



**MURDOCH RESEARCH REPOSITORY**

<http://researchrepository.murdoch.edu.au>

*This is the author's final version of the work, as accepted for publication following peer review but without the publisher's layout or pagination.*

**Donnelly, R.P. , Lyons, T.J. and Flassak, T. (2009) Evaluation of results of a numerical simulation of dispersion in an idealised urban area for emergency response modelling. Atmospheric Environment, 43 (29). pp. 4416-4423.**

<http://researchrepository.murdoch.edu.au/2018>

Copyright © 2009 Elsevier Ltd.  
It is posted here for your personal use. No further distribution is permitted.

# Accepted Manuscript

Title: Evaluation of results of a numerical simulation of dispersion in an idealised urban area for emergency response modelling

Authors: R.P. Donnelly, T.J. Lyons, T. Flassak

PII: S1352-2310(09)00454-3

DOI: [10.1016/j.atmosenv.2009.05.038](https://doi.org/10.1016/j.atmosenv.2009.05.038)

Reference: AEA 9081

To appear in: *Atmospheric Environment*

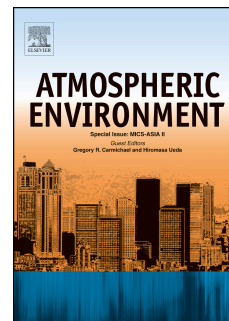
Received Date: 25 November 2008

Revised Date: 17 May 2009

Accepted Date: 18 May 2009

Please cite this article as: Donnelly, R.P., Lyons, T.J., Flassak, T. Evaluation of results of a numerical simulation of dispersion in an idealised urban area for emergency response modelling, *Atmospheric Environment* (2009), doi: [10.1016/j.atmosenv.2009.05.038](https://doi.org/10.1016/j.atmosenv.2009.05.038)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Evaluation of results of a numerical simulation of dispersion in an idealised urban area for emergency response modelling.

R.P. Donnelly<sup>a</sup>, T.J. Lyons<sup>a</sup>, T. Flassak<sup>b</sup>

<sup>a</sup>*School of Environmental Science, Murdoch University, Murdoch, WA 6150, Australia*

<sup>b</sup>*Ingenieurbuero Lohmeyer GmbH & Co. KG, D-76229 Karlsruhe, Germany*

---

## Abstract

WinMISKAM is evaluated from an emergency response perspective. Comparisons are made between ground level concentrations observed during selected Mock Urban Setting Test (MUST) field trials and predictions generated by the model. The model was driven by 5 minute averaged on-site meteorological data, and used minimum grid spacing of 0.5 m in both the horizontal and vertical. The code was found to perform well, with 46% of all predictions (paired in time and space) and 83% of arc maxima predictions within a factor of two of observed concentrations. The model was found to perform better for neutral cases than stable cases with 27% of stable case predictions and 57% of neutral case predictions within a factor of two when compared in time and space.

**Key words:** Urban dispersion; modelling; WinMISKAM; MUST

---

## 1. Introduction

During the last decades, with the growth in urban terrorist incidents, interest has grown in understanding and predicting pollutant pathways within built up urban areas (Milliez and Carissimo, 2007; Britter and Hanna, 2003). In such areas, traditional methods of predicting pollutant transport and diffusion (T&D), such as Gaussian plume modelling can fail due to critical assumptions not being met, e.g. complex morphology disrupting flows (Gailis and Hill, 2006). With advances in computing power, computational fluid dynamics (CFD) models can now resolve individual buildings and predict wind pathways through such complex terrain as an urban centre. Such models are increasingly being used to simulate the T&D of pollutants within urban areas, where the population is

---

*Email addresses:* R.Donnelly@murdoch.edu.au (R.P. Donnelly)

at risk (Milliez and Carissimo, 2007). If such models are to be used in emergency contexts, where success is measured in lives and health, it is of interest to evaluate model performance.

For urban dispersion problems, validation comes traditionally in the form of tracer release and capture studies. Whilst the optimum tracer study for validation of an urban T&D model occurs in a real urban centre within the planetary boundary layer (PBL), cost and difficulty have limited the number of such full scale investigations to a handful (Grimmond, 2006; Batchyarova and Gryning, 2006). In the same way that homogeneous terrain studies have invaluable aided understanding of more complex situations (Fernando et al., 2001), so too does the study of flows within a stylised area aid in understanding flow in more complex geometries, such as a real urban area. The Mock Urban Setting Test (MUST) (Biltoft, 2001b) investigations provide a simplified or stylised urban area, allowing the investigation of the underlying physics involved in urban T&D, without the overwhelming complexities observed in a real urban area (Biltoft, 2001b). The data also capture T&D under a number of different atmospheric conditions, along with a wealth of meteorological data, resulting in an ideal dataset for the validation of urban dispersion models.

From an emergency response perspective, the MUST dataset provides a unique opportunity to investigate the performance of dispersion models, with a view to determining the best response options available to emergency responders in the case of accidental or deliberate airborne releases within urban areas. With the growth of computing power showing no signs of abating, CFD models will soon be able to be run in much shorter timeframes, giving responders access to hitherto unknown information with which to make crucial decisions regarding the preservation of human life and health within highly populated urban areas.

The MUST dataset has been used extensively in the formulation and validation of digital models (e.g., Brook et al., 2003; Milliez and Carissimo, 2007; Santiago and Martilli, 2007; Goricsan et al., 2007). While Goricsan et al. (2007) compared the windfields generated by WinMISKAM to those measured on site during the MUST campaign, and Milliez and Carissimo (2007) analysed the concentration predictions of the CFD code *Mercure\_Saturne* (Archambeau et al., 2004) against MUST observations, no-one has yet compared concentration predictions by WinMISKAM against the MUST data.

This paper investigates the performance of WinMISKAM (Lohmeyer et al., 2002a) when run against selected MUST trials. The MUST experiment is outlined in Section 2 and the WinMISKAM model is described briefly in Section 3.1 with the simulation details outlined in Section 3.2. The statistical procedures used to evaluate the model's performance are based on Chang and Hanna (2004) and are outlined in Section 4. Section 5 presents results from the statistical comparison between observed ground level concentrations ( $C_{ground}$ ) of the passive tracer propylene taken in the MUST investigation and WinMISKAM predictions. Concluding remarks in Section 6 outline possible applications of the model for emergency response situations, both now and in the future.

## 2. The MUST Investigation

The MUST trials were designed to test the effects of an array of roughness elements (buildings) on the flow and dispersion of pollutants within an idealised urban morphology under a range of atmospheric stabilities (Biltoft, 2001a). MUST set up and details of all trials conducted are fully described in Biltoft (2001b). A brief description of parameters important to this study is given below.

