



The Spatial Variation of the Maximum Possible Contaminant Concentration From a Steady Line Source

Nils Mole¹ · Rick Munro²

Received: 22 May 2018 / Accepted: 15 January 2019
© The Author(s) 2019

Abstract

When considering the possible hazard or nuisance associated with a release of toxic or malodorous gas into the atmosphere, large concentrations are especially important, but relatively little work has been done on measuring or modelling the probability distribution of large concentrations of a contaminant dispersing in a turbulent flow. We have previously applied statistical extreme-value theory to field experiment measurements, analyzing large concentrations at a handful of positions. We have also been involved in using a related moment-based method to give more comprehensive spatial coverage for a steady line source in a wind tunnel. In the latter case, however, we used a method which would be expected to have shortcomings, and we did not calculate confidence intervals for the estimates. In the present paper we use an improved method, with confidence intervals derived from bootstrapping. Again we analyze the measurements from steady line-source wind-tunnel experiments, with particular emphasis on the spatial variation of the estimated maximum possible concentration. We show that the moment-based method agrees well with maximum-likelihood fitting to exceedances of a high threshold. We find that the centreline maximum concentration, normalized by the centreline mean concentration, increases downwind from a value just greater than 1 near the source, to a peak value of 5–6, before decreasing with distance from the source. Across the plume the maximum concentration only varies slowly for the downwind distances in these experiments. These observations are explained in terms of the physical processes of turbulent advection and molecular diffusion.

Keywords Atmospheric dispersion · Generalized Pareto distribution · Maximum concentration · Molecular diffusion

✉ Nils Mole
n.mole@sheffield.ac.uk

¹ School of Mathematics and Statistics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK

² Faculty of Engineering, University of Nottingham, Coates Building, University Park, Nottingham NG7 2RD, UK

1 Introduction

Atmospheric boundary-layer flows are turbulent, i.e. random, so the concentrations of any pollutant gases released into such flows are also random. These concentrations need to be analyzed in terms of their probability distributions. For releases of toxic or malodorous gases, large concentrations are of particular importance for the assessment of hazards or nuisance. In this paper attention will be focussed on large concentrations, so it is the high-concentration tails of the probability distributions which are of interest.

The most appropriate framework for high concentrations is provided by statistical extreme-value theory, for which we outline the relevant aspects below. Let $P(\theta; \theta_T)$ be the distribution function of the concentration Γ , conditional on Γ being above a threshold θ_T , i.e.

$$P(\theta; \theta_T) = \text{Prob}(\Gamma \leq \theta | \Gamma > \theta_T). \tag{1}$$

The corresponding probability density function (p.d.f.) is

$$p(\theta; \theta_T) = \frac{d}{d\theta} P(\theta; \theta_T). \tag{2}$$

Subject to some regularity conditions, Pickands (1975) showed that for large threshold θ_T

$$p(\theta; \theta_T) \approx g(\theta - \theta_T; k, a) \quad \text{for } \theta > \theta_T, \tag{3}$$

where $g(s; k, a)$ is the p.d.f. of the generalized Pareto distribution, given by

$$g(s; k, a) = \frac{1}{a} \left(1 - \frac{ks}{a} \right)^{\frac{1}{k}-1}. \tag{4}$$

Sections 4.1 and 4.2 of Coles (2001) deal with this. Here k is the shape parameter and $a (> 0)$ is the scale parameter.

This result has previously been applied to the statistical modelling of extreme values by fitting, usually by maximum likelihood, to exceedances of a high threshold. Examples of this approach are given by Davison and Smith (1990), Leadbetter (1991) and, for turbulent dispersion, Mole et al. (1995), Anderson et al. (1997), Schopflochler (2001) and Munro et al. (2001). For pollutant concentration the maximum possible value, θ_{MAX} , is finite (bounded above by the largest source concentration). Thus we expect that $k > 0$, so that $g(s; k, a)$ has a finite upper endpoint a/k . We then have

$$\theta_{MAX} = \theta_T + \frac{a}{k}. \tag{5}$$

Figure 1 of Mole et al. (2008) shows possible shapes for $g(s; k, a)$. If we let θ_T vary, while keeping it large so that (3) still applies, then for two large values θ_{T1} and θ_{T2} of θ_T , with $\theta_{T2} > \theta_{T1}$, we have

$$p(\theta; \theta_{T2}) = \frac{p(\theta; \theta_{T1})}{\int_{\theta_{T2}}^{\theta_{MAX}} p(\theta; \theta_{T1}) d\theta}. \tag{6}$$

If the parameters corresponding to θ_{T1} and θ_{T2} are (k_1, a_1) and (k_2, a_2) respectively, then (see Appendix 1 for details) this can only be satisfied if $k_2 = k_1 (= k \text{ say})$ and

$$a_2 - a_1 = -k(\theta_{T2} - \theta_{T1}). \tag{7}$$

Thus, as θ_T varies, the form of the asymptotic distribution $g(s; k, a)$ and the value of the shape parameter k cannot change, but the scale parameter a must decrease as θ_T increases. The result for a can also be obtained from (5).

Ideally a physical model would give the concentration p.d.f., with a finite upper endpoint for the concentration. Thus this would include probabilities of the exceedances of high thresholds. In practice, however, this is difficult, and models are much more likely to give the first few concentration moments. Motivated by this, Mole et al. (2008) introduced a different method for estimating properties of the distribution for large concentrations. The central argument there was that high-order moments m_n are dominated by the largest concentration values, so that

$$m_n = E\{\Gamma^n\} \approx \eta \int_0^{\theta_{MAX}} \theta^n g(\theta; k, a) d\theta \tag{8}$$

for some constant (at a fixed point in space) η . (Without loss of generality θ_T was taken to be zero—otherwise a can simply be rescaled.) Rather more detail, including the spatial variation of η , is given in Sects. 2.2 and 2.4 of Mole et al. (2008), who showed that this implies that, for sufficiently large n ,

$$\frac{m_{n-1}}{m_n} \approx \frac{1}{a} \left(\frac{1}{n}\right) + \frac{k}{a} = \frac{1}{a} \left(\frac{1}{n}\right) + \frac{1}{\theta_{MAX}}. \tag{9}$$

The linear relationship between the ratio of successive moments, m_{n-1}/m_n , and $1/n$ then allows the estimation of the parameters k , a and θ_{MAX} . Mole et al. (2008) used this method to estimate the parameters from a theoretical model for the concentration moments, and from the experimental data of Sawford and Tivendale (1992), for a steady line-source release in a wind tunnel.

