18th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 October 2017, Bologna, Italy

ANALYTICAL MODELLING OF DISPERSION FOR BAYESIAN SOURCE TERM ESTIMATION

Andrew R. Webb and Paul B. Westoby

Dstl, Porton Down, UK.

Abstract: The downwind hazard from a Chemical or Biological (CB) release is calculated via Atmospheric Dispersion Models (ADMs), which can provide estimates of the mean concentration of an agent and its variance as a function of time and spatial location for a release of known properties. These models can vary substantially in complexity and computational requirements and the estimates of concentration are often used within complex computer models for source term estimation, hazard prediction or sensor placement applications generally known as decision support tools, where the ADM can be run millions of times but answers are required in near real time to protect lives. Therefore, the computation time of a dispersion model is a critical factor in how effective it will be in various applications.

The aim of this study is to assess a simple, but rapidly evaluated dispersion model that has utility within real time DSTs. If accurate enough, the rapid model could be used as a pre-processor in DSTs to reduce the number of complex dispersion model runs required.

A simple, rapid ADM based on a Gaussian plume model is formulated for a continuous, but finite duration, release. The dispersion model is developed for specific parametric forms for the standard deviations σ_x, σ_y and σ_z, of the stream wise, lateral and vertical concentrations. The model is developed within a statistical learning framework whereby model parameters are determined by optimising a cost function using historical data. This is done for both the urban and open-terrain environments. The approach is assessed on FFT07 and JU2003 field trial data using several different forms for σ_x, σ_y and σ_z; different cost functions, different methods of nonlinear optimisation and several performance measures.

Key words: *analytic dispersion model; source term estimation; Bayesian methods; Gaussian plume model; performance measures; FFT07 and JU2003 field trial data.*

INTRODUCTION

The motivation for the current study is the inverse dispersion modelling problem of a particular Decision Support Tool (DST), where the aim is to determine the properties of a release of a (potentially) hazardous CB material, given sensor measurements in the field. This is known as Source Term Estimation (STE; Robins, P. *et al*., 2009; Yee, E., 2012; Bieringer, P.E. *et al*., 2015). The purpose of this paper is to describe the optimisation of a simple analytical dispersion model that can be used as a pre-processor in STE studies for a wide range of operating conditions. In STE applications, the dispersion model is typically evaluated millions of times and in the case of STE solutions must be obtained rapidly to allow mitigation of the potential hazard. The analytic model could be used to minimise the use of the complex long running dispersion model.

The analytic model proposed in this work is based on the Gaussian puff model. It is a development of ideas in Hanna and Baja (2009) and Luhar (2011). The model includes the following features:

- It is not steady state – it includes the start of the release as a model parameter (Luhar, 2011);
- The release is of finite duration – the release duration is a model parameter;
- Reflections from the ground and boundary layer top are included (following Luhar, 2011);
- There are non-zero initial values for σ_x, σ_y and σ_z – for use in an urban environment to allow for mixing around buildings near the source (Hanna and Baja, 2009).

The approach taken to parameter estimation is to define a cost function measuring the discrepancy between observations and the model predictions. Several cost functions and methods of optimisation are assessed. Historical trials data from the Fusion Field Trial 2007 (FFT07, Allwine and Flaherty, 2006) and Joint Urban 2003 field trial (JU2003, Storwold, 2007; Clawson *et al.*, 2005) are used to validate the procedure and assess how widely applicable the ADM is once it has been trained.

DISPERSION MODEL OPTIMISATION

Approach

For a particular terrain type (open, urban, dense urban), the rapid ADM is trained using experimental data, where the source terms are known. This trained ADM can then be used within an STE system to identify unknown source term parameters from downwind sensor data. During both model training and STE the meteorological conditions are considered unknown and are optimised.

Let the predicted concentration at location (x, y, z) at time t be denoted by $\bar{c}(x, y, z, t;\ \theta, \phi, \nu)$, where x, y and z denote the streamwise, lateral and vertical directions. We distinguish between three types of model parameters, θ, ϕ and ν:

- θ refers to the dispersion model parameters that are optimised as part of the model training procedure. These parameters depend on the environment (terrain type, release conditions) and not the source terms.
- ϕ represents the source terms, which are known for the training procedure, but are optimised as part of the STE process (for which θ, the model parameters, is known).
- ν represents the meteorological conditions, which are optimised for each scenario.