120 shipping containers (each 12.2 m x 2.42 m x 2.54 m) were placed in a regular formation of 10 lines of 12 containers forming an approximately 200 m x 200 m square array. Meteorological data was sampled at a number of locations, including four 6 m towers which were distributed within the array, one in each quadrant, each holding two 3-D sonic anemometers (sonics), one at 4 m and the other at 6 m (see Fig.1). Propylene was used as the tracer and 40 photoionisation detectors (digiPIDs) were arranged in four sampling ‘lines’ or ‘arcs’ approximately 25, 60, 95 and 125 metres from release locations; these were placed at a height of 1.6 m above ground level (AGL). The sampling calibration range of these detectors was 0.04 ppm(v) - 1000ppm(v) (Biltoft, 2001b).

## 3. The WinMISKAM Model

### 3.1. Model Description

MISKAM is a 3-D non-hydrostatic flow and dispersion model for the prediction of the T&D of passive scalars within complex geometric environments, such as those found in built up urban areas (Eichhorn et al., 1988). The model solves the Reynolds-averaged Navier-Stokes equations with a modified  $k-\epsilon$  closure scheme in a non-uniform Cartesian grid and uses the Eulerian dispersion equation to calculate the concentration of scalars across that grid (Eichhorn, 2008). With the Windows interface designed by Lohmeyer Consulting Engineers, WinMISKAM can upload morphology directly from ArcView shape files and requires a single wind vector, along with a vertical potential temperature gradient and information on surface roughness values, to generate a windfield. The model has undergone extensive validation (e.g. Schatzmann and Leitl, 2002; Lohmeyer et al., 2002b; Dixon et al., 2006; Goricsan et al., 2007; Eichhorn and Kniffka, 2007). Version 2.1.2 (with MISKAM 5.02) of WinMISKAM has been used in this study.

Goricsan et al. (2007) compared velocity profiles from the full scale MUST and the wind tunnel simulation performed by Bezpalcova and Harms (2005) to those profiles predicted by two CFD models, MISKAM 5.02 and FLUENT 6.3.26 for a case where the wind blew perpendicularly onto the long sides of the containers. The WinMISKAM model performed well for horizontal windspeeds and directions, with only slight underpredictions near the surface and slight overpredictions at higher altitudes.

Although WinMISKAM has a vertical potential temperature stability setting, the model is designed primarily for use in urban areas where neutral stability dominates (Lohmeyer et al., 2002a). This assumption of urban neutrality

can be seen to be valid in a range of studies showing that strong diurnal changes in stability are absent in urban areas (Oke, 1987) and slightly unstable to neutral stability conditions are the norm (Bowne and Ball, 1970; Yersel and Goble, 1986; Kahn and Simpson, 1997; Britter and Hanna, 2003; Salmond et al., 2005; Hanna et al., 2006; Harman and Belcher, 2006; Lundquist and Chan, 2007).

Taking one wind vector as model input, along with roughness and stability information, WinMISKAM assumes a logarithmic wind profile at inflow boundaries. This is consistent with other windfield and dispersion models which do not require numerical weather prediction windfield input such as CAMEO/ALOHA (EPA, 1999), Hotspot (Homann, 1994) and A2C (Yamada, 2004). In this study a dynamical aspect is included by updating the windfield every 5 minutes and dispersing a theoretical tracer in this changing windfield. The tracer release has been modelled at a height in accordance with the release height of the individual runs, which varied between 0.15 m and 5.2 m. Information on release heights are presented in Table 1.

### 3.2. Numerical Simulations

The WinMISKAM grid used to model the MUST dispersion contained 400 x 400 x 30 cells, representing 314 x 300 x 130 m, with 50% of grids set to the minimum horizontal grid spacing of 0.5 m and a maximum horizontal grid spacing of 2 m. The minimum vertical grid spacing was 0.5 m, and the maximum at the top of the computational domain was 21.6 m; the lowest 10 cells were set to 0.5 m, with a stretching factor of 1.21 (the maximum recommended by (COST, 2007)) above this level.

Modelled containers (each 12.2 m x 2.42 m x 2.54 m) required between 23 x 5 x 5 grid cells at 0.5 m horizontal spacing and 12 x 2 x 5 cells at 1 m horizontal spacing. All containers were modelled within the area of the grid with 1 m horizontal spacing or lower. The containers used in the field trials were aligned slightly imperfectly to normal. The grid mimics this as closely as possible, with all modelled containers aligned parallel to the numerical grid, requiring no 'steps' along any sides. Minimum vertical grid height was set to 50 cm, allowing a maximum domain height of 130 m, or 50 times the height of the containers (H). Ketzel et al. (2000) showed that results from WinMISKAM V3.6 were more accurate when the height of the domain far exceeded the height of the obstacles (in their case increasing domain height from 5H to 25H times improved results substantially).

Twenty-one 200 s quasi-steady state periods have been identified by Yee and Bilitoft (2004); these have been selected for their remarkably low lateral velocity standard deviations. The lateral velocity fluctuations were found to increase gradually with sampling time, and large scale 2-D motions can mask the effect of the array on turbulent diffusion over longer sampling times (Yee and Bilitoft, 2004). Thus the 200 s sampling times showed the most useful steady state periods useful for model intercomparison. Twenty of these time averaged concentration files have been used by Milliez and Carissimo (2007) for comparison with the Mercure\_Saturne model and 19 of these have been selected

and modelled herein so that comparisons can be made on the performance of the models on (almost) the same dataset (Table 1).

Warner et al. (2006) evaluated the HPAC code (DTRA, 2001) against the MUST dataset, utilising 5 minute average meteorological data from a number of on-site locations. They showed that data from the sonic located within the array, 16 m above the canopy, resulted in the best predictions.

Data from the four 6 m sonics was used in this study to capture a more representative flow across the array than one single measurement could. Meteorological data from the four 6 m sonics were vectorially averaged to generate the single wind vector and vertical potential temperature gradient required for model input. It should be noted that this data, taken at  $z \approx 2.4H$ , could be influenced by the obstacle flow deviations. Coceal et al. (2006) found that the flow over a group of cubic obstacles was much less inhomogeneous above  $2H$  than below, thus the flow deviations caused by the MUST array at 6 m would not be likely to be sufficiently large so as to discount the improvement in accuracy obtained by the use of the vectorially averaged wind data.

There were cases where the data from each sonic was significantly different, such as for run 2672150 where the three sonics available (A B and D) reported 6 m wind directions of 109, 119 and 124 degrees respectively. Thus, even a ‘perfect’ model would not result in ‘perfect’ statistics, in part due to the stochastic nature of the atmosphere (Hanna et al., 1998), but also in part due to these meteorological input uncertainties.

The Monin-Obukhov length ( $L$ ) was also made available for each 5 minute period. This was used to estimate the vertical potential temperature gradient by first assigning a Pasquill-Gifford (PG) stability class (Golder, 1972) and then assigning a vertical potential temperature gradient (Mohan and Siddiqui, 1998).