The concentration model in Mole et al. (2008) gave all moments, in terms of some parameters which would have to be modelled. Other modelling approaches to calculating moments include meandering plume (e.g., Gifford 1959; Luhar et al. 2000; and Cassiani and Giostra 2002), large-eddy simulation (e.g., Xie et al. 2004) and p.d.f. micromixing (e.g., Pope 1985; Sawford 2004). Most models would only give the lowest few moments, whereas the moment-based method described here would be expected to require higher moments. This is a problem that requires further investigation.

In the present paper we extend the work of Mole et al. (2008) by applying this moment-based method to the Sawford and Tivendale (1992) wind-tunnel data in a more thorough manner, including the estimation of confidence intervals for the parameter estimates. We also compare these results with those obtained by the more conventional method of maximum-likelihood fitting to exceedances of a high threshold. We have two main aims. The first is to explore the variation in space of properties of the distribution of high concentrations, in particular θ_{MAX} and k . The second is to establish how well the moment-based method performs, with applications to models in mind.

2 The Estimation Methods

2.1 The Moment-Based Method

Mole et al. (2008) fitted (9) by least squares, using the values $n = 4$ to $n = 8$. The straight-line fit was quite good, but there was a suggestion of curvature which would have altered the estimates if larger values of n had been used.

For a finite-size dataset, for large n the moments m_n become dominated by the largest measured value Γ_{MAX} . Thus, using very large n to fit (9) will result in θ_{MAX} being underestimated (as Γ_{MAX}). Conversely, at small n (9) would not be expected to hold. So in fitting (9)

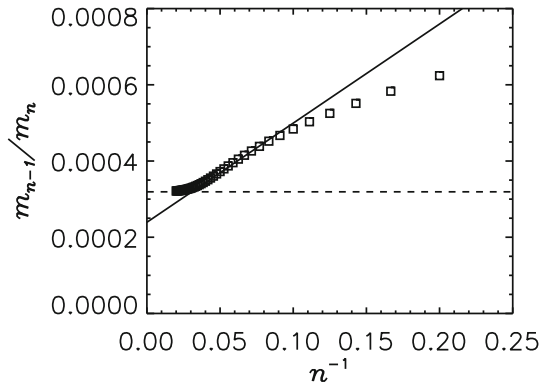


Fig. 1 The ratio of successive moments, m_{n-1}/m_n , against $1/n$. The squares are the values calculated from the data, and the solid line is the fit of (9) using the maximum gradient. The dashed line is $1/\Gamma_{MAX}$, where Γ_{MAX} is the largest measured concentration. This case is for the wind-tunnel data of Sawford and Tivendale (1992), on the mean-plume centreline, at a distance of $X = 100$ mm downwind of the source (i.e. $X/X_0 = 0.323$, where X_0 is the distance from the grid to the source)

to the data a balance must to be struck between using small and large values of n . We have carried out some preliminary analysis, on simulated datasets, which suggests that a reasonable choice is to fit the straight line through the point of maximum gradient when m_{n-1}/m_n is plotted against $1/n$. For typical cases this will give the largest estimate for θ_{MAX} that is possible from (9), and for hazard-assessment purposes will therefore tend to be on the conservative side. Figure 1 shows a typical plot of m_{n-1}/m_n against $1/n$, with the straight line fitted using the maximum gradient.

Confidence intervals for θ_{MAX} and k are calculated by bootstrapping. We use the studentized bootstrap method with a log transformation of the data (details of this method can be found in e.g. Davison and Hinkley 1997).

2.2 The Maximum-Likelihood Method Using Exceedances

With this method we fit the generalized Pareto distribution (4) to excesses above a high threshold, using maximum likelihood. The thresholds are chosen using mean-excess plots (i.e. mean-residual-life plots)—see e.g. Coles (2001, p. 79). Confidence intervals are estimated using profile likelihood (Davison and Smith 1990). To take account of possible dependence between concentration exceedances when estimating the confidence intervals, declustering (see, e.g., Coles 2001, p. 99) is carried out. This involves choosing a cluster separation time τ . If the time between successive exceedances is greater than τ then they are deemed to belong to separate clusters, and only the maximum concentration value from each cluster is used in the maximum-likelihood fitting. We choose τ using the method of Ferro and Segers (2003), which is based on estimating the extremal index. For details of the extremal index see Leadbetter et al. (1983, p. 67); it can be interpreted as the reciprocal of the mean cluster size (Leadbetter 1983).

In a few cases the profile-likelihood method does not converge, and instead we estimate confidence intervals by assuming that the likelihood estimator is approximately multivariate normal (see, e.g., Sect. 4.2.1 of Davison 2003, or Sect. 2.6.4 of Coles 2001). This happens far downwind or in the plume fringes, where there are fewer data on large concentration

values. In these cases the estimated confidence intervals will automatically be symmetric, whereas the real ones will probably not be. In particular, we would expect that the upper end of the confidence interval should be larger than that indicated by the multivariate-normal approximation.

3 Results

The two estimation methods are applied to the experimental data of Sawford and Tivendale (1992), obtained from a steady line source in wind-tunnel grid turbulence, with the turbulence approximately uniform in the direction normal to the line source and the mean-flow direction. Further details of the experiments are given in Appendix 2 here, and also in Sawford and Sullivan (1995) and Mole et al. (2008). Concentration (which in this case is temperature) measurements were made at many downwind and crosswind positions.