We define a cost function, $\mathcal{L}$, that measures the difference between the predictions and the observations at a set of locations and times,

$$\mathcal{L} = \mathcal{L}\big(\boldsymbol{c}(t, \theta, \phi, \nu), \boldsymbol{m}(t), \{\boldsymbol{u}_j\}\big)$$

where $\boldsymbol{c}(t, \theta, \phi, \nu)$ is a vector of predictions at a set of sensor locations, with $c_i(t, \theta, \phi, \nu) = \bar{c}(x_i, y_i, z_i, t;\ \theta, \phi, \nu)$, the predicted concentration at sensor i at time t for model parameters θ, ϕ and ν; $m_i(t)$ is the measured concentration at sensor i at time t and $\{\boldsymbol{u}_j\}$ is the set of wind direction measurements. For each field trial experiment, i, the source term parameters (ϕ_i) are known and $\mathcal{L}$ is optimised with respect to the dispersion model parameter, θ, and meteorological parameters (ν_i).

Gaussian puff model

We adopt the ensemble-averaged Gaussian puff model and express the mean concentration at location (x, y, z) at time t since the start of the emission as

$$\bar{c}(x, y, z, t; \theta, \phi, \nu = \bar{u}) = \frac{Q}{(2\pi)^{3/2}} \int_{\max\{0, t-T\}}^{t} \frac{1}{\sigma_x \sigma_y \sigma_z} \exp\left\{-\frac{(x - \bar{u}t')^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{z'^2}{2\sigma_z^2}\right\} dt' \qquad (1)$$

where T is the duration of the release; Q is the rate of emission (so that QT is the total mass released); $z' = z - z_s$, where z_s is the effective source height; $\bar{u}$ is the vector averaged wind speed. The quantities σ_x, σ_y and σ_z represent the standard deviations of the concentration distributions in the stream wise, lateral and vertical directions respectively. Several models have been proposed for σ_x, σ_y and σ_z and we propose the model

$$\frac{\sigma_x}{\sigma_u} = \frac{\sigma_y}{\sigma_v} = \frac{\sigma_z}{\sigma_w} = \begin{cases} \alpha + t' & t' \le 2\tau \\ [(\alpha + 2\tau)(\alpha + t')]^{\frac{1}{2}} & t' > 2\tau \end{cases} \qquad (2)$$

where τ is the Lagrangian interval time scale (Luhar, 2011). For $\alpha = 0$, this reduces to the model of Thomson and Manning (2001) (see also Luhar, 2011), whereby the parameters grow linearly with time for short travel times and as the square root of time for long travel times. For $\alpha > 0$, there are initial values for σ_x, σ_y and σ_z that account for mixing around building obstacles near the source (Hanna and Baja, 2009). The integral in equation (1) can be evaluated analytically (Webb and Westoby, 2017).

EXPERIMENTAL PROCEDURE

Training the model

Two forms for the cost function, $\mathcal{L}$, were assessed: a square error measure and a log likelihood cost function that employed a simple variance model (after Bisignano *et al.*, 2013). We adopted a common value, σ_{uvw}, for σ_u, σ_v and σ_w. From the final converged solution, the following parameters are derived:

$$\sigma_{x0} \triangleq \alpha\sigma_{uvw} \tag{3}$$

and

$$\gamma \triangleq \frac{\sigma_{uvw}}{|u|} \tag{4}$$

where σ_{x0} is the initial size of the puff (in m) and γ is the rate of expansion (in m.m^{-1}). Hanna and Baja (2009) adopt values of 40m for σ_{x0} and 0.25 m.m^{-1} (day) and 0.08 m.m^{-1} (night) for γ for the JU2003 trials.

Evaluation

The goal of the model evaluation is (ultimately) to assess the performance of the model within an STE framework as an initial filtering for a more refined STE model (Westoby, Delle Monache & Silk, 2017). In this case the value of $\theta = \theta^*$, and $\mathcal{L}(\boldsymbol{c}(t, \theta^*, \phi, \nu), \boldsymbol{m}(t), \{\boldsymbol{u}_j\})$, would be optimised on data, $\boldsymbol{m}(t)$ and $\{\boldsymbol{u}_j\}$, from a withheld field trial experiment with respect to the source term parameters, ϕ, and meteorological parameters, ν. The optimised values, ϕ^*, are then compared with the true source term values, thus providing a measure of performance within an STE context.

As a step towards this, we assess the effectiveness of the rapid ADM via the commonly-used measures: Fractional Bias (FB), Normalised Mean Square Error (NMSE) and Factor of 2 (FAC2) using the criteria in Chang and Hanna (2004). Since we are comparing a predicted mean concentration with often a single realisation of concentration, we consider two thresholding procedures based on model prediction $\boldsymbol{c}$ and observations $\boldsymbol{m}$**:** 1) use in the evaluation criteria only values for which the observations and the predictions are above a threshold (exclude 'false positives' and 'false negatives' (threshold criterion, T1)); 2) use in the evaluation criteria only values for which the observations or the predictions are above a threshold (include 'false positives' and 'false negatives' (T2)).

Results

The approach has been assessed using JU2003 puff release (gas analyser) data; JU2003 extended release (integrated gas sampler) data; FFT07 puff release data and FFT07 extended release data. We present results here for the model trained using the maximum likelihood procedure on the JU2003 puff release data and using the T1 threshold to calculate statistics. Further results are presented in (Webb and Westoby, 2017).

Here, Figure 1 shows predicted traces and summary plots (Quantile-Quantile (QQ) and Cumulative Factor (CF) plots) (Herring and Huq, 2017) for one of the puff releases (IOP3 R3), these show reasonable agreement between the rapid approximation and the true values.

Figure 2 summarises the results for all the puff releases for the parameters σ_{x0} and γ. Figure 2 (a) plots σ_{x0} as a function of the normalised mean square error for each puff release in the JU2003 trial. The mean value of σ_{x0} is 38m which is comparable to the results assumed in Hanna and Baja (2009) (40m) and there is little difference between day and night values. Figure 2 (b) plots the expansion rate, γ, as a function of the normalised mean square error for each puff release in the JU2003 trial. Apart from one night-time outlier, the mean value is 0.22, again these results are not correlated with the value of γ and are commensurate with Hanna and Baja (2009).

Errore. L'origine riferimento non è stata trovata. gives the number of puff release solutions that satisfy all of the criteria given in Chang and Hanna (2004) for the JU2003 puff and extended releases for both of the threshold procedures T1 and T2. The success rate, even for the stricter T2 criterion, is remarkably good and suggests that the model may provide a suitable pre-processor in an STE context.

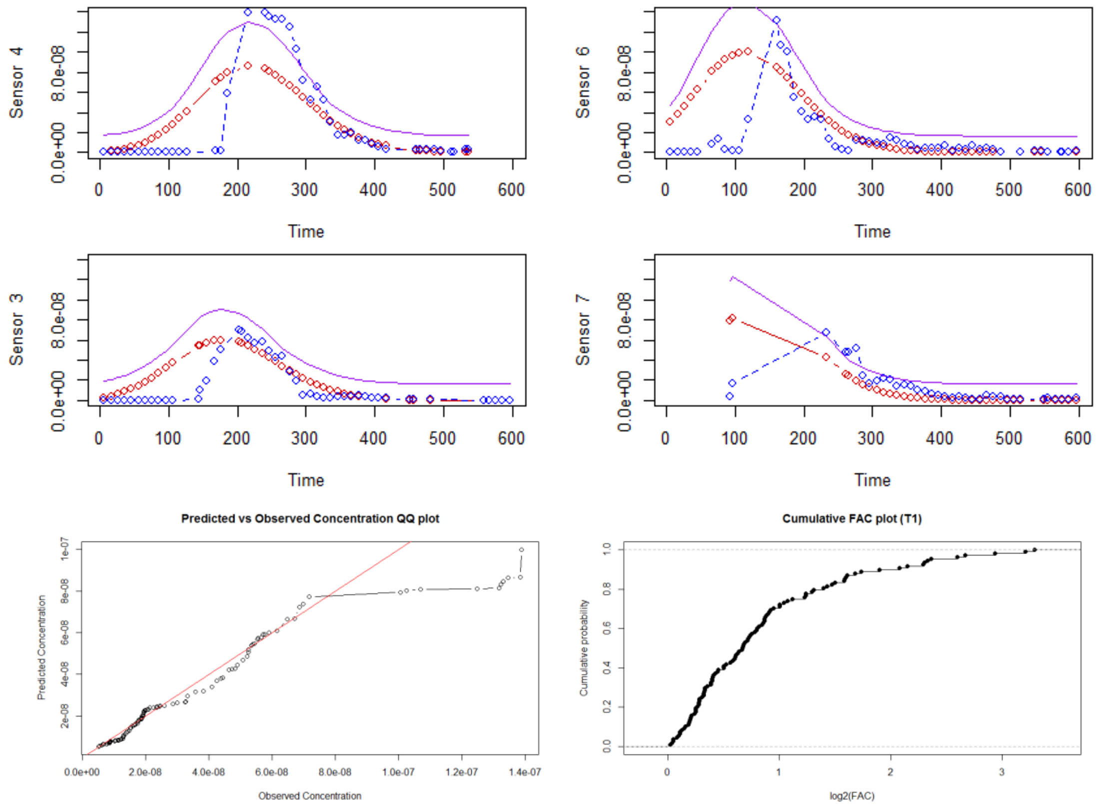

Figure 1. Predicted traces for the four maximally responding sensors (top – blue: measured; red: model; solid line: model plus standard deviation) and QQ and CF summary plots (bottom) for JU2003 puff release IOP3 R3.

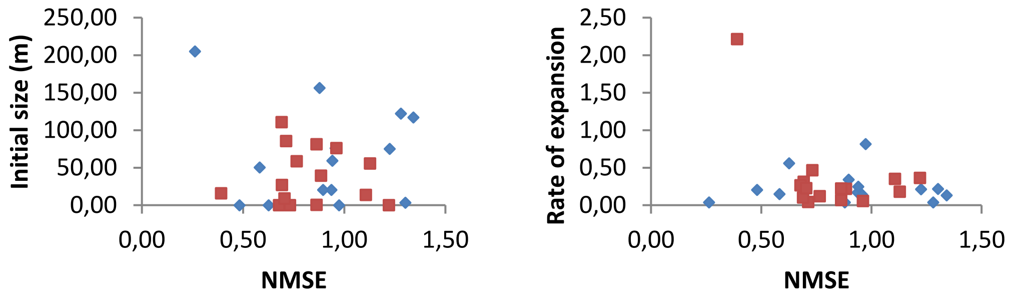

Figure 2. Left: plot of σ_{x0} against NMSE for each puff release in the JU2003 trial. Right: plot of the expansion rate γ against NMSE for each puff release in the JU2003 trial. Blue markers: daytime; red: night.

Table 1. Proportion of solutions satisfying all criteria for each of the threshold schemes.

	T1	**T2**
JU2003 Puff Release	29/29	22/29
JU2003 Extended Release	13/15	9/15

DISCUSSION

Initial results of an analytic model optimised on JU2003 trials data have been presented. Whilst encouraging results showing agreement between the model and observations have been achieved, further assessment against a wider range of scenarios is required. The next stage is to assess the model in a more sophisticated Bayesian STE procedure as a pre-processor for STE based on more detailed atmospheric dispersion models.

REFERENCES

Allwine, K.J. and Flaherty, J.E. (2006). Joint Urban 2003: study overview and instrument locations. Pacific Northwest National Laboratory report PNNL-15967, August 2006.

Bieringer, P.E., Rodriguez, L.M., Vandenberghe, F., Hurst, J.G., Bieberbach, G., Sykes, I., Hannan, J.R., Zaragoza, J. and Fry, R.N. (2015). Automated source term and wind parameter estimation for atmospheric transport and dispersion applications. *Atmospheric Environment*, 122, 206-219.

Bisignano, A., Mortarini, L., Ferraro, E., & Alessandrini, S. (2013). Offline approach for higher order concentration moments. *Proc. of the 15th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, HARMO 2013*, (pp. 518-523)

Chang, J.C. and Hanna, S.R. (2004). Air quality model performance evaluation. *Meteorology and Atmospheric Physics*, 87, 167-196.

Clawson, K. L., Carter, R. G., Lacroix, D. J., Biltoft, C. A., Hukari, N. F., Johnson, R. C., et al. (2005). *Joint urban 2003 (JU03) SF6 atmospheric tracer field tests.* NOAA Technical Memorandum OAR ARL-254, National Oceanic and Atmospheric Administration, Office of Oceanic and Atmospheric Research.

Hanna, S. and Baja, E. (2009). A simple urban dispersion model tested with tracer data from Oklahoma City and Manhattan. *Atmospheric Environment*, 43, 778-786.

Herring, S. and Huq, P. (2017). Assessing the performance of atmospheric dispersion models. (in preparation).

Luhar, A.K. (2011). Analytical puff modelling of light-wind dispersion in stable and unstable conditions. *Atmospheric Environment*, 45, 357-368.

Raftery, A.E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133:1155–1174.

Robins, P., Rapley, V.E. and Green, N. (2009). Realtime sequential inference of static parameters with expensive likelihood calculations. *Applied Statistics*, 58(5), 641-662.

Storwold, D.P, 2007: Detailed Test Plan for the Fusing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT 07), West Desert Test Center, U.S. Army Dugway Proving Ground, WDTC Document No. WDTC-TP-07-078

Thomson, D.J. and Manning, A.J. (2001). Along-wind dispersion in light wind conditions. *Boundary-Layer Meteorology*, 98, 341-358.

Webb, A.R. and Westoby, P.B. (2017). Analytical modelling of dispersion and its assessment within Bayesian Source term estimation. (in preparation).

Westoby, P.B. Delle Monache, L., & Silk, D.S. (2017). Coupling adjoint and Bayesian approaches for source term estimation. *Proc. of the 18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, HARMO 2017.*

Yee, E. (2012). Inverse dispersion for an unknown number of sources: model selection and uncertainty analysis. *International Scholarly Research Network ISRN Applied Mathematics*, Volume 2012.