Each of these 5 minute averaged meteorological datasets were used to generate a steady state windfield. Tracer releases were modelled at heights in accordance with release heights, which varied from release to release (Table 1). Consecutive 1 minute dispersion runs were conducted for the entirety of each release. Each of these dispersion runs (after the first) took output from the previous minute’s dispersion as input concentration data. The 200 s observations were compared with model output averaged over the corresponding three minutes. Only ground level detectors arranged in the four lines or arcs have been investigated. With 40 samplers for each of the 19 runs, a total of 760 data points are presented for comparison in Section 5.

#### 4. Statistical Methods Used in Evaluation

Model estimates, although acknowledged by the modelling community as not being ‘absolute answers’, will be taken as exactly that by emergency responders. Even if the modeller understands these limitations and passes this information on to the actual responders, they will in all likelihood not understand or know how to deal with this level of uncertainty (NOAA, 2004). With this in mind, when examining model performance, it should be noted that maximum concen-

trations *and* time averaged concentrations within the area of interest are equally important (NOAA, 2004).

It is important, however, for scientists involved in the evaluation of such predictive models to divorce themselves from this philosophy, and appreciate that model output is a prediction of just one of a range of probable outcomes, in the same way that a Gaussian plume is an ensemble of all the instantaneous plumes which many exist over a long enough timeframe. This understanding can be manifested by appreciating the nature of input errors in atmospheric modelling, such as model limitations and errors associated with predicting an essentially chaotic flow (Hanna et al., 1998). The expectation of perfect agreement between model predictions and observations is not realistic (Chang and Hanna, 2004). By examining a range of statistics evaluating the performance of the model against standardised and idealised datasets however, the underlying physics can be evaluated, model intercomparisons made and areas for model improvement identified.

Data from the current study has been analysed and presented in two ways; (a) Point to point, where all predictions were compared with the observation at that point in space and time and (b) Arcmax - whereby only the maximum modelled and observed concentration on each arc for each averaging period was compared, regardless of its location along that arc. Chang and Hanna (2004) suggest that for emergency response-oriented evaluation, analysis should focus on paired in time and space comparison, as the location of a plume, as well as its concentration, is of importance. Many authors have utilised this approach to analyse model performance against the MUST data (e.g. Warner et al., 2006; Milliez and Carissimo, 2007) while other authors have used the Arcmax approach for the MUST data (Hanna et al., 2004a, 2006, 2007, 2008), and for full scale urban tracer data, e.g. Madison Square Garden (Hanna et al., 2004b) and JointUrban2003 (Allwine et al., 2004). Both point to point and Arcmax analyses have been presented here to aid in inter-model comparison.

Chang and Hanna (2004) state that of the metrics suggested by them for model evaluation, all have advantages and disadvantages, thus an array of metrics should be used to glean a comprehensive evaluation of model performance. Of these metrics, some, (e.g. fractional bias and normalised mean square error) are linear, leading to the possibility of infrequently occurring values exerting undue influence on the measure. Others (e.g. mean geometric bias and mean geometric variance) are logarithmic and deal better with data spanning many orders of magnitude. Factor of two metric can be seen to be the most robust, as it is least affected by outliers (Chang and Hanna, 2004), but becomes more difficult to achieve as observed values become smaller.

As the logarithmic metrics return undefined results for zeros, Chang and Hanna (2004) suggest removing these values from the dataset by setting all values below the detection limit of the samplers to that detection limit. For the purposes of emergency modelling, false negatives, where models predict zero concentration while observations show concentrations above detection limits, are of the most interest as these may lead to health impacts or loss of life. False positives are also of importance, as these can lead to poor personnel deployment



in an emergency, where resources are scarce. Points of agreement between model and observations where both are below the detection limit are clearly of lesser importance. For this reason, all pairs where one of a pair (e.g. observation) shows tracer concentrations below 0.04 ppm(v) and the other (e.g. prediction) is greater than 0.04 ppm(v), the lower value has been set to the detection limit. Where both agree that concentrations are below the detection limit, both have been set to zero and removed from the dataset. This removes the bias associated with setting many points to the same values and can be seen to be the most rigorous way to deal with paired in time and space comparison. This has only been required for point to point comparison, as maximum observations and predictions for each trial were above this detection limit.

The BOOT statistical examination software (Chang and Hanna, 2004) was used to analyse the data, where the fractional bias ( $FB$ ), the mean geometric bias ( $MG$ ), the normalised mean square error ( $NMSE$ ), the geometric variance ( $VG$ ), and the percentage of data within a factor of two of observations ( $FAC2$ ) are as outlined in Chang and Hanna (2004). A factor of 5 test was also conducted to aid comparison with other model evaluations. These are defined as

$$FB = \frac{(\overline{C}_o - \overline{C}_p)}{0.5(\overline{C}_o + \overline{C}_p)}, \quad (1)$$

$$MG = \exp(\overline{\ln C}_o - \overline{\ln C}_p), \quad (2)$$

$$NMSE = \frac{(\overline{C}_o - \overline{C}_p)^2}{\overline{C}_o \overline{C}_p}, \quad (3)$$

$$VG = \exp \left[ \overline{(\ln C_o - \ln C_p)^2} \right], \quad (4)$$

The FAC2 metric refers to the fraction of data points satisfying  $0.5 \leq \frac{C_p}{C_o} \leq 2$ , and FAC5 the fraction of data points satisfying  $0.2 \leq \frac{C_p}{C_o} \leq 5$ ,

where  $C_p$  are model predictions of concentration,  $C_o$  observations of tracer concentration and  $\overline{C}$  represents the mean of the dataset. Perfect model agreement would be indicated by MG, VG, FAC2 and FAC5 values of 1 and FB and NMSE values of zero. FB can be interpreted by the following equation:

$$\frac{\overline{C}_p}{\overline{C}_o} = \frac{1 - \frac{1}{2}FB}{1 + \frac{1}{2}FB}, \quad (5)$$

and MG can be interpreted as  $\langle C_p \rangle / \langle C_o \rangle = 1/MG$ , where  $\langle C \rangle$  refers to the geometric mean of the dataset. All metrics are dimensionless with the exception of the MG and VG, which are in the units of measurement (ppm(v)).

Chang and Hanna (2004), after statistical examination of many dispersion datasets, have concluded that for ‘research grade’ field experiments (such as MUST), a range of metric values to indicate ‘acceptable model performance’ is as follows:  $-0.3 < FB < 0.3$ ,  $0.7 < MG < 1.3$ ,  $NMSE < 4$ ,  $VG < 1.6$ , and

$FAC2 > 0.5$  for Arcmax comparison where values below the detection limit have been set to this value. No criteria are outlined for point to point comparison or for datasets with zero points removed.

Section 5 discusses the WinMISKAM model’s performance against all horizontal samplers for the selected runs as analysed by Arcmax and point to point comparison. Table 2 shows the number of observation pairs ( $\#$  obs) and the number of pairs used in evaluation after zero pairs were removed (n). Table 3 show only n as no pairs were removed.

## 5. Results and Discussion

Of primary interest when evaluating a model for emergency response applications are false positives and false negatives. Fig.2b shows all pairs of observations and model prediction used in this analysis. The cases where the prediction was below the detection limit and has been set to this value as the corresponding observation showed a discernible tracer concentration can be clearly seen on this figure. These false negatives result from the plume direction being mis-modelled. The majority of these points lie at the lower end of the scale; the model has predicted all but a few of the higher concentrations. These pairs represent the worst kind of error in an emergency response model. Understanding that these are in fact caused by small differences in plume direction is reassuring, but given the much more complex nature of real urban areas, remains a cause for concern. The false positives are also skewed to the lower end of the scale, supporting this hypothesis.

Of secondary interest in an emergency response evaluation is the model’s tendency to over or underpredict concentrations. Linear measures to evaluate this, such as FB, can be influenced by infrequent outliers. Table 2 shows statistics for the point to point comparison, where FB shows the slight mean overprediction of the neutral cases ( $\overline{C_p}/\overline{C_o}=1.14$ ), underprediction of the stable cases ( $\overline{C_p}/\overline{C_o}=0.57$ ) and an overall underprediction of  $\overline{C_p}/\overline{C_o}=0.77$ . The large NMSE of 3.4 for all horizontal samplers paired in time and space shows that the distribution of data is not normal, but log-normal (Chang and Hanna, 2004), so the logarithmic metrics of MG and VG are more appropriate when analysing the data. The MG values, which show only systematic errors (Chang and Hanna, 2004), can also return favourable values despite compensating errors of over and under prediction. MG shows more significant underprediction than does the FB (another systematic error metric), suggesting that all regimes are underpredicted ( $\langle C_p \rangle / \langle C_o \rangle = 0.35$  for stable cases, 0.84 for neutral cases, and 0.61 overall); this underprediction can be seen clearly in Fig.2.

The net underprediction on line 1 for all values paired in time and space is shown by the  $MG = 2.0$  in Table 2. This line also shows the largest scatter of all lines, with a  $VG$  of 50.8, suggesting that the plume is not being sufficiently dispersed near the source, leading to higher plume centreline maxima (evidenced in Table 3,  $MG=0.88$ , or overprediction of line 1 maxima of approximately 14%) and lower values towards the edge of the plume. Plume direction mismodelling

has also affected the level of scatter as high predictions are often paired with very low observations and vice versa.

This supports the findings of Dixon et al. (2006) who report sharper WinMISKAM modelled peaks than seen in observations and overestimation of maximum values. They cite lower levels of predicted turbulence than seen in observations and reduced vertical velocities as the reason for this model behaviour. Santiago and Martilli (2007) found similar model behaviour for the FLUENT code when compared to MUST observations, especially in the first rows, as did Camelli et al. (2004) for FEFLO-URBAN. This effect has also been noted in this study, particularly in runs where the incident wind was less than  $30^\circ$  to the array, such as runs 2640246, 2672150 and 2672213 (not shown).

The mismodelling of plume direction can also be seen to be attributable to poor predictions of vertical flows close to obstacles. This is supported by further examination of runs 2681829 and 2681849. Here the release location was the same in the x,y plane (Fig.1), and the wind direction, windspeeds and stabilities were also similar but release heights varied; being 1.8 m and 0.15 m respectively (Table 1). While 2681849 had a slightly larger angle of wind incidence and lower windspeed, the effects of these differences are likely to be small compared to the effect of the change in release height, which has influenced the initial modelled plume deviation and channeling, but has had remarkably little effect on observed plume shape and distribution.

Fig.3a and b show WinMISKAM output for runs 2681829 and 2681849 respectively, focussed on the the near-source region. Also shown is the wind direction (solid arrow) and the plume centreline (dashed arrow). In both instances the incident wind angle was significantly greater than  $20^\circ$  to normal to the array and the plumes have deviated away from normal to the array, in accordance with findings of Yee and Biltoft (2004) and Gailis and Hill (2006). There is a marked shift in the apparent source between the output of the two model runs, seen by tracing back along the modelled plume centreline (Fig.3).

Fig.4a and b show observations and model predictions for line 1 for runs 2681829 and 2681849. Despite the difference in the release height, the observations between the two runs are very similar. The model predictions however, are quite varied. The lower released plume (2681949) has been modelled impacting more fully onto the long edge of container K8 and has been caught in a recirculation zone and transported down the corridor between container rows K and L significantly more than has the 2681929 plume.

It is likely that the higher plume (2681829) has been swept over container K8 as the release height was  $\approx 71\%$  of the container height. The similarities in observations suggest that the 2681849 plume has also been swept over K8 by vertical wind at the container face. The predictions for run 2681929 are very similar to the observations, suggesting that this case has been well modelled, although the modelled plume can be seen to be more dispersed and slightly displaced along the sampling line than the observed plume. The slight overprediction, shown by the larger area under the model prediction curve in Fig.4a, indicates that more propylene has been lost to vertical transport in observations than predictions, supporting the interpretation that the vertical transport was

underpredicted.

The predictions for run 2681949 however, show much more significant displacement and dispersion than can be seen in observations. This indicates again that vertical transport has been underpredicted, supporting Dixon et al. (2006)'s assertion regarding vertical windspeeds in WinMISKAM. That this plume has been modelled to impact so fully on the container face, where the observed plume can be seen to behave in a large part without interference from the container supports this view.

This example directly supports findings of Baechlin et al. (1992) and Theurer (1995) who showed that the specific geometries of near source buildings are integral in the dispersion of pollutants in the near field and highlights the importance of using obstacle resolving models for dispersion within built up areas, where flow patterns can be highly complex (DePaul and Sheih, 1986; Kastner-Klein et al., 2004).

Under stable cases, the line one values show much more scatter, when considered point to point, than predictions under neutral conditions and a more pronounced tendency to underpredict, which can be seen by the red crosses representing stable line 1 predictions in Fig.2b. The poor performance for stable cases for all lines is evidenced by more significant underprediction, greater scatter, lower FAC2 and FAC5 values and significantly greater NMSE values than the neutral cases for point to point comparison, (Table 2). Arcmax values also show underprediction for stable cases, contrasting the good performance for neutral cases, (Fig.2a) and in the MG metrics (Table 3).

The overpredicted maxima effect is less obvious under stable conditions, where the model has not captured the higher concentrations caused by the stable atmosphere; supporting Gailis and Hill (2006) who found that the effects of a stably stratified atmosphere on flows were not negligible when modelling in high turbulence areas such as an urban style array. Under these conditions any tendency for the model to overpredict maxima is masked by the general tendency of the model to underpredict all concentrations. Some instances of narrow plumes without the overpredicted maxima include runs 2672101 and 2672033 (not shown). Poor model performance for stable cases is contrasted with the adequate performance in neutral conditions. As these are the conditions which tend to dominate in urban areas due to the influence of building induced turbulence (Bowne and Ball, 1970; Yersel and Goble, 1986; Lundquist and Chan, 2007), this result is acceptable for urban emergency response modelling.

The tendency in point to point comparison for model underprediction to decrease with further array penetration (MG = 2, line 1; MG = 1.17, line 4), coupled with the lowest FAC2 and FAC5 values in line 1 suggests poor prediction close to the source. Similar results were published for Mercure\_Saturne (Milliez and Carissimo, 2007), while Hanna et al. (2004a) found little difference with downwind distance for MUST simulations with the FLACS model.

Further from the source (line 4), the maxima are significantly underpredicted for both stable and neutral cases, with underprediction of all  $C_{ground}$  at these downwind distances. The very low observed concentrations on the sampling lines further from the source lead to increased difficulty in predicting maxima

within a factor of 2 or 5, and thus lower FAC2 (0.63) and FAC5 (0.84) values and a higher degree of scatter ( $VG = 2.77$ ) than can be seen for any other line. In the two cases showing the worst performance, which can be seen on the bottom left of Fig.2a (runs 2671852 and 2692223), the tracer has blown almost entirely off the grid; these small differences are exaggerated on the log plot.

The model has predicted best in Lines 2 and 3, with point comparisons showing the lowest scatter ( $VG$  and  $NMSE$ ) and highest FAC2 and FAC5 values. These lines do not exhibit the difficulties associated with predicting very low concentrations, or near source modelling issues.

The performance of the WinMISKAM model against MUST observations is well within the bounds proposed by Chang and Hanna (2004) for acceptable Arcmax performance. To compare with other models' performance, it is easiest to look at overall statistics on all  $C_{ground}$ . Warner et al. (2006) investigated the performance of the HPAC software with the urban dispersion module (UDM) turned on - regarded by Gailis and Fulford (2001) as the best operational tool for hazard predictions available at that time - when run against the MUST data, reporting an overall MG of 1.14, VG of 6579 and a FAC2 of 0.34 for point to point comparison. These statistics reflect comparison against 37 MUST runs in the horizontal and vertical planes, as opposed to the horizontal samplers for 19 runs investigated herein.

The FLACS MUST evaluation, conducted by Hanna et al. (2004a), cites Arcmax MG = 1.57, VG = 1.7 and a FAC2 of 0.64 for 37 trials. Mercure\_Saturne was evaluated for point to point comparison against the MUST dataset by Milliez and Carissimo (2007) who show MG = 1.01, VG = 1.84 and a FAC2 of 0.7 for all horizontal samplers for 20 trials (19 of which are modelled herein). The WinMISKAM model thus performed well compared to these models, and the ease of use of the model (sources can be easily selected with a mouse click and only a small number of choices must be made by the user for a dispersion run) also lends itself well towards its selection as an emergency response tool. The much better results obtained by Milliez and Carissimo (2007) are consistent with the more demanding input parameters required by Mercure\_Saturne.

The high computational cost (a single windfield took approximately 11 hours to generate on a 2.99 Ghz Intel Xeon system with 2GB RAM and 1 minute dispersion runs took between 40 minutes and 2 hours on the same system) currently precludes this model from use in emergency response situations where windfields must be calculated in real-time. These windfields could, however, be pre-modelled and stored in a look up table, from which the most similar situation to the atmospheric conditions prevailing in an emergency situation, as a function of wind speed and direction, could be selected.

To use such a model in an emergency situation would require, in the first instance, an increase in grid cell size. This study was run on a very fine grid of 0.5 m cells, whereas a typical central urban grid may have cells of 2 - 5 m, allowing dispersion runs to be conducted much more quickly. Dixon et al. (2006) showed that the model is not overly sensitive to changes of domain size, but more so to the changes in turbulent kinetic energy effected by upwind roughness elements - thus highlighting the importance of a grid covering an area well in excess of the

area of interest. Thus, while larger grids may be used, a larger area must also be covered to provide this upstream turbulent kinetic energy spectrum. While this study does not evaluate the performance of the model on differently sized domains, it does show the performance of the model in comparison to other models for a standardised situation.

Secondly, with recent increases in computing power showing no signs of abatement, the time taken to conduct dispersion runs will be greatly reduced in coming years, especially if codes can be optimised for multi-core processing. This would enable the use of obstacle resolving dispersion models, with pre-modelled windfields, for real-time predictions in emergency situations.

The model, despite resolving obstacles to a fine resolution, does not capture all important features of the flow to a sufficient degree, in particular with regards to vertical transport and recirculation zones. There are clearly some emergency response applications where this will be seriously detrimental to model application. However, the MUST array uses individual blocks with a height to width ratio of  $\approx 0.2$  separated by channels, whereas in a real urban situation, buildings typically have much greater height to width ratios and are arranged in blocks, where the vertical transport and recirculation is likely to be of lesser importance than building induced channeling.

## 6. Conclusions

An investigation into the performance of WinMISKAM on selected MUST trials, driven by on-site meteorological data and compared to on-site ground level observations showed the model has performed well within the bounds of ‘good model performance’ outlined by Chang and Hanna (2004). The model has, in general, predicted the maximum concentrations across a given sampling line in neutral conditions very well, and has underpredicted maxima in stable conditions. Modelled plumes are seen to be generally narrower than observed plumes, resulting in overall underprediction by the model for point to point comparison. Underprediction was much more pronounced for stable conditions than neutral condition. As complex geometries in urban centres push stability towards neutral, this result can be seen to be acceptable for emergency response modelling in built-up areas. Plume direction was most poorly represented, indicating a possible area for model development to pursue. Poor representation of vertical flow is suggested as the cause of the narrower plumes and plume direction mismodelling.

The model is well suited to emergency response planning and preparation, due to its ease of use. It is recommended that emergency responders be briefed in model interpretation before being expected to deal with model output, and that model predictions for atmospheric conditions far from neutral be treated with caution. Although is currently too slow to be of practical use for real time emergency prediction, increases in computing power in accordance with those seen over the past decades would enable its use in such situations within years, not decades.

## Acknowledgements

This work was funded by the Australian Research Council (project LP0455832), Chemistry Centre of W.A. and the Fire and Emergency Service of Western Australia. The authors would like to thank Ingenieurbuero Lohmeyer GmbH & Co. KG for the use of the WinMISKAM model. All this assistance is gratefully acknowledged.

## References

- Allwine, K., Leach, M., Stockham, J., Shinn, J., Hosker, R., Bowers, J., and Pace, J. (2004). "Overview of Joint Urban 2003 - An atmospheric dispersion study in Oklahoma City." In "Pre-prints, Symposium on Planning, Nowcasting and Forecasting in the Urban Zone.", American Meteorological Society, Seattle, Washington.
- Archambeau, F., Mechtoua, N., and Sakiz, M. (2004). "Code\_Saturne : a finite volume code for the computation of turbulent incompressible flows - industrial applications." *Int. J. Finite Volumes*, 1: 1 – 62.
- Baechlin, W., Theurer, W., and Plate, E. J. (1992). "Dispersion of gases released near the ground in built up areas: experimental results compared to simple numerical modelling." *J. Wind Eng. Ind. Aero.*, 41-44: 2721 – 2732.
- Batchvarova, E. and Gryning, S. E. (2006). "Progress in urban dispersion studies." *Theoretical and Applied Climatology*, 84(1): 57–67.
- Bezpalcova, K. and Harms, F. (2005). "EWTL Data Report / Part I: Summarized test description mock urban setting test." Technical report, Environmental Wind Tunnel Laboratory, Center for Marine and Atmospheric Research.
- Biltoft, C. A. (2001a). "Abbreviated test plan for customer test - Mock Urban Setting Test." Technical Report WDTC-TP-01-028, West Desert Test Centre, U.S. Army Dugway Proving Ground, Dugway, UT.
- Biltoft, C. A. (2001b). "Customer report for Mock Urban Setting Test." Technical report, DPG Document No. WDTC-FR-01-121, US army Dugway Proving Ground, Dugway, Utah.
- Bowne, N. E. and Ball, J. T. (1970). "Observational comparison of rural and urban boundary layer turbulence." *Journal of Applied Meteorology*, 9: 862 – 873.
- Britter, R. E. and Hanna, S. R. (2003). "Flow and dispersion in urban areas." *Annual Review of Fluid Mechanics*, 35(1): 469–496.



- Brook, D. R., Felton, N. V., Clem, C. M., Strickland, D. C. H., Griffiths, I., Kingdon, R., Hall, D. J., and Hargrave, J. M. (2003).** “Validation of the Urban Dispersion Model (UDM).” *International Journal of Environment and Pollution*, 20(1 - 6): 11 – 21.
- Camelli, F. E., Lohner, R., and Hanna, S. (2004).** “Feflo-Urban CFD Model Evaluation Of The Must Experiment.” *Mecanica Computacional*, 23(16): 1701 – 1725.
- Chang, J. and Hanna, S. (2004).** “Air quality model performance evaluation.” *Meteorology and Atmospheric Physics*, 87(1-3): 167–196.
- Coccal, O., Thomas, T., Castro, I., and Belcher, S. (2006).** “Mean Flow and Turbulence Statistics Over Groups of Urban-like Cubical Obstacles.” *Boundary-Layer Meteorology*, 121(3): 491–519.
- COST (2007).** “COST Action 732: Best practice guideline for the CFD simulation of flows in the urban environment.” Technical report, University of Hamburg Meteorological Institute Centre for Marine and Atmospheric Sciences.
- DePaul, F. T. and Sheih, C. M. (1986).** “Measurements of wind velocities in a street canyon.” *Atmospheric Environment*, 20: 455 – 459.
- Dixon, N. S., Boddy, J. W. D., Smalley, R. J., and Tomlin, A. S. (2006).** “Evaluation of a turbulent flow and dispersion model in a typical street canyon in York, UK.” *Atmospheric Environment*, 40(5): 958–972.
- DTRA (2001).** “The Hazard Prediction and Assessment Capability (HPAC) users guide, version 4.0.3.” Technical Report HPAC-UGUIDE-02-U-RAC0, Prepared for the Defense Threat Reduction Agency by Science Applications International Corporation.
- Eichhorn, J. (2008).** “MISKAM Manual for Version 5.” Technical report, Giese-Eichhorn Environmental Meteorological Software.
- Eichhorn, J. and Kniffka, A. (2007).** “An improved version of the microscale flow model MISKAM - evaluation according to VDI guideline 3783/9.” In “11th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes,” Cambridge, UK, July 2-5, 2007.
- Eichhorn, J., Schrodin, R., and Zdunkowski, W. (1988).** “Three-Dimensional numerical simulations of the urban climate.” *Beitr. Phys. Atmosph.*, 61(3): 187 – 203.
- EPA, U. S. (1999).** “ALOHA User’s Manual.” Technical report, Environmental Protection Agency, Office of Energy Management, Washington D.C. <http://www.epa.gov/emergencies/content/cameo/aloha.htm>.



- Fernando, H. J. S., Lee, S. M., Anderson, J., Princevac, M., Pardyjak, E., and Grossman-Clarke, S. (2001).** “Urban Fluid Mechanics: Air Circulation and Contaminant Dispersion in Cities.” *Environmental Fluid Mechanics*, 1(1): 107–164.
- Gailis, R. and Fulford, G. (2001).** “Agent Dispersion and Prediction.” Technical Report DSTO-GD-0293, Aeronautical and Maritime Research Laboratory.
- Gailis, R. and Hill, A. (2006).** “A Wind-Tunnel Simulation of Plume Dispersion Within a Large Array of Obstacles.” *Boundary-Layer Meteorology*, 119(2): 289–338.
- Golder, D. (1972).** “Relations among stability parameters in the surface layer.” *Boundary-Layer Meteorology*, 3(1): 47–58.
- Goricsan, I., Balczo, M., Czader, K., Rakai, A., and Tonko, C. (2007).** “Simulation of flow in an idealised city using various CFD codes.” In “11th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes,” Cambridge, UK.
- Grimmond, C. S. B. (2006).** “Progress in measuring and observing the urban atmosphere.” *Theoretical and Applied Climatology*, 84(1): 3–22.
- Hanna, S., Baja, E., Flaherty, J. E., and Allwine, E. (2008).** “The Use of Tracer Data from the Madison Square Garden 2005 Field Experiment to Test a Simple Urban Dispersion Model.” In “15th Joint Conference on the Applications of Air Pollution Meteorology with the A&WMA,” New Orleans, LA.
- Hanna, S., White, J., Zhou, Y., and Kosheleva, A. (2006).** “Analysis of Joint Urban2003 and Madison Square Garden 2005 meteorological and tracer data.” In “paper J7.1, Sixth Symposium on the Urban Environment AMS Forum .”, American Meteorological Society, Boston, Ma.
- Hanna, S. R., Chang, J. C., and Fernau, M. E. (1998).** “Monte carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables.” *Atmospheric Environment*, 32(21): 3619–3628.
- Hanna, S. R., Chang, J. C., White, J. M., and Bowers, J. F. (2007).** “Evaluation of the HPAC dispersion model with the Joint Urban 2003 (JU2003) tracer observations.” In “Seventh Symposium on the Urban Environment (7URBAN),” San Diego, CA.
- Hanna, S. R., Hansen, O. R., and Dharmavaram, S. (2004a).** “FLACS CFD air quality model performance evaluation with Kit Fox, MUST, Prairie Grass, and EMU observations.” *Atmospheric Environment*, 38(28): 4675–4687.

- Hanna, S. R., Reynolds, R. M., Heiser, J., and Bornstein, R. (2004b).** “Plans for MSG04 tracer experiment in Manhattan.” In “Paper 1.5 at AMS Urban Environment Conference,” Available at AMS web site ([www.ametsoc.org](http://www.ametsoc.org)) and on CD from AMS, 45 Beacon St., Boston, MA 02215., Vancouver.
- Harman, I. N. and Belcher, S. E. (2006).** “The surface energy balance and boundary layer over urban street canyons.” *Quar. Jour. Royal Meteorol. Soc.*, 132: 2749–2768.
- Homann, S. G. (1994).** “HOTSPOT Health Physics Codes for the PC.” Technical report, Lawrence Livermore National Laboratory, University of California, Livermore.
- Kahn, S. M. and Simpson, R. W. (1997).** “Modelling mixing height from routinely measured surface and upper air meteorological data.” *Environmental Modeling and Assessment*, 2: 191200.
- Kastner-Klein, P., Berkowicz, R., and Britter, R. (2004).** “The influence of street architecture on flow and dispersion in street canyons.” *Meteorology and Atmospheric Physics*, 87(1): 121–131.
- Ketzel, M., Berkowicz, R., and Lohmeyer, A. (2000).** “Comparison of Numerical Street Dispersion Models with Results from Wind Tunnel and Field Measurements.” *Environmental Monitoring and Assessment*, 65(1): 363–370.
- Lohmeyer, A., Eichhorn, J., Flassak, T., and Kunz, W. (2002a).** “WinMISKAM 4.2, microscale flow and dispersion model for build up areas, recent developments.” In “11th International Symposium Transport and Air Pollution. vol. 2.”, Graz, Austria.
- Lohmeyer, A., Mueller, W. J., and Baechlin, W. (2002b).** “A comparison of street canyon concentration predictions by different modellers: final results now available from the Podbi-exercise.” *Atmospheric Environment*, 36(1): 157–158.
- Lundquist, J. K. and Chan, S. T. (2007).** “Consequences of Urban Stability Conditions for Computational Fluid Dynamics Simulations of Urban Dispersion.” *Journal of Applied Meteorology and Climatology*, 46(7): 1080–1097.
- Milliez, M. and Carissimo, B. (2007).** “Numerical simulations of pollutant dispersion in an idealized urban area, for different meteorological conditions.” *Boundary-Layer Meteorology*, 122: 321–342.
- Mohan, M. and Siddiqui, T. A. (1998).** “Analysis of various schemes for the estimation of atmospheric stability classification.” *Atmospheric Environment*, 32(21): 3775–3781.

- NOAA (2004).** “Federal Research and Development Needs and Priorities for Atmospheric Transport and Diffusion Modeling.” Technical Report FCM-R23-2004, National Oceanic and Atmospheric Administration Joint Action Group for Atmospheric Transport and Diffusion Modeling (Research and Development Plan).
- Oke, T. R. (1987).** *Boundary Layer Climates*. Methuen, London.
- Salmond, J. A., Oke, T. R., Grimmond, C. S. B., Roberts, S., and Offerle, B. (2005).** “Venting of heat and carbon dioxide from urban canyons at night.” *J. Appl. Meteorology and Atmospheric Physics*, 44: 1180–1194.
- Santiago, J. L. and Martilli, A. (2007).** “Simulation of MUST experiment using RANS K-Epsilon model: Validation against wind tunnel measurements and analysis of spatial average properties.” In “Proceedings of the 11th International Conference on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes,” Cambridge, United Kingdom.
- Schatzmann, M. and Leitl, B. (2002).** “Validation and application of obstacle-resolving urban dispersion models.” *Atmospheric Environment*, 36(30): 4811–4821.
- Theurer, W. (1995).** “Point sources in urban areas: modelling of neutral gas clouds with semi-empirical models.” In C. J. E., editor, “Proc. NATO Advanced Studies Institute, Wind Climate in Cities,” pages 485–502. Kluwer, Dordrecht.
- Warner, S., Platt, N., Heagy, J. F., Jordan, J. E., and Bieberbach, G. (2006).** “Comparisons of transport and dispersion model predictions of the mock urban setting test field experiment.” *Journal of Applied Meteorology and Climatology*, 45(10): 1414–1428.
- Yamada, T. (2004).** “Merging CFD and atmospheric modeling capabilities to simulate airflows and dispersion in urban areas.” *Computational Fluid Dynamics Journal*, 13(2)(47): 323 – 335.
- Yee, E. and Biltoft, C. (2004).** “Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles.” *Boundary-Layer Meteorology*, 111: 363 – 415.
- Yersel, M. and Goble, R. (1986).** “Roughness effects on urban turbulence parameters.” *Boundary-Layer Meteorology*, 37(3): 271–284.

Table 1: List of selected trials with release rates (Q), release duration, Monin-Obukhov length (L), release height, average wind speed and average wind direction onto the face of the array.

Run Name	Q (l/min)	Release Duration (min)	L <sup>†</sup> (m)	Release Height (m)	Wind Speed (ms <sup>-1</sup> )	Wind Angle (deg)*
2640138	175	15	91	0.15	1.6	-19
2640246	200	15	62	0.15	2.4	23
2671852	200	22	330	0.15	3.0	-64
2671934	200	15	5.8	1.8	1.8	-61
2672033	200	15	4.8	1.8	2.7	-38
2672101	200	14	7.7	0.15	2.2	-23
2672150	200	16	150	0.15	2.3	27
2672213	200	15	150	1.8	2.7	29
2672235	200	15	48	2.7	2.5	27
2672303	200	19	74	1.8	2.8	19
2681829	225	15	28000	1.8	7.7	-44
2681849	225	15	2500	0.15	6.4	-47
2682256	225	15	240	0.15	4.2	-48
2682320	225	15	170	2.7	3.6	-42
2692054	225	22	170	1.3	3.2	39
2692131	225	17	220	1.3	4.0	34
2692157	225	15	130	2.7	3.2	37
2692223	225	15	120	1.3	3.2	33
2692250	225	17	130	1.3	3.8	32

† Source Milliez and Carissimo (2007)

\* Indicates angle of wind onto the array where a Southerly wind (180°) would appear as -27°, a wind from (153°) would appear as 0° and a wind from 126° would appear as 27°. Wind speeds and directions calculated from vector averaged 6 m on-site sonic data, taken over the time of the 200 s sample.

Table 2: Statistical evaluation of WinMISKAM performance in predicting ground level concentrations for all modelled MUST trials paired in time and space.

Statistical Measure	Line 1	Line 2	Line3	Line 4	All Horizontal	All Stable	All Neutral
MG	2.00	1.50	1.80	1.17	1.63	2.88	1.19
VG	50.8	8.46	7.68	10.9	15.9	120	5.10
FB	0.21	0.32	0.43	0.26	0.26	0.55	-0.14
NMSE	2.79	1.46	1.39	1.47	3.40	3.36	1.50
FAC2	0.43	0.50	0.47	0.45	0.46	0.27	0.57
FAC5	0.66	0.78	0.78	0.72	0.73	0.58	0.81
n	171	129	102	102	511	184	327
‡ obs	228	171	171	190	760	280	480

Table 3: Statistical evaluation of WinMISKAM performance in predicting ground level Arc-Max concentrations for all modelled MUST trials.

Statistical Measure	Line 1	Line 2	Line3	Line 4	All Horizontal	All Stable	All Neutral
MG	0.88	0.98	1.32	1.36	1.11	1.33	1.01
VG	1.34	1.16	1.21	2.77	1.51	1.47	1.53
FB	0.04	0.08	0.35	0.31	0.09	0.29	-0.26
NMSE	0.60	0.44	0.53	0.55	1.12	0.92	0.66
FAC2	0.84	0.95	0.90	0.63	0.83	0.71	0.90
FAC5	1.00	1.00	1.00	0.84	0.96	1.00	0.94
n	19	19	19	19	76	28	48

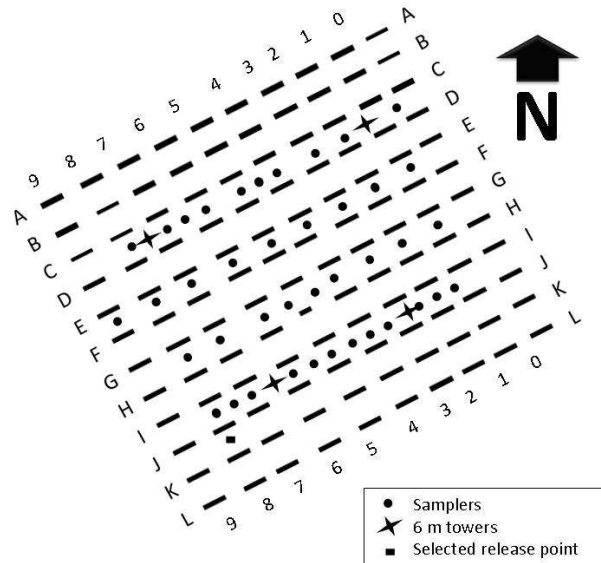


Figure 1: Diagram of MUST array showing 120 containers, four 6 m towers holding the sonics used in this study (stars) and sampler locations (dots) arranged in 4 lines - line 1 is towards the bottom of the figure between rows I and J and line 4 is towards the top between rows C and D. The release point for runs 2681829 and 2681849 is shown as a small rectangle between containers K8 and J8.

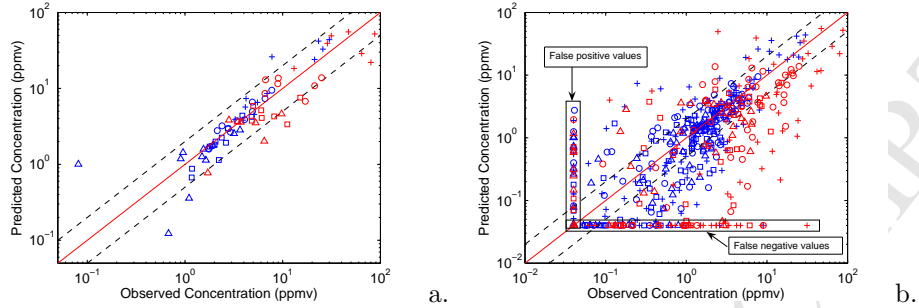


Figure 2: WinMISKAM predictions against MUST observations (a) Arcmax and (b) Paired in space and time. Crosses (+) show pairings for line 1, circles (o) show line 2, squares ( $\square$ ) show line 3 and triangles ( $\triangle$ ) show line 4. Stable cases are shown in red and neutral cases are shown in blue.

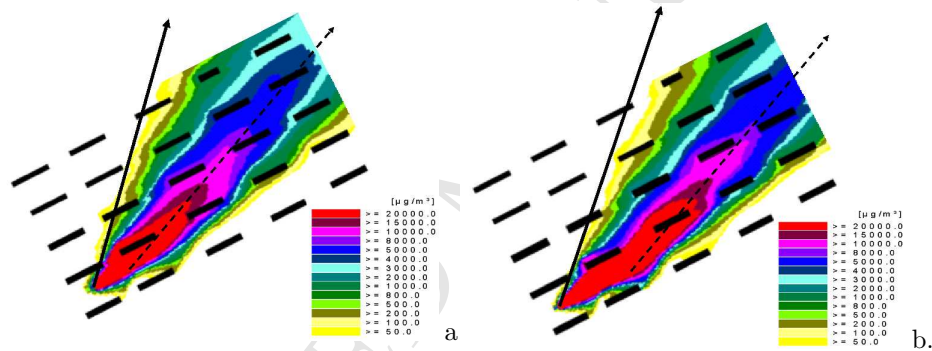


Figure 3: Near source region of WinMISKAM output for run 2681829 (a) and 2681849 (b) during the 200 s sample time.

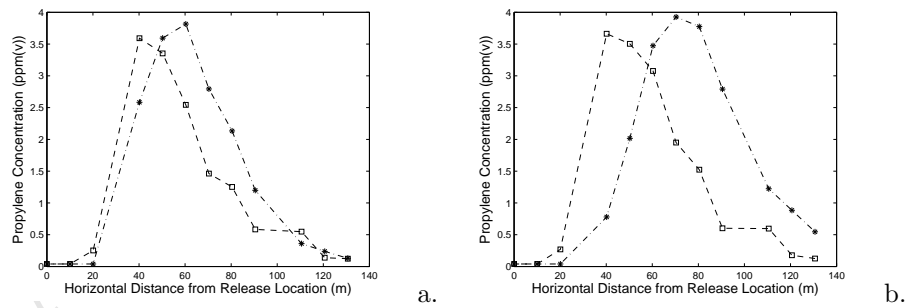


Figure 4: Line 1 observed propylene concentration (dashed line with boxes) and WinMISKAM predictions (broken dashed line with stars) for runs 2681929 (a) and 2681849 (b). Zero on the x axis shows the horizontal position of the release location, with positive values Northeast from the release location along the sampling line.