The wind-tunnel boundaries were sufficiently distant that the temperature was effectively a conserved scalar. Close to the source the mean temperatures were up to 50 K above the background temperature, similar to the experiments of Stapountzis et al. (1986), but much higher than in Warhaft (1984). Therefore, near the source the scalar was not passive, because of the buoyancy. Sawford (2004) showed that the intensity of fluctuations was reduced by comparison with the Warhaft data, but that in the far field these source effects became unimportant. This is less clear-cut here since we are concerned with the largest concentrations. Turbulent advection means that, except in the far field, the largest concentrations decay less quickly downwind than the mean and standard deviation. This means that non-passive effects on the largest concentrations may well persist to larger distances from the source. However, further from the source large temperatures will be found in relatively small volumes, so we expect that non-passive effects will be small. The subsequent figures for cross-plume variations do not show obvious asymmetry resulting from buoyancy (note that the cross-plume direction is in the vertical). This is also the case in Fig. 1 of Sawford and Sullivan (1995) for the first four moments of concentration. So in subsequent attempts to explain the results physically we assume that a good first approximation is provided by taking the scalar to be passive.

Figure 2 shows the estimates of θ_{MAX} on the mean-plume centreline, as a function of non-dimensional downwind distance from the source, X/X_0 , where X_0 is the distance from the grid to the source. (Note that turbulent length scales will be of the order of the hole spacing in the grid, i.e. of order 0.08 when normalized by X_0 .) The value of θ_{MAX} is normalized by the centreline mean concentration μ_0 . Up to about $X/X_0 = 0.3$ the two methods agree very well. Beyond $X/X_0 = 0.3$ the two methods differ more, but the confidence intervals are larger, and the differences are probably not very significant. The quantity θ_{MAX}/μ_0 increases steadily from a value just over 1 at $X/X_0 \approx 0.01$, to a maximum of about 5–6 at $X/X_0 \approx 0.8$, and then decreases to about 3.5 at $X/X_0 \approx 8$.

The downwind variation of θ_{MAX}/μ_0 agrees qualitatively with the results of Mole et al. (2008) (in their Fig. 6), obtained using the moment ratios from $n = 4$ to $n = 8$, but the values here are roughly 20% larger. For the moment-based method the larger values here are to be expected, since we are fitting to the maximum gradient, but the agreement with the maximum-likelihood estimates suggests that the maximum-gradient method is superior to the method used by Mole et al. (2008).

Mole et al. (2008) argued that the variation of θ_{MAX}/μ_0 with downwind distance X shown in Fig. 2 is what would be expected on physical grounds. For a steady source, close to the

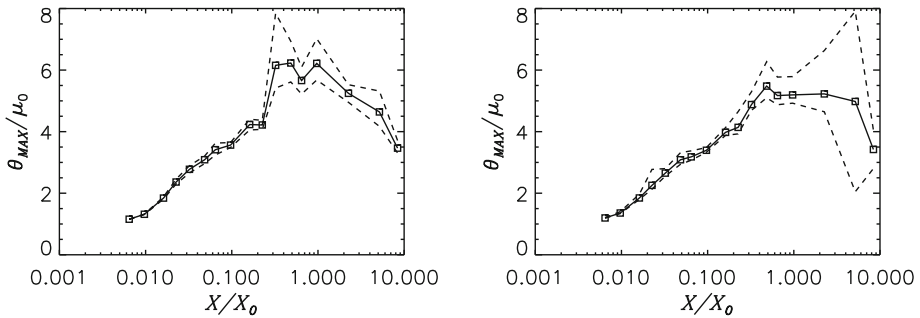


Fig. 2 All values here are on the mean-plume centreline. Left panel: the moment-based estimate of the maximum possible concentration θ_{MAX} , as a function of non-dimensional downwind distance from the source. θ_{MAX} is normalized by the centreline mean concentration μ_0 . The squares and solid line are the estimates of θ_{MAX}/μ_0 , and the dashed lines give the 95% confidence intervals, estimated using bootstrapping. Right panel: as left panel, but these are the maximum-likelihood estimates from exceedances of high thresholds. The confidence intervals (again 95%) are estimated using profile likelihood where possible

source we expect θ_{MAX} to be close to the largest source concentration θ_2 , and μ_0 to be close to the mean source concentration θ_1 . Thus, as $X \rightarrow 0$, we expect $\theta_{MAX}/\mu_0 \rightarrow \theta_2/\theta_1 \geq 1$. For most sources θ_2/θ_1 is likely to be close to 1.

For a passive conserved scalar the physical processes involved are advection by the turbulent velocity, and molecular diffusion (which is the only process which can change the concentration in a fluid element). The variation of θ_{MAX}/μ_0 as one goes away from the source is determined by the balance between the effects of these processes. Except very close to the source and very far downwind the mean concentration μ is hardly affected by molecular diffusion, and so is controlled by turbulent advection. The maximum concentration θ_{MAX} can only be altered through the action of molecular diffusion.

Most atmospheric releases have large Péclet number $Pe = ul/\kappa$ near the source, where u is a velocity scale for the turbulent fluctuations, l is a length scale for the turbulent concentration fluctuations, and κ is the molecular diffusivity. The ratio of the size of the advection and diffusion terms in the dispersion equation is Pe , so near the source advection acts much more quickly than diffusion. Further downwind, as the plume becomes stretched out into thin sheets and strands by turbulent advection, the concentration length scale l becomes smaller than near the source, and diffusion acts more quickly, as described by Batchelor (1959). The conduction cut-off length is defined as $(\nu\kappa^2/\epsilon)^{1/4}$, where ν is the kinematic viscosity and ϵ is the turbulent-energy dissipation rate per unit mass. The Schmidt number is defined as ν/κ . For pollutants with Schmidt number of order 1 or greater (which is usually the case, and is true for the experimental data used here) once the concentration length scale reaches the conduction cut-off length then a balance is reached between advection and diffusion, where their time scales are the same and the concentration length scale does not decrease any more.

Thus, provided the source size is large compared with the conduction cut-off length (which would usually be the case for atmospheric releases), near the source advection acts much faster than diffusion. This means that μ_0 is reduced much more quickly than θ_{MAX} , so θ_{MAX}/μ_0 increases away from the source. Further downwind the concentration length scale decreases, and since the largest concentrations are found in the thin sheets and strands, diffusion acts to reduce θ_{MAX} more quickly. Conversely, as the plume width increases downwind, advection reduces μ_0 more slowly. Eventually the rate of reduction of μ_0 becomes less than that of θ_{MAX} , so θ_{MAX}/μ_0 reaches a peak and then decreases with downwind distance.

In the grid turbulence found in the experiments discussed here, the turbulent energy (and hence the velocity scale u) and the dissipation rate ϵ decay with downwind distance from the source (see, e.g., Sect. 5.4.6 of Pope 2000). This implies that the Péclet number Pe will decrease faster downwind than in a non-decaying flow, and the conduction cut-off length will increase downwind. Using the values in Table I of Sawford (2004), in these experiments Pe at the source is of order 3 (in the atmosphere the source is likely to be larger, so Pe would be several orders of magnitude larger). It also gives a conduction cut-off length comparable with the source size (for the atmospheric applications the conduction cut-off length would probably be similar, but the source size considerably greater).

However, at the source the turbulence length scale in these experiments is of order 70 times the source size. So up to moderate distances downwind from the source the instantaneous plume is swept from side to side by the turbulence on length scales greater than the width of the instantaneous plume, thus increasing the mean plume width L and decreasing μ_0 downwind. Very close to the source (up to $X/X_0 \approx 0.01$ in these experiments) L is increased more by molecular diffusion than by turbulent advection, but at larger downwind distances its increase is dominated by turbulent advection. Reasonably near the source the instantaneous plume is broadened mainly by molecular diffusion, and not by turbulence (which mostly exists at length scales greater than that of the instantaneous plume width). Thus θ_{MAX} (which is found in the instantaneous plume) decreases more slowly than μ_0 with increasing distance from the source. Further downwind L increases, and turbulent advection reduces μ_0 more slowly. Far enough downwind we expect θ_{MAX}/μ_0 to reach a peak and then decrease. This behaviour of θ_{MAX}/μ_0 is what Fig. 2 shows. For most atmospheric releases we expect the source size to be rather greater than the conduction cut-off length, and as described above we expect similar behaviour of θ_{MAX}/μ_0 .

If there is an upper bound on the turbulent length scales for the velocity, as is the case for wind-tunnel grid turbulence, then far downwind the mean-plume width is much greater than the largest turbulent length scale for velocity and, as argued by Mole et al. (2008), we expect all concentrations to tend to the local mean. Thus $\theta_{MAX} \rightarrow \mu$ as $X \rightarrow \infty$, so on the centreline we expect $\theta_{MAX}/\mu_0 \rightarrow 1$. In practice, wind tunnels are not long and wide enough to approach this limit, and in field experiments one would expect non-stationarity and inhomogeneity to make this limit difficult to observe.

Figure 3 shows the moment-based and maximum-likelihood estimates of the shape parameter k . The moment-based estimates show a rapid decrease in values within $X/X_0 \approx 0.02$ of the source. Beyond this there is no obvious pattern, with values generally between 0.1 and 0.2. The maximum-likelihood estimates have no obvious pattern throughout, with values between 0.1 and 0.3. The moment-based estimates give narrower confidence intervals, and often the estimates of k with the two methods differ by more than the estimated confidence intervals. It appears that there is a systematic difference between the two types of estimator.

The suggestion is that k does not vary much with downwind distance, except perhaps close to the source, with values of order 0.2 or a little less. This is different from the results of Mole et al. (2008), where k was roughly equal to the reciprocal of θ_{MAX}/μ_0 , decreasing to about 0.2 at $X/X_0 \approx 0.8$, before increasing again with downwind distance. At most distances, the suggestion here is that the shape of the generalized Pareto distribution (4) is as in Fig. 1a of Mole et al. (2008), with zero gradient at the upper endpoint. The one case giving an estimate of $k > 0.5$, corresponding to infinite gradient at the upper endpoint, is for the moment-based estimate at the position closest to the source.

Theoretically, close to the source we expect most of the weight of the p.d.f. of concentration to be close to the source concentration values, and away from zero. For a uniform source with concentration θ_0 , this p.d.f. would be $\delta(\theta - \theta_0)$ at the source, and for a non-uniform source

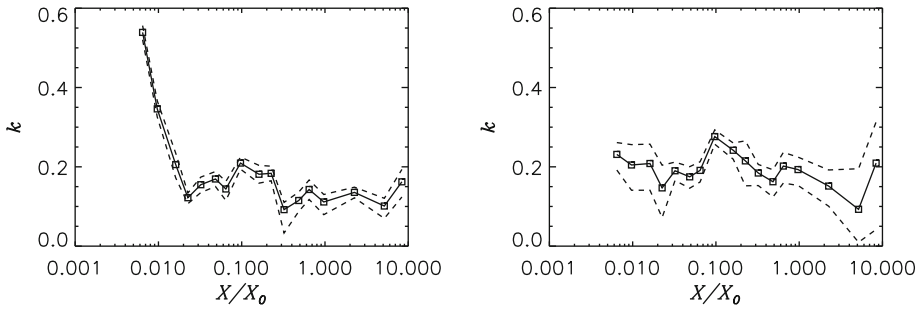


Fig. 3 All values here are on the mean-plume centreline. Left panel: the moment-based estimate of k , as a function of non-dimensional downwind distance from the source. The squares and solid line are the estimates of k , and the dashed lines give the 95% confidence intervals, estimated using bootstrapping. Right panel: as left panel, but these are the maximum-likelihood estimates. The confidence intervals (again 95%) are estimated using profile likelihood where possible

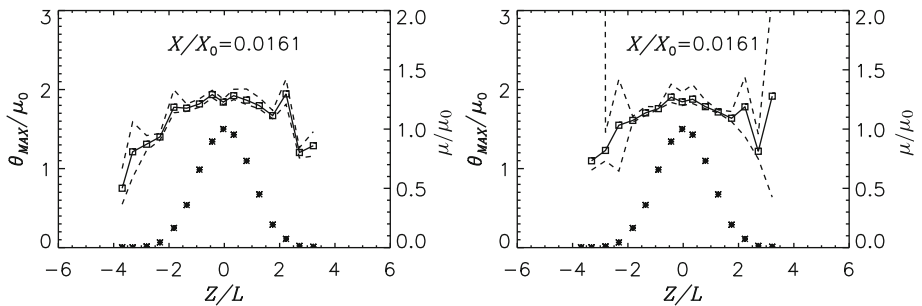


Fig. 4 Left panel: the moment-based estimate of θ_{MAX}/μ_0 , at non-dimensional downwind distance $X/X_0 = 0.0161$, as a function of crosswind distance Z from the mean-plume centreline, normalized by the mean-plume width L . The squares and solid line are the estimates of θ_{MAX}/μ_0 , and the dashed lines give the 95% confidence intervals, estimated using bootstrapping. The asterisks show the mean concentration μ , estimated by the centreline value μ_0 . Right panel: as left panel, but these are the maximum-likelihood estimates. The confidence intervals (again 95%) are estimated using profile likelihood where possible

we expect the p.d.f. to be broadened slightly from this. For (4) to give a peak away from $\theta = 0$ requires $k > 1$, so very close to the source we would expect to have large values of k . Far downwind we expect the concentration p.d.f. to tend to $\delta(\theta - \mu)$, so the same argument for k holds. From these arguments we would anticipate that k would be large close to the source, decrease to a minimum, and then increase again very far downwind. Figure 3 does not show any evidence for the latter, which may just be because we do not have measurements sufficiently far downwind. The moment-based estimates in Fig. 3 do show evidence for larger values of k near the source, but since this is not supported by the maximum-likelihood estimates, it is not clear how reliable this is. The moment-based estimates in Mole et al. (2008) do show larger values of k at small and large downwind distances, but the suggestion from Fig. 2 is that the present results are more reliable.

Figures 4, 5 and 6 show the variation of θ_{MAX}/μ_0 , as a function of crosswind distance Z (with $Z = 0$ on the centreline), at three downwind distances; Z is normalized by the mean-plume width L , defined as the standard deviation of the crosswind profile of mean concentration (which is close to Gaussian: the normalized mean concentration μ/μ_0 is also shown in the figures). Again, both moment-based and maximum-likelihood estimates are

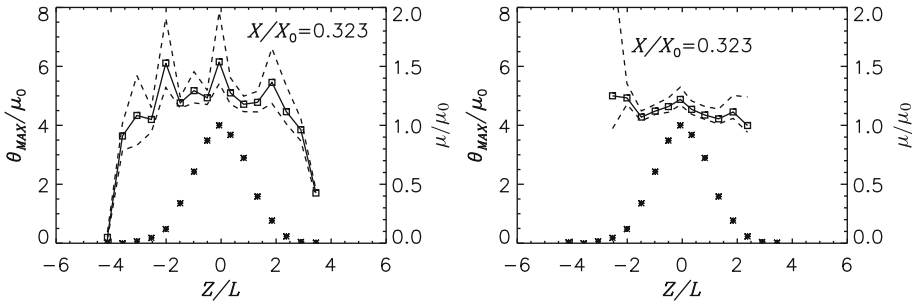


Fig. 5 As Fig. 4, but for non-dimensional downwind distance $X/X_0 = 0.323$

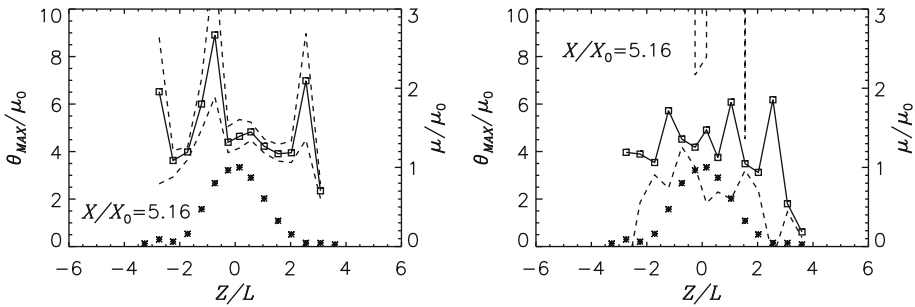


Fig. 6 As Fig. 4, but for non-dimensional downwind distance $X/X_0 = 5.16$

shown. The downwind distances are $X/X_0 = 0.0161$, close to the source, $X/X_0 = 0.323$, which is close to the maximum in θ_{MAX}/μ_0 , and $X/X_0 = 5.16$, which is far downwind. In some cases in the plume fringes convergence was not obtained, so points are not shown. In some other cases for the maximum-likelihood method, the profile-likelihood method for the confidence intervals did not converge, and confidence intervals estimated from normal approximations to the likelihood estimator are used instead. This is especially the case for $X/X_0 = 5.16$, for which the number of exceedances of the threshold θ_T tends to be smaller. For both estimators the width of the confidence intervals is quite variable across the plume, especially further from the source. Measurements at different positions were made in different runs, so variability between releases may perhaps explain this.

For $|Z/L| < 2 - 3$, the moment-based and maximum-likelihood methods give similar results for θ_{MAX}/μ_0 . At $X/X_0 = 0.0161$ the maximum-likelihood values tend to be slightly smaller, and at $X/X_0 = 5.16$ the uncertainty is greater than the difference in the estimated values.

At all downwind distances, for $|Z/L| \gtrsim 3$, θ_{MAX}/μ_0 appears to be fairly constant, allowing for the larger variability and larger confidence intervals at the far-downwind distance. Far from the centreline larger concentrations occur only rarely, so to obtain accurate estimates of their probability distribution will require longer time series than for positions close to the centreline. This is reflected in the lack of convergence in some cases, and in the generally larger confidence intervals far from the centreline, especially for the maximum-likelihood method.

The crosswind variation of μ/μ_0 , and the value of L , is determined by turbulent advection. Very far from the centreline the positions are much further from the source than are those positions near the centreline, so diffusion has had longer to act. Thus we expect $\theta_{MAX} \rightarrow 0$ as

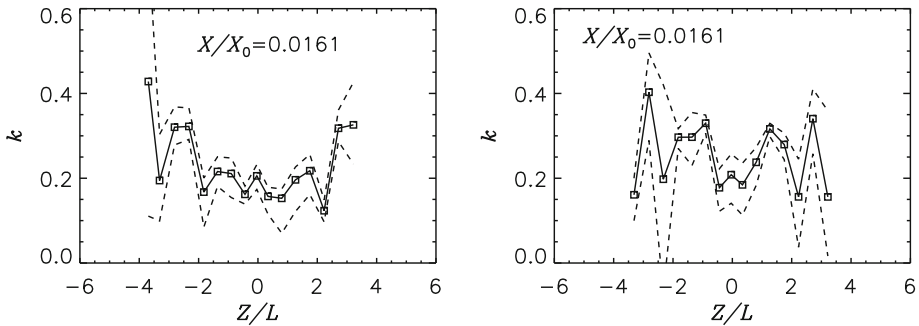


Fig. 7 Left panel: the moment-based estimate of k , at non-dimensional downwind distance $X/X_0 = 0.0161$, as a function of Z/L . The squares and solid line are the estimates of k , and the dashed lines give the 95% confidence intervals, estimated using bootstrapping. Right panel: as left panel, but these are the maximum-likelihood estimates. The confidence intervals (again 95%) are estimated using profile likelihood where possible

$|Z/L| \rightarrow \infty$. At small values of X , where advection reduces μ_0 more quickly than diffusion reduces θ_{MAX} , we expect that the decrease of θ_{MAX} away from the mean-plume centreline will be on a larger length scale than L , so we expect θ_{MAX}/μ_0 to be fairly constant for $|Z/L| < 2$ or 3, as seen in Fig. 4.

At values of X close to the downwind peak in θ_{MAX}/μ_0 , we still expect this to be true because it depends on the cumulative effect while the fluid elements travel from the source to these downwind distances. So this argument is in agreement with the observations in Fig. 5.

Very far downwind we expect that $\theta_{MAX} \rightarrow \mu$, so $\theta_{MAX}/\mu_0 \rightarrow \mu/\mu_0$, and θ_{MAX}/μ_0 will decrease on the same length scale L as μ/μ_0 . Figure 6 suggests that the wind-tunnel measurements do not extend far enough downwind to identify this regime.

Figures 7, 8 and 9 show the crosswind variation of the moment-based and maximum-likelihood estimates of k . There is no clear pattern to the results, but the maximum-likelihood values tend to be a little larger than the moment-based ones. Far from the centreline we believe the results are less reliable, for the reasons discussed above for θ_{MAX}/μ_0 . The only cases where the estimates of k are greater than 0.5 are for the moment-based method, in the plume fringes at $X/X_0 = 0.323$. Here we would expect $k < 0.5$, so we believe these results are a reflection of the lack of reliability. The evidence seems to be that k only varies slowly across the plume. This is what we would expect from the physical arguments given above, except very far downwind, since k is determined by the distribution of large concentrations, which are mainly affected by molecular diffusion.

In about 80% of the cases shown, the moment-based estimate of θ_{MAX} is larger than that obtained using maximum likelihood. There also appears to be a tendency for the moment-based confidence intervals to be narrower than those found with maximum likelihood. The suggestion is that there is some systematic difference between the two types of estimator. For applications to toxic or malodorous gases this difference means that the moment-based estimates of θ_{MAX} are likely to be conservative.

4 Discussion

We have used two methods for estimating the maximum possible concentration θ_{MAX} and the shape parameter k of the distribution of large concentrations. One method is based on the expected behaviour of high-order concentration moments, and the other fits the generalized

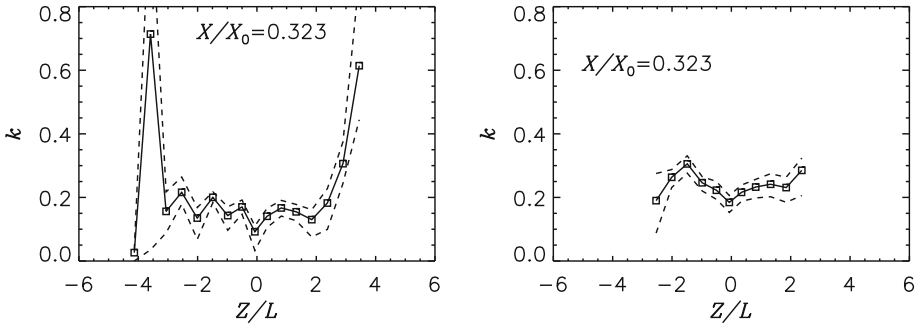


Fig. 8 As Fig. 7, but for non-dimensional downwind distance $X/X_0 = 0.323$

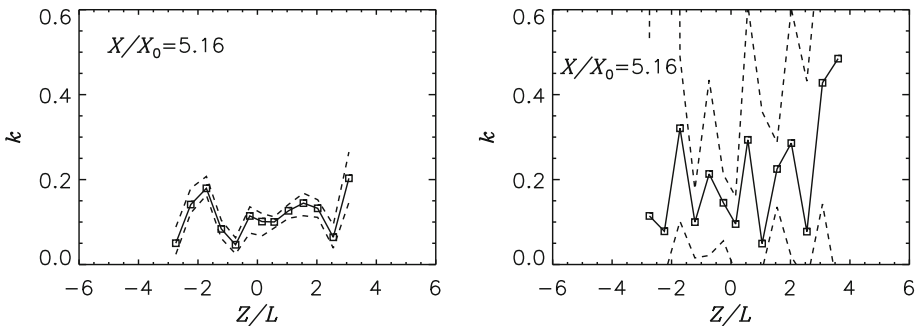


Fig. 9 As Fig. 7, but for non-dimensional downwind distance $X/X_0 = 5.16$

Pareto distribution (4) directly to concentrations above a high threshold, using maximum likelihood. The two methods agree reasonably well for all the cases shown, and are in very close agreement for θ_{MAX} on the mean-plume centreline up to $X/X_0 \approx 0.3$. The moment-based method used here appears to give better results than that used in Mole et al. (2008).

The results show that on the centreline θ_{MAX}/μ_0 increases from a value slightly larger than 1 very near the source, to a value of 5–6 at $X/X_0 \approx 0.8$, before decreasing again. In the crosswind direction both θ_{MAX}/μ_0 and k vary much more slowly than μ/μ_0 , as we would expect if they are affected much more by molecular diffusion than by turbulent advection. Very far downwind we would expect diffusion to dominate if there is an upper bound on the velocity length scales. In this case θ_{MAX}/μ_0 and k would vary on the same scale as μ/μ_0 . The results suggest that the measurements do not extend far enough downwind to reach this regime.

These results provide encouragement to proceed with attempts to develop quantitative models, in particular those based on concentration moments, for θ_{MAX}/μ_0 and for other properties of the distribution of large concentrations. Such models would then enable hazard assessment to be carried out for practical applications involving releases of toxic and malodorous gases in the atmosphere. Some discussion of possible modelling approaches was given in Mole et al. (2008). In general limits on the quality of results given by this method will be imposed by the ability of models or experiments to give accurate results for concentration moments. The more moments are required, the worse this problem is likely to be.

Grid turbulence provides an approximation to homogeneous isotropic turbulence, for which it is relatively easy to interpret results and develop models. However, in practical applications releases occur in inhomogeneous boundary-layer turbulence which, unlike grid turbulence, does not decay with downwind distance. An alternative would be to use wind-tunnel boundary-layer releases such as those analyzed by Xie et al. (2007).

To avoid the limitation on downwind distances imposed by the dimensions of a wind tunnel, and to consider real boundary-layer conditions, it would be desirable to repeat the analysis for field experiments. There are, however, some difficulties with attempting this. In addition to the usual problems with field experiments of non-stationary conditions and terrain effects, there is a specific difficulty relating to measuring large concentrations. Large concentrations are found at the smallest spatial scales present in the concentration field, so sensors with very good spatial resolution are needed to measure them. Such sensors will be ones which can only make measurements at a single point, so obtaining a wide spatial coverage, especially at positions where long time series are needed, is expensive. On the other hand, a measurement system such as lidar which gives wide spatial coverage cannot resolve sufficiently small scales to give reliable measurements of the largest concentrations. Perhaps this problem may be overcome by resolving at the scales of an averaging human breath.

Acknowledgements We would like to thank Paul Sullivan and Clive Anderson for many useful discussions over the years on these topics.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

We have

$$p(\theta; \theta_{T2}) = \frac{p(\theta; \theta_{T1})}{\int_{\theta_{T2}}^{\theta_{MAX}} p(\theta; \theta_{T1}) d\theta} \tag{10}$$

and, from (3) and (4),

$$p(\theta; \theta_T) \approx \frac{1}{a} \left[1 - \frac{k(\theta - \theta_T)}{a} \right]^{\frac{1}{k} - 1}. \tag{11}$$

If the parameters corresponding to θ_{T1} and θ_{T2} are (k_1, a_1) and (k_2, a_2) respectively, then

$$\frac{1}{a_2} \left\{ 1 - \frac{k_2(\theta - \theta_{T2})}{a_2} \right\}^{\frac{1}{k_2} - 1} = \frac{\frac{1}{a_1} \left\{ 1 - \frac{k_1(\theta - \theta_{T1})}{a_1} \right\}^{\frac{1}{k_1} - 1}}{\left[- \left\{ 1 - \frac{k_1(\theta - \theta_{T1})}{a_1} \right\}^{\frac{1}{k_1}} \right]_{\theta_{T2}}^{\theta_{MAX}}}$$

$$= \frac{\frac{1}{a_1} \left\{ 1 - \frac{k_1(\theta - \theta_{T1})}{a_1} \right\}^{\frac{1}{k_1}-1}}{\left\{ 1 - \frac{k_1(\theta_{T2} - \theta_{T1})}{a_1} \right\}^{\frac{1}{k_1}}} \tag{12}$$

using (5). Both sides only give the same function of θ if $k_1 = k_2$ ($= k$ say), in which case we have

$$\frac{1}{a_2} \left\{ 1 - \frac{k(\theta - \theta_{T2})}{a_2} \right\}^{\frac{1}{k}-1} = \frac{\frac{1}{a_1} \left\{ 1 - \frac{k(\theta - \theta_{T1})}{a_1} \right\}^{\frac{1}{k}-1}}{\left\{ 1 - \frac{k(\theta_{T2} - \theta_{T1})}{a_1} \right\}^{\frac{1}{k}}} \tag{13}$$

Rearranging this gives

$$a_2^{-\frac{1}{k}} \{a_1 - k(\theta_{T2} - \theta_{T1})\}^{\frac{1}{k}} = \left\{ \frac{a_1 - k(\theta - \theta_{T1})}{a_2 - k(\theta - \theta_{T2})} \right\}^{\frac{1}{k}-1} \tag{14}$$

The left-hand side is a constant, so the right-hand side must be too. To remove the θ dependence this constant must be 1, giving

$$a_2 - a_1 = -k(\theta_{T2} - \theta_{T1}) \tag{15}$$

This makes the left-hand side of (14) also equal to 1, so it is consistent.

Appendix 2

Sawford and Tivendale (1992) conducted experiments for a steady line source in wind-tunnel grid turbulence. The source was horizontal and across the wind tunnel, normal to the horizontal flow direction and in the middle of the cross-section. The main details are summarized here.

The experiments were carried out in a wind tunnel with a rectangular test section 0.69 m high, 1.07 m wide and 3.3 m long. A grid with circular holes of diameter 0.0208 m in a hexagonal pattern was used to generate turbulence. The hole spacing was $M = 2.54 \times 10^{-2}$ m, giving a solidity ratio of 0.39. The source was a heated Nichrome wire of diameter 0.213 mm placed a distance $X_0 = 12.2M$ downstream of the grid. The mean air speed was $U = 5.0 \text{ m s}^{-1}$, with a corresponding Reynolds number UM/ν of 8500, where ν is the kinematic viscosity of unheated air ($1.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ at 20 °C).

Temperatures (i.e. scalar concentrations here) were measured using a platinum cold wire which had a diameter of 1.27 μm and a length of 0.4 mm. In time the sampling was at 4096 Hz, with the signal low-pass filtered at 2 kHz. Statistics were calculated from 20 separate 1-s samples, i.e. from a total of 81,920 points. According to Sawford and Sullivan (1995) frequencies up to 1 kHz accounted for about 90% of the temperature variance near the source, and over 99% of the variance far downstream.

Chan (2009) dealt with possible shortcomings of the temperature measurements caused by the mounting of the cold wire (see, for example, Paranthoen et al. 1982 and Tsuji et al. 1992), and concluded that when the measurements were normalized by the centreline mean, as in the present paper, they were unlikely to cause a problem.

References

- Anderson CW, Mole N, Nadarajah S (1997) A switching Poisson process model for high concentrations in short-range atmospheric dispersion. *Atmos Environ* 31:813–824
- Batchelor GK (1959) Small-scale variation of convected quantities like temperature in turbulent fluid Part 1. General discussion and the case of small conductivity. *J Fluid Mech* 5:113–133
- Cassiani M, Giostra U (2002) A simple and fast model to compute concentration moments in a convective boundary layer. *Atmos Environ* 36:4717–4724
- Chan LT (2009) Fluctuations in concentration of dispersing hazardous gases. PhD thesis, University of Sheffield, Sheffield, UK
- Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London, 208 pp
- Davison AC (2003) Statistical models. Cambridge University Press, Cambridge, 726 pp
- Davison AC, Hinkley DV (1997) Bootstrap methods and their application. Cambridge University Press, Cambridge, 592 pp
- Davison AC, Smith RC (1990) Model for exceedances over high thresholds. *J R Stat Soc B* 52:393–442
- Ferro CAT, Segers J (2003) Inference for clusters of extreme values. *J R Stat Soc B* 65:545–556
- Gifford F (1959) Statistical properties of a fluctuating plume dispersion model. *Adv Geophys* 6:117–137
- Leadbetter MR (1983) Extremes and local dependence in stationary sequences. *Z Wahrsch Ver Geb* 65:291–306
- Leadbetter MR (1991) On a basis for ‘peaks over thresholds’ modelling. *Stat Probab Lett* 12:357–362
- Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and series. Springer, New York, 352 pp
- Luhar AK, Hibberd MF, Borgas MS (2000) A skewed meandering plume model for concentration statistics in the convective boundary layer. *Atmos Environ* 34:3599–3616
- Mole N, Anderson CW, Nadarajah S, Wright C (1995) A generalized Pareto distribution model for high concentrations in short-range atmospheric dispersion. *Environmetrics* 6:595–606
- Mole N, Schopflocher TP, Sullivan PJ (2008) High concentrations of a passive scalar in turbulent dispersion. *J Fluid Mech* 604:447–474
- Munro RJ, Chatwin PC, Mole N (2001) The high concentration tails of the probability density function of a dispersing scalar in the atmosphere. *Boundary-Layer Meteorol* 98:315–339
- Paranthen P, Petit C, Lecordier JC (1982) The effect of the thermal prong-wire interaction on the response of a cold wire in gaseous flows (air, argon and helium). *J Fluid Mech* 124:457–473
- Pickands J (1975) Statistical inference using extreme order statistics. *Ann Stat* 3:119–206
- Pope SB (1985) Pdf methods for turbulent reactive flows. *Prog Energy Combust Sci* 11:119–192
- Pope SB (2000) Turbulent flows. Cambridge University Press, Cambridge, 771 pp
- Sawford BL (2004) Micro-mixing modelling of scalar fluctuations for plumes in homogeneous turbulence. *Flow Turbul Combust* 72:133–160
- Sawford BL, Sullivan PJ (1995) A simple representation of a developing contaminant concentration field. *J Fluid Mech* 289:141–157
- Sawford BL, Tivendale CM (1992) Measurements of concentration statistics downstream of a line source in grid turbulence. In: 11th Australasian fluid mechanics conference. 14–18 December 1992, University of Tasmania, Hobart, Australia, pp 945–948
- Schopflocher TP (2001) An examination of the right-tail of the PDF of a diffusing scalar in a turbulent flow. *Environmetrics* 12:131–145
- Stapountzis H, Sawford BL, Hunt JCR, Britter RE (1986) Structure of the temperature field downwind of a line source in grid turbulence. *J Fluid Mech* 165:401–424
- Tsuji T, Nagano Y, Tagawa M (1992) Frequency response and instantaneous temperature profile of cold-wire sensors for fluid temperature fluctuation measurements. *Exp Fluids* 13:171–178
- Warhaft Z (1984) The interference of thermal fields from line sources in grid turbulence. *J Fluid Mech* 144:363–387
- Xie Z-T, Hayden P, Robins AG, Voke PR (2007) Modelling extreme concentrations from a source in a turbulent flow over a rough wall. *Atmos Environ* 41:3395–3406
- Xie Z-T, Hayden P, Voke PR, Robins AG (2004) Large-eddy simulation of dispersion: comparison between elevated and ground-level sources. *J Turbul* 5:article no. 031, 16 pp