Locating a compact odor source using a four-channel insect electroantennogram sensor

2011 Bioinspir. Biomim. 6 016002

(http://iopscience.iop.org/1748-3190/6/1/016002)

Bioinsp. Biomim. **6** (2011) 016002 (16pp) doi:10.1088/1748-3182/6/1/016002

Locating a compact odor source using a four-channel insect electroantennogram sensor

A J Myrick and T C Baker

Chemical Ecology Laboratory, Department of Entomology, Pennsylvania State University, University Park, PA 16802, USA

Received 3 July 2010
Accepted for publication 15 November 2010
Published 15 December 2010
Online at stacks.iop.org/BB/6/016002

Abstract
Here we demonstrate the feasibility of using an array of live insects to detect concentrated packets of odor and infer the location of an odor source (∼15 m away) using a backward Lagrangian dispersion model based on the Langevin equation. Bayesian inference allows uncertainty to be quantified, which is useful for robotic planning. The electroantennogram (EAG) is the biopotential developed between the tissue at the tip of an insect antenna and its base, which is due to the massed response of the olfactory receptor neurons to an odor stimulus. The EAG signal can carry tens of bits per second of information with a rise time as short as 12 ms (K A Justice 2005 *J. Neurophiol.* **93** 2233–9). Here, instrumentation including a GPS with a digital compass and an ultrasonic 2D anemometer has been integrated with an EAG odor detection scheme, allowing the location of an odor source to be estimated by collecting data at several downwind locations. Bayesian inference in conjunction with a Lagrangian dispersion model, taking into account detection errors, has been implemented resulting in an estimate of the odor source location within 0.2 m of the actual location.

1. Introduction

The development of methods for locating the sources of volatile compounds has many applications, including environmental monitoring, security, and drug enforcement. The source localization problem is a complex one that depends on proper modeling of the flow of the medium around the source. As a result, the problem of estimating the location of an odor source has been an active area of research [1–14]. Two main approaches to locating odor sources have appeared in the literature: robotic approaches, addressed here, and those that employ distributed arrays of static sensors. Odor localization studies based on distributed odor sensors utilize the expected values of concentration averaged over long periods of time. While distributed array methods generally do not use anemometric measurements, in robotic applications many times velocity measurements of the transporting medium are made available because of the closer range involved. This difference makes the use of high time resolution odor sensors such as the electroantennogram (EAG) critical to the efficient use of the velocity measurements. The EAG is the electrical potential measured at the severed tip of a living insect antenna relative to the base of the antenna or other electrically accessible region of the insect when a physiologically relevant odor is passed over the antenna. High time resolution allows the detector to resolve high concentration parcels (due to what is described as turbulent intermittency) of odor containing air in a plume emanating from a source and to differentiate between parcels of air that have originated at the source and those that have not. See [15] for high-resolution measurements and characterization of intermittent plumes. Robotic approaches to odor source localization with regard to the inverse problem of locating the source from limited data have lagged behind source localization studies that utilize a set of distributed odor sensors. Recent work in the area of source localization using static sensors includes Keats *et al* [1], and later Yee *et al* [2] who use Bayesian inference in conjunction with an atmospheric dispersion model to infer source parameters using distributed chemo sensors in an urban environment. See also Guo *et al* [16] who have extended these methods to unsteady conditions. Some work has also been done on strategic placement of sensors [17], which

has applications to robotic source localization. While these studies utilized detailed dispersion models, they are meant to be applied over a larger time scale than a robotic entity might use.

The robotic approach to source localization has been an active area of research since the early 1990s. As a result of turbulent intermittency, many localization algorithms under turbulent conditions have focused on impulsive approaches such as moving upwind upon detection of the source odorant. See for instance [3–5]. The EAG has been used as a detector to guide robots that mimic moth behavior toward an odor source [7, 8]. Also, recently, chemosensors have been used to guide a robot toward an odor source [9], as well as photo-ionization detectors [10]. A recent review on robotic odor localization includes references to many more of these interesting approaches [6]. Accurate estimates of the probability density of source parameters (such as location and strength) allow a robotic planner to better strategize its movements. A model also allows data collection locations to be chosen to be far away from an odor source for standoff detection of dangerous materials such as explosives [18]. Several authors [11–14] have begun development of statistical models and search methods for the estimation of the location of an odor source. Pang and Farrell [12] used a model of particles emanating from a source in reverse time in conjunction with Bayesian inference to create a posterior density function for the location of an odor source. Their model assumed that the fluid environment moved as a unit with the measured velocity over the sensor in a strictly two-dimensional problem and lacked documented model parameters. In another similar study, Jakuba *et al* [19] created a forward model for the probability of detecting intermittent strands without velocity measurements of surrounding underwater fluid which was then used to fill an occupancy grid map, or the posterior probability density of source locations.

Here, we use an approach to model turbulent dispersion that is based on the Langevin equation. The Langevin equation is commonly used in what are known as Lagrangian atmospheric dispersion models to model odor parcel movement. Lagrangian dispersion models have been primarily used to model atmospheric transport of pollutants or other material by tracking the position of individual particles. Predictions of concentration fields usually employ Monte Carlo techniques by simulating the position of many particles. See [20, 21] and references therein for more examples of these types of models. However the same model may be used to infer source parameters by reversing time. Reverse-time Lagrangian models were first used to quantify source strength by Flesch *et al* [22]. However like other distributed sensor studies, the model relied on averaged concentrations without anemometric measurements.

Here, we use Bayesian inference to quantify, based on an atmospheric model, the uncertainty in the source parameter estimates. To utilize the extra information available from high-time resolution measurements, time-dependent solutions to the Langevin equation are used in reversed time to compute the probability density of where a parcel with a measured velocity might have originated. This probability density is combined with Bayesian inference to estimate the location of a single source in an open field. This provides a more informative model than has been previously used to model odor dispersion and whose parameters are based on atmospheric measurements, allowing the source to be located through triangulation.

In the problem of detecting a signal in the presence of noise, there is always the possibility of making decisions that a signal is present when it is not. These incorrect decisions are known as false alarms. We take the effect of these false alarms of odor detection into account by incorporating an unknown probability of false alarm into the model and inferring its value using Bayesian inference. False alarms tend to reduce the effective range of the detector, where distant sources may give rise to less detections than the false alarm rate itself.

A valuable tool for the development of methods requiring sensitive and high bandwidth odor detection is the EAG. The use of the EAG measurement was first published in 1957 [23] and has since been used for many purposes, including the investigation of insect behavior and identification of insect pheromones and host odors for pest management. The values of the EAG for locating odor sources are its high speed and sensitivity, the same properties that insects rely on to locate food and mates. For instance, in moths, the EAG's rapid rise time of approximately 15–75 ms in response to turbulent plumes of pheromonal components was measured in two moth species [24]. The same study estimated that the information carrying capacity of those moth EAGs ranged from about 18 to 37 bits s^{-1}, depending on species and pheromonal component. In *Drosophila* [25], EAG channel capacities in excess of 50 bits s^{-1} were estimated. High olfactory channel capacity is evidently important to some insects for relaying accurate high time resolution concentration measurements as a function of time.

EAG potentials, on the order of -1 mV at 1 MΩ, are likely the result of the summation of many receptor neuron activities that result when an odor is passed over the antenna [26]. Note also there is a net movement of charge from the small amount of extracellular fluid into a large number of rapidly firing olfactory neurons. Some empirical models have been put forward to describe the non-linear dose–response relationship [27] of EAGs. Dynamically, at frequencies above 1 Hz in *Drosophila*, the EAG voltage seems to be well approximated by a linear low-pass filter at an operating point [28].

Although used only for the detection of one compound, here we have employed EAG recordings from multiple live insects' antennae which can be used in an 'artificial nose' configuration. The artificial nose approach has previously been proposed for the use in the odor localization problem [29, 30]. The artificial nose was first proposed by Persaud and Dodd in 1982 [31] and has since been utilized in many applications including, among others, environmental monitoring, medical diagnostics, food quality assessment, explosives detection, and fragrance assessment [32]. For reviews of sensors and pattern recognition algorithms applied to artificial olfaction, see [32–35]. The EAG response of different species of moths and insects exhibit fast, broadly tuned responses to different odorants [36–38] and thus an array

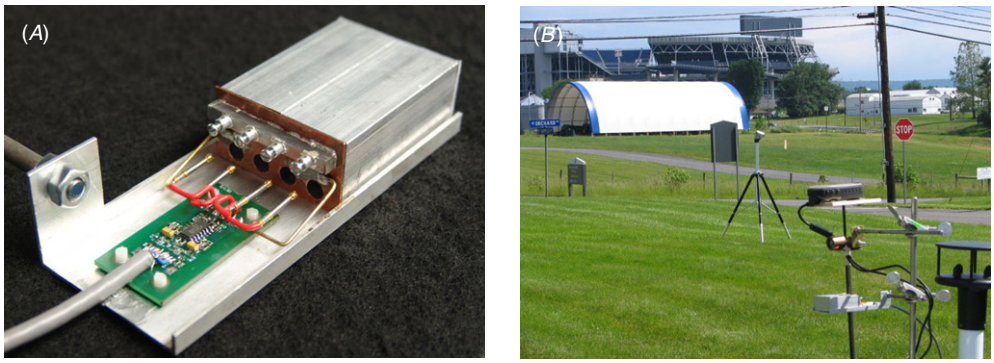

Figure 1. (*A*) Photograph of the four-channel EAG pre-amplifier and holder for live moths. (*B*) Representative photograph of an experimental setup showing the pheromone source (attached to the tripod), the GPS, the EAG sensor and the anemometer.
(This figure is in colour only in the electronic version)

of antennae from different species [37] or different sensory neurons [39] may be used in an artificial nose configuration. It has been shown that a four- or eight-channel array of excised insect antennae is useful for both odor detection and recognition [40]. When recorded from live insects, consistent EAG responses may be obtained for longer periods of time [36], up to 8 h in aphids [38] and several days in moths such as *Helicoverpa zea* and *Trichloplusia ni* [41].

2. Methods

2.1. Hardware

To demonstrate the utility of the methods outlined in this paper in an outdoor environment, a pheromonal odor source whose location was to be inferred was placed in an open grassy field. Instrumentation placed on a hand-pushed mobile cart included a four-channel antennal sensor an Airmar PB200 WeatherStation and a Gill WindSonic ultrasonic anemometer. Included in the weatherstation is a 2D ultrasonic anemometer that measures wind speed and direction, a WAAS-enabled GPS unit and a three-axis digital compass (see figure 1). The anemometric measurements obtained from the PB200 in its current proprietary configuration lacked sufficient information carrying capacity and therefore anemometric data were obtained from the Gill anemometer.

To record from live insects, an insect holder was constructed. Live insects were immobilized in tapered, aerated plastic tubes and placed in a custom four-channel preamplifier, shown in figure 1(*A*). With an electrolytically sharpened tungsten electrode inserted into the eye as a ground reference, the antennae (tips intact) were draped over the amplifier electrodes using electroconductive gel (Spectra 360, Parker Laboratories Inc., USA) to establish a connection to the amplifier.

As described in [40], the antennal sensor comprises a low noise, high impedance four-channel CMOS pre-amplifier connected through a second-stage amplifier with first-order high- (0.15 Hz) and low-pass (34 Hz) filters. The total in-band gain was 110 V/V, effectively swamping any digitizer noise. EAG data were digitized for storage using the National Instruments DAQCard-6036E into a Dell Inspiron 8200 Laptop PC running Labview 6.1. Data were sampled at 50 kSa s^{-1}, filtered using a 25 Hz Butterworth low-pass filter (20 pole IIR) and decimated to 100 Sa s^{-1}.

Data from instruments were collected in synchrony using a laptop PC. Communications between the PC and the PB200 were established over an RS-232 interface. Measurements from the compass, GPS and anemometer were sent via NMEA 0183 messages to the PC at a rate of 10 Hz. Wind velocity data from the Gill WindSonic anemometer were sent to the PC over a second RS-232 connection. This anemometer samples at 20 Sa s^{-1} and block averages to 4 Sa s^{-1}. GPS, compass, and anemometer data were recorded with timestamps that correspond to the EAG sensor sample number (i.e. referenced to the acquisition card sample clock).

2.2. Experimental setup

Experimental measurements were made on 10/08/09 starting at 2:36 PM eastern standard time in State College, PA. The sensor cart was moved to 13 locations for approximately 8 to 13 min each during recording. While standing at location 0, background EAG activity was recorded (source located downwind). Following this, the source was waved upwind from the sensor to cause EAG depolarizations for the purpose of collecting training data for the classifier. Subsequently, each recording location was generated from a normal density centered 15 m downwind from the source with a standard deviation of 10 m in the downwind direction and 5 m in the crosswind direction. The mean downwind direction was calculated using the last 10 min of anemometer data prior to moving from the previous location. Navigation to these locations was accomplished using the GPS, however its accuracy was not sufficient to be utilized in the determination of the source location. When the randomly generated location was within about 1 m of a previously used location, that location was used instead because the GPS lacked the navigation accuracy to get to the new location. Instead of using GPS measurements, distances between markers at each location were measured using a laser rangefinder (Craftsman model 320.48277) so that the locations could later be determined accurately. Recording from location 1 was started at 3:05 PM, and recording while standing at the last

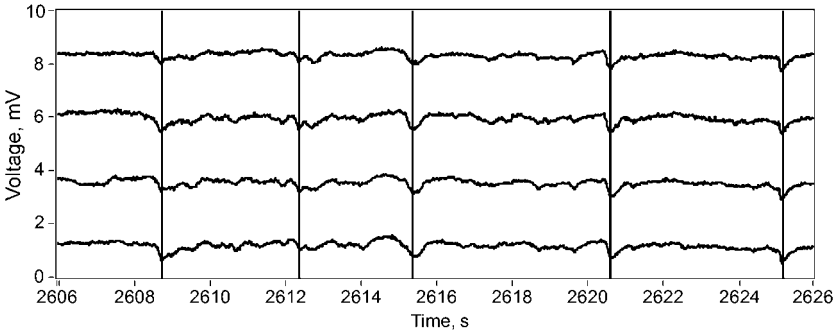

Figure 2. EAG recordings (prior to filtering) obtained while standing at location 2 (low p_{FA} analysis). EAG channels are 1–4 from bottom to top. Vertical lines indicate times when the classifier made the decision that pheromone had caused EAG depolarizations.

location ended at 5:45 PM. Using Turner's stability criteria [42], weather conditions were mildly unstable, changing over to neutral at 5:15 PM.

The EAG was simultaneously recorded from four moths of two different species for the purpose of detecting the presence of (Z)-11-hexadecenal (Z11–16:Ald). Channels 1 and 2 each measured signals from two male *Heliothis virescens* (tobacco budworm), while channels 3 and 4 measured responses from two male *H. zea* (corