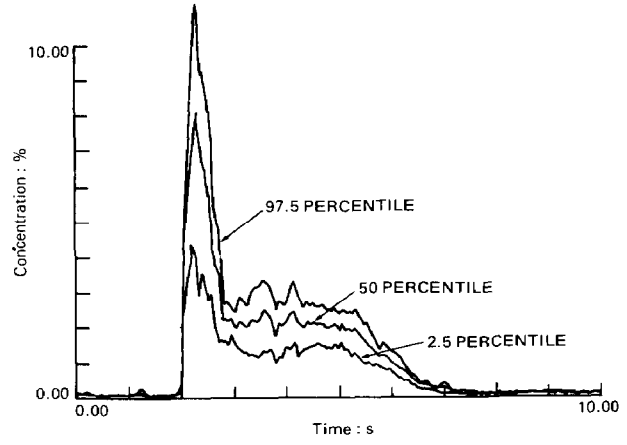

Fig. 5. Bootstrapped 95% confidence bounds for the ensemble standard deviation: re-aligned dataset.

tenuous, as Puttock and Colenbrander [6] point out. Whether this is likely to be the case in practice remains to be seen. Carn et al. [7] emphasize, however – quite correctly, in the author's view – that when fitting a box model to a set of experimental realisations the choice of ensemble is a matter for the user of the model, but once the ensemble has been specified the fitting of the model can only proceed according to established statistical methods, which take into account the variation in concentration over the ensemble thus chosen.

Variation in the ensemble standard deviation $\sigma(C)$ may be treated in exactly the same way as variation in the ensemble mean $\mu(C)$. Figure 4 shows the results of bootstrapping the standard deviation in a manner analogous to Fig. 2. It will be seen that as in Fig. 2 there appears to be a period of time during which $\sigma(C)$ is to all intents and purposes constant. Using the same criterion

as before, the period in question is roughly 2.4–5.4 s. Practically the same conclusion follows from Fig. 5, which shows the time variation of standard deviation in the realigned dataset: the stable period is 2.6–5.6 s.

The picture thus emerges that for the WSL dataset the flow assumes a statistically stationary pattern for a period after the passage of the initial peak: during this period the intensity $I(C)=\sigma(C)/\mu(C)$ is constant. The area-averaged results of Wheatley et al. [8] for the Phase I Thorney Island trials suggest that intensity is roughly constant, at least near the centroid of the cloud, a result compatible with the proposal of Chatwin and Sullivan [9], made on the basis of theoretical considerations. The implication of this finding for the optimising of box models is: relative prediction errors should be used rather than absolute prediction errors. The justification for this claim is the subject of the following section. As an aside, it may be remarked that quite apart from its significance for box model optimisation, the intensity is widely agreed to be a useful measure of statistical variability.

3. The optimisation of box models

On the assumption of Gaussian errors, the systematic optimisation of a box model M, in which the mean concentration field $\mu(C(x,t))$ is approximated by a function $f(x,t;\theta)$, where θ is a vector of empirical constants, is achieved by minimising an objective function of the form

$$\phi(\theta)=N^{-1}\sum_{x,t}\left(\frac{C(x,t)-f(x,t;\theta)}{\sigma(C(x,t))}\right)^2$$

with respect to variation in σ where the summation is over N suitably chosen points in (x,t) space. If, throughout the summation, we may set $\sigma(C(x,t))=I\mu(C(x,t))$ then we can write

$$\phi=N^{-1}\,I^{-2}\,\Sigma\left(\frac{C-f(\theta)}{\mu}\right)^2$$

where the dependence on x and t has been suppressed for notational convenience. If, furthermore, the model is at all representative of reality, we can write $\mu=f$ when $\theta=\theta^*$ (the optimum parameter value), and so are led to consider a slightly revised objective function of the form

$$\phi=N^{-1}\,\Sigma\left(\frac{C-f(\theta)}{f(\theta)}\right)^2$$

which is proportional to the sum of squares of relative prediction errors.

It should be noted that the use of absolute prediction errors in the optimization of box models would only be appropriate if the variance were reasonably constant throughout the summation. This is known not to be the case: Chatwin

[2] has shown that, averaged over the cloud, $\sigma^2(C)$ is a decreasing function of time.

An estimate of the intensity in any given case may be obtained by the following argument. Assuming that optimization takes place at $\theta=\theta^*$, we can write

$$\phi(\theta^*)=\phi^*=N^{-1}\sum\left(\frac{C-f^*}{f^*}\right)^2\approx N^{-1}\sum\left(\frac{C-f^*}{\mu}\right)^2$$

since $f^*\approx\mu$ and hence

$$\phi^*\approx N^{-1}\,I^2\sum\left(\frac{C-f^*}{\sigma}\right)^2$$

since by assumption $\sigma=I\mu$ throughout the summation. Assuming further – see below – that $(C-f^*)/\sigma\sim N(\alpha,1)$, a Normal distribution with mean α and variance unity, it follows from the properties of the non-central chi-square distribution [10], that

$$E[\phi^*]=N^{-1}\,I^2\,E\left[\left(\frac{C-f^*}{\sigma}\right)^2\right]=I^2(1+\alpha^2)=I^2+\beta^2$$

where $\beta=\alpha\,I=E[(C-f^*)/\mu]$ is the (relative) bias or systematic error in the model. By setting $E[\phi^*]=\phi^*$ in the usual way, it follows that with respect to the model M the intensity I and the bias β may be estimated from $I^2=$variance $[(C-f^*)/f^*]$ and $\beta=$mean $[(C-f^*)/f^*]$. An immediate corollary of the above argument is that the minimum possible value of ϕ^* occurs for a 'bias-free' model with $\beta=0$, in which case we would have $\phi^*=I^2$, a purely physical parameter. In all other cases our estimate of I^2 would be affected unavoidably by the presence of model bias, though it would always be the case that $\phi^*>I^2$, so that ϕ^* would act in practice as an upper bound.

Another corollary is that if we have two models, M_1: $\mu=f_1(x,t;\theta)$ and M_2: $\mu=f_2(x,t;\theta)$ for which the resulting minimum objective function values were ϕ_1^* and ϕ_2^* and $\phi_1^*<\phi_2^*$, then M_1 should be preferred to M_2, even though it might be the case that $\beta_1>\beta_2$, since ϕ^* measures the overall performance of the model — variance + (bias)^2 — and neither aspect should be neglected at the expense of the other.

4. Distributional properties of concentration in the WSL dataset

In the previous section we were concerned with some of the statistical properties of the WSL dataset which can be computed directly without needing to know the underlying probability distributions. In this section we describe the results of trying to fit three well-known probability models to the dataset, and discuss some of the implications.

TABLE 1

Variation in 99th percentile of concentration

Time (s)	$C^{N}_{0.99}$	$C^{B}_{0.99}$	$C^{LN}_{0.99}$
2.1	24.5	29.2	35.9
2.2	28.2	33.9	43.3
2.3	26.5	32.9	37.6
2.4	19.6	23.4	28.2
2.5	17.0	20.0	24.3
2.6	15.2	17.7	19.1
2.7	14.1	16.2	16.3
–	31.8	35.8	41.3

The three probability models are the Normal (N), Lognormal (LN) and Beta (B) models [10]. All are two-parameter models and so are able to cope to a greater or lesser extent with the location and spread of the data to which they are fitted, but beyond this have no power to reflect more detailed aspects of the distributional shape.

The models are fitted using the method of moments, i.e. by making the first two theoretical moments equal to the first two sample moments — and then ranking the fits thus obtained in terms of the maximum absolute deviation over the whole range of concentrations between each of the three model cumulative distribution functions (CDFs), and the empirical CDF. The results of applying this procedure over the period 1.7–5 s (i.e. from the arrival of the signal to the end of the 'plateau' after the main signal has decayed) may be summarised as follows:

(i) in the critical period 1.7–2.7 s which contains the peak concentration the Lognormal model is (almost) uniformly superior to the other two models;

(ii) in the period 2.7–3.6 s the Lognormal and Beta distributions vie for superiority, with no clear picture emerging;

(iii) subsequently, from 3.6 s to 5 s the Beta and Normal distributions alternate as best fits.

The behaviour of the higher percentiles which result from adopting each of the three probability models in turn as the true model may be seen in Table 1. In this table we show the variation in the 99th percentile of concentration over the period 2.1–2.7 s. The main feature is the fact that at $t = 2.2$ s — the time at which the sample mean reaches its maximum value — the predicted 99th percentile is no less than 50% higher with the preferred Lognormal model than with the Normal model.

The last row of Table 1 was included to show the effect of fitting the same three probability models are fitted to the maximum value in each replication, irrespective of the time at which it occurred. This row should be compared with

TABLE 2

Pooled 99-percentiles and local ignition prbabilities

	ECDF	N	B	LN
$C_{0.99}$ (%)	17	12.5	14	17
$\int_{0.05}^{0.15} P(C)\,dC$	0.52	0.55	0.52	0.49

the second row of the table, corresponding to 2.2 s from start up. It will be seen that the two sets of 99-percentiles are very similar.

It is interesting to observe what happens when the concentration values for the period $t=2.4$ s to $t=4.4$ s — when both $\mu(C)$ and $\sigma(C)$ are steady — are pooled together. Table 2 shows the results of estimating the 99th percentile of concentration and the point ignition probability (over the usual flammable range for methane, i.e. 5% to 15%) from the pooled data, which amount to 841 readings. The 'ECDF" column refers to the estimate derived from the empirical cumulative distribution; the remaining three columns to the estimates derived from the Normal, Beta and Lognormal models. The point ignition probability was determined directly from the empirical CDF and by numerical integration of the fitted distribution functions in the case of the three probability models. Comparing these models with the pooled data, it will be seen that the Lognormal model gives an accurate estimate of $C_{0.99}$, and an acceptable figure for the point ignition probability. This is confirmed by the finding that over the same period we have $\sigma(C)/\mu(C) = 0.45 \pm 0.08$ and $\sigma(\ln C) = 0.53 \pm 0.13$.

5. Distributional properties of dosage in the WSL dataset

As far as dosages are concerned, the picture is more clear cut. Defining the dosage as $D(t) = \int_0^t C\,dt'$, where $t=0$ corresponds to start-up, we find that within the limits set by sampling theory the dosage is pretty well Lognormal. The basis for this claim is the behaviour of the skewness, γ_1, and the flatness, γ_2, of the successive samples of size 21 of $\ln D(t)$ as time increases. The skewness and flatness of a probability distribution are defined, respectively, as

$$\gamma_1 = \mu_3/{\mu_2}^{3/2},\ \gamma_2 = (\mu_4/{\mu_2}^2) - 3$$

where μ_2, μ_3 and μ_4 are the 2nd, 3rd and 4th moments about the mean — see, e.g. [11]. Evidently, γ_1 measures departure from symmetry, since if the probability distribution were symmetric we would have $\mu_3=0$, and hence $\gamma_1=0$. Similarly, γ_2 measures how flat or how peaked the distribution is relative to the Normal distribution, for which $\gamma_2=0$ by definition. The sharply-peaked

two-sided negative exponential distribution, for example, has $\gamma_2 = 3$, while the broad-shouldered uniform distribution has $\gamma_2 = -6/5$. Note that these results contain no reference to mean and standard denotion, since by definition γ_1 and γ_2 are independent of location and scale.

The skewness $\gamma_1(\ln D)$ varies from -0.47 to -0.1, while the flatness $\gamma_2(\ln D)$ varies between 0.3 and -0.3. Assuming Lognormality, the sampling standard deviations are ± 0.53 for γ_1, and ± 1.1 for γ_2, given a sample of size 21, the ideal values being $\gamma_1 = \gamma_2 = 0$, of course. The same analysis performed on the dosage defined as $D(t) = \int_0^t C^{2.75}\, dt'$ shows that, if anything, the dosage is then even more closely Lognormal: $\gamma_1(\ln D)$ varies between -0.24 and 0.14 and $\gamma_2(\ln D)$ between -0.1 and 0.1. Moving the origin of the integration to $t = 2.1$ s (just before the signal arrives) makes no difference to the skewness and flatness.

The standard deviation of the log-dosage shows a consistent tendency to decrease with time. Calculations demonstrate that throughout the signal $\sigma(\ln \int C^{2.75}\, dt') = 2.75\, \sigma(\ln \int C dt')$, a result which, if generally true, would have important repercussions for hazard assessors, since it suggests that substances within a wide range of 'toxicity exponents' can be handled, providing that something is known of the standard deviation $\sigma(\ln \int C dt')$ of the log-dosage as usually defined. In our case, this quantity is closely comparable to the standard deviation of the log-concentration, which previous analysis has shown to equal the concentration intensity: $I(C) = \sigma(C)/\mu(C) = \sigma(\ln\ C)$, because of Lognormality.

No conclusions of a specifically statistical nature emerge from a consideration of the behaviour of the mean dosage. On the other hand, a potentially useful parameterization of the mean concentration field does emerge. By definition, the mean dosage at any given position is an increasing function of time, and by definition would be sigmoidal if the mean concentration profile were Gaussian. However, this is not the case: the mean concentration profile is in general strongly positively skewed. It is, however, well known that a logarithmic transformation is able to roughly symmetrise a positively skewed function, in which case we would find that $\mu(C(t))$ is roughly proportional to $\phi[(1/\sigma) \ln\ (t - t_a)/t_m - t_a)]$, where $\phi(\ )$ is the standard Gaussian function — not to be confused with the 'ϕ' used earlier to denote the objective function. In this expression, t_a is the arrival time of the mean profile, t_m is the time at which the mean profile attains its maximum value, and σ is the spread of the mean profile in logarithmic time.

More generally, we could set

$$\ln \mu(C(x,t)) = a_0\ (x) + \sum_{k=2}^{k=\infty} a_k\ (x) \left(\frac{1}{\sigma} \ln \frac{t - t_a(x)}{t_m(x) - t_a(x)} \right)^k$$

say, and retain the first few terms in the summation, which would probably be sufficient in practice, in view of the property of the logarithmic transformation referred to above. It should be noted that $a_1(x) = 0$ by definition. By setting $t = t_m(x)$ we find that

$$a_0(x) = \ln C(x, t_m(x)) = \ln \mu_m\ (x)$$

say, so that $\mu_m = \exp a_0(x)$ gives the maximum mean concentration as a function of position. Substituting for $a_0(x)$ in the proposed mean-field model, we find that finally

$$\ln \frac{\mu(C)}{\mu_m} = \sum_{k=2}^{k=\infty} a_k \left(\frac{1}{\sigma} \ln \frac{t - t_a}{t_m - t_a} \right)^k$$

This parameterization is invariant with respect to a change in time origin and time unit, and is centred on the physically significant aspects of the concentration profile. With a few terms retained in the summation, it should give a good representation of $\mu(C)$, at least for $t \approx t_m$, more accuracy being achievable by including further terms. Finally, it should be noted that the formulation is directly in terms of the logarithm of the mean concentration field which, according to the earlier evidence, has a constant variance over the cloud, so that the model could be fitted by a simple unweighted least squares procedure.

6. Comparison with other datasets

The results described so far have been derived from a study of the WSL dataset, largely because it consisted of a fair number of replicates made under tightly controlled conditions. Additional datasets are also, of course, available from the Thorney Island series of field trials [1] and from some work described by Meroney and Lohmeyer [4]. The results of analysing these are set out below.

The analysis will be carried out in terms of C_{max}, since this is an easily determined summary statistic of the concentration record at a sensor, and is clearly influential in trying to determine the probability that a local ignition will occur at the sensor in the case of a flammable release. The other summary statistic, which is important for toxic releases, is the total dose $\int C dt'$. In both cases the statistic removes the time element, and thereby considerably reduces the amount of data to be considered.

6.1 The Thorney Island dataset

Since the Thorney Island trials were not replicated, the most straightforward way of deriving information on the intensity in any given trial is to take the maximum concentration C_{max} observed by each sensor that 'saw' gas, to fit a statistical model to the data thus accumulated, and finally to compute the root-mean-square (RMS) variation $\sigma(\ln C_{max})$ about the model. (It may be

TABLE 3

Statistics of C_{max} for Phase I Thorney Island Trials

	Trial No.									
	15	19	14	16	07	13	17	18	09	12
$\sigma(\ln C_{max})$	0.45	0.46	0.48	0.53	0.56	0.57	0.58	0.59	0.74	0.79
Ri_{10}	1.9	3.7	2.1	3.0	9.3	2.2	13.8	1.7	26.5	25.2
Windspeed ($m\ s^{-1}$)	5.4	6.4	6.8	4.8	3.2	7.5	5.0	7.4	1.7	2.6
Atm. Stab.	C/D (2.5)	D/E (3.5)	C/D (2.5)	D (3)	E (4)	D (3)	D/E (3.5)	D (3)	D/E (3.5)	E (4)

shown – see Appendix 1 – that to a first order of approximation, $\sigma(\ln C_{max}) = I(C_{max})$.) The statistical model adopted takes advantage of the apparent Lognormality of the concentration near the maximum (as discussed in Section 4) and assumes that in the cloud $\ln C_{max}$ is a quadratic function of position plus a Normally distributed homogeneous random component.

The sensors which did not 'see' gas could have been included in the statistical model by regarding them as sensors in which C_{max} was less than some lower limit of detection, say 0.1%, but was otherwise unknown. Such an extension to the model would have turned a straightforward linear optimisation problem into a potentially difficult non-linear problem, and was not thought justifiable in the circumstances.

The results of this exercise carried out on 10 of the Phase I Thorney Island series are given in Table 3, in which the trials are arranged in order of increasing RMS variation, $\sigma(\ln C_{max})$, for clarity. There is no evidence in the table of any linear or power law relationship between the RMS variation and the initial bulk Richardson number, the windspeed, or the Pasquill atmospheric stability category (when coded as shown).

6.2 The Meroney and Lohmeyer [4] dataset

This dataset consisted of maximum concentration readings from 100 wind tunnel replicates taken at an initial bulk Richardson number of 14.9 and a windspeed of 0.4 m s^{-1}. On applying the distributional analysis described in Section 4 above, it was found that the three probability models were ranked in the order 'Normal better than Beta better than Lognormal', although there was little to choose between them. Meroney and Lohmeyer [4] classified the distribution as Normal. The 99-percentiles of C_{max} under the three models were 10.2%, 10.5% and 11.5%, in order of preference. Comparing these results with those from the line for 2.2 s in Table 1, it will be seen that the Meroney and

TABLE 4

Statistics of C_{max} for other datasets

	Experiment	
	WSL	Meroney and Lohmeyer
Intensity	0.52	0.21
Ri_{10}	1.9	14.9
Windspeed ($m\,s^{-1}$)	0.84	0.4

Lohmeyer maxima show much less sensitivity to the distributional model than do the WSL maxima. In addition, the intensity in the Meroney and Lohmeyer dataset was found to be 21% ±2%.

Table 4 shows the WSL and Meroney and Lohmeyer results for comparison with those in Table 3. That such a comparison may be made follows from the fact that in each of the trials considered in Table 3 the systematic effects on C_{max} due to sensor position have been removed by fitting a quadratic function of position as in Section 6.1, and the value of the intensity $\sigma(\ln C_{max})$ computed from the remaining random component. Fitting a least-squares straight line to the intensities in Tables 3 and 4 taken together and using bulk Richardson number and windspeed in turn as 'explanatory' variables showed that the slope coefficient was insignificantly different from zero because of the magnitude of the dispersion about the line. There are thus no grounds on the evidence considered here for assuming any functional relationship between intensity and either of the explanatory variables considered, the contrast between the Meroney and Lohmeyer intensity and the remaining intensities being more apparent than real. The overall mean intensity turned out to be 54%.

6.3 Discussion

The level of intensity noted above has important consequences for the use of box models in the assessment of flammable hazards. It is the practice to allow a 'margin of safety' in any given case by computing not only the LFL contour but also the 1/2-LFL and possibly 1/4-LFL contours, thereby acknowledging implicitly the stochastic nature of the dispersion process. The degree of conservatism represented by this practice has not, however, been quantified, as far as the author is aware. The solution to this problem lies in recognising the distinctions between contours of equal concentration and contours of equal chance of local ignition, and developing a statistical argument to establish the connection between them. The steps in this argument are set out briefly below.

The first step, with reference to Table 3, is to determine the value of the ratio

TABLE 5

Values of $r=\mu C_{LFL}$ for Phase I Thorney Island trials

	Trial									
	15	19	14	16	07	13	17	18	09	12
$\sigma(\ln C_{max})$	0.45	0.46	0.48	0.63	0.56	0.57	0.58	0.59	0.74	0.79
$r=\mu/C_{LFL}$	0.25	0.24	0.23	0.19	0.18	0.17	0.17	0.16	0.10	0.09

$r=\mu/C_{LFL}$ for each of the trials such that with μ as the ensemble median value of C_{max}, the chance of a local ignition would be, say, 1-in-1000, assuming as above that C_{max} is distributed Lognormally and that a local ignition would occur if $C_{max} > C_{LFL}$. The value of μ we are seeking satisfies

$$pr\{C_{max} > C_{LFL} \mid \mu\} = 1 - \Phi\left(\frac{\ln C_{max} - \ln \mu}{\sigma(\ln C_{max})}\right) = p$$

say, where we have set $p=0.001$ and $\Phi(\ )$ is the standard Normal CDF. By inverting this relationship we have

$$\frac{\ln C_{max} - \ln \mu}{\sigma(\ln C_{max})} = \Phi^{-1}(1-p)$$

which leads to

$$r=\mu/C_{max} = \exp\{-\sigma(\ln C_{max})\ \Phi^{-1}(1-p)\}$$

The results are given in Table 5.

The second step is to study the variation of r from trial to trial with a view to determining a value of r that has a reasonably high chance of being exceeded in the ensemble represented by the Phase I trials – say the lower 5 percentile, $r_{0.05}$, which has a roughly 20-to-1 chance of being exceeded and which we would therefore expect to be exceeded in only 1 case out of 20. Assuming that r is Lognormal we find that $r_{0.05}=0.09$, while assuming that r is Normal we obtain $r_{0.05}=0.08$, so that it appears reasonable to assume on the basis of the data available that $r_{0.05}=0.08$, i.e. $\mu=1/12\ C_{LFL}$.

The third step is to note that in box models the concentration is always a decreasing function of distance from the source. It follows from this that there is a roughly 20-to-1 chance that the 1/12-LFL contour will contain the 1-in-1000 local ignition contour. At the other end of the range of r it is easy to show that $r_{0.09}=0.25$ under both Normality and Lognormality, so that there is a 9-to-1 chance that the $\frac{1}{4}$-LFL contour will lie inside the 1-in-1000 local ignition contour. Similarly in the case of the 1/2-LFL contour it can be shown that there is a 50/50 chance that it coincides with the 1-in-8 chance of local ignition contour. Should the corresponding chance of a local ignition propogating

through the cloud be high then these results indicate that the use of the 1/2-LFL contour for the assessment of flammable hazards needs to be reviewed. The work of Birch et al. [12] shows that in the case of a turbulent jet the 1/2-LFL contour could lie within the flammable boundary within which an initial flame kernel once formed will give rise to a total light-up of the jet, but outside which the kernel will be extinguished suggests that in more complex phenomena the probability that a local ignition will give rise to a generalised ignition will decrease with distance from the source.

I have assumed in the argument above that box models predict the ensemble median concentration, as opposed to the ensemble mean. If, however, it is assumed that they predict the ensemble mean then these conclusions would have to be modified slightly to allow for the fact that in a Lognormal distribution the mean always exceeds the median by an amount depending on the variance. In the cases considered here the relative difference (mean-median)/median is never greater than 25% and is usually much less: the general trend of the conclusions is not disturbed.

7. Conclusions

Several conclusions follow from the analysis as here described. They are:

(i) Within the limits of sampling error, the ensemble mean and standard deviation of concentration in the WSL dataset – and hence the ensemble intensity – have been shown to be constant for a period after the initial arrival of the concentration signal.

(ii) The intensity of maximum concentration found in the WSL dataset is comparable to that found in the Phase I Thorney Island series and in the Meroney and Lohmeyer dataset. There is no evidence of any linear or power law relationship between the intensity and either the initial bulk Richardson number, the windspeed or the Pasquill stability category. The overall geometric mean intensity of these data sets was found to be 0.52.

(iii) Assuming constancy of concentration intensity throughout the cloud, a form of objective function based on relative prediction errors has been proposed when optimizing box models, and a statistical argument has been brought forward to show how both the concentration intensity and the model bias may be estimated as a by-product of the optimizing process.

(iv) Three probability models of gas concentration have been considered: the Normal, the Lognormal and the Beta. A comparative goodness-of-fit test has shown that of these three, the Lognormal model is to be preferred in the initial development of the signal. This preference gradually gives way to a preference for the Normal model, with the Beta model occupying a consistently intermediate position.

(v) A parallel distributional study of the dosage shows that the Lognormal model is adequate throughout the signal, the more so for higher 'toxicity exponents'.

(vi) A relation has been found to exist between the geometric variances of dosage for different 'toxicity exponents' which, if true more generally, would bring about a great simplification in the assessment of toxic hazard.

(vii) A general parametric model has been proposed for the ensemble mean concentration field, based on the similarity between concentration/time profiles and the Lognormal probability density function.

(viii) The variation in intensity levels within the Thorney Island Trials (7, 9, 12–19) has been used to quantify the degree of conservatism built into criteria used for the assessment of flammable hazards. It is concluded that there is a 50/50 chance that the 1/2-LFL contour coincides with the 1-in-8 chance of local ignition contour and that there is roughly a 10% chance that the 1/4-LFL contour contains the 1-in-1000 local ignition contour.

It is recommended that further datasets are needed to verify the conclusions and to test the proposals reported here. Such datasets should be gathered using

(i) more replicates than the 20 of the WSL dataset, in order to provide for a more accurate analysis of distributional properties;

(ii) more instrumentation, in order to study the interaction between the concentration field and the wind field; and

(iii) a variety of initial conditions, in order to confirm or refute the finding that the intensity of maximum concentration is insensitive to variation in the initial bulk Richardson number, the windspeed and the Pasquill stability category.

Acknowledgements

The author wishes to thank Dr. J. McQuaid for drawing this attention to Efron's work, to Dr. D.J. Hall for supplying results of his wind tunnel data on magnetic tape and to Professor R.N. Meroney for supplying additional data from the work reported by Meroney and Lohmeyer.

References

1 J. McQuaid and B. Roebuck, Large scale field trials on dense vapour dispersion, Report No. EUR 10029, Commission of the European Communities, Brussels, Belgium, 1985.

2 P.C. Chatwin, The statistical description of the dispersion of heavy gas clouds, Report on Contract No. 1189/01.01, Research and Laboratory Services Division, Health and Safety Executive, Sheffield S3 7HQ, Great Britain, 1981.

3 D.J. Hall, E.J. Hollis and H. Ishaq, A wind tunnel model of the Porton dense gas spill field trials, Report No. LR394(AP), Warren Spring Laboratory, Stevenage, Great Britain, 1982.
4 R.N. Meroney and A. Lohmeyer, Gravity spreading and dispersion of dense gas clouds released suddenly into a turbulent boundary layer, Gas Research Institute Research Report GRI-81/0025, Chicago, IL, U.S.A., 1982, Section 4.
5 B. Efron, Review, SIAM 21(4) (1979) 460–480.
6 J.S. Puttock and G.W. Colenbrander, Thorney Island data and dispersion modelling, J. Hazardous Materials, 11 (1985) 391–397.
7 K.K. Carn, S.J. Sherrell and P.C. Chatwin, Analysis of Thorney Island data: variability and box models, IMA Conference, Chester, Great Britain, 1986.
8 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, Ref. No. SRD R355, Safety and Reliability Directorate, UKAEA, Culcheth, Great Britain, 1986.
9 P.C. Chatwin and P.J. Sullivan, The relative diffusion of a cloud of passive contaminant in incompressible fluid flow, J. Fluid Mech., 91 (1979) 337–356.
10 J. Patel, C.M. Kapadia and D.B. Owen, Handbook of Statistical Distributions, Marcel Dekker, New York 1976, pp. 27, 28 and 32.
11 G.E.P. Box and G.C. Tiao, Bayesian Inference in Statistical Analysis, Addision-Wesley, 1973, p. 150.
12 A.D. Birch, D.R. Brown and M.G. Dodson, Ignition probabilities in turbulent mixing flows, Report No. MRS E374, British gas Corp., Midlands Research Station, June 1980.
13 A.A. Sveshnikov, Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions, Dover, New York, 1978.

Appendix 1

Derivation of $\sigma(\ln C_{max}) = I(C_{max})$

Sveshnikov [13] shows by the usual linearization argument that if X is a random variable and $g(\ )$ is any differentiable function, then the variance of the random variable $Y=g(X)$ is related to the variance of X by the expression

$$\sigma_Y^2 \approx (\dot{g}(\mu_X))^2 \sigma_X^2$$

Substituting $g(X) = \ln X$ it follows that $\dot{g}(\mu_X) = \mu_X^{-1}$ and $\sigma_Y^2 = \sigma_{\ln X}^2 \approx \mu_X^{-2}\sigma_X^2 = I_X^2$, since by definition $I_X = \sigma_X/\mu_X$. The result follows by taking $X = C_{max}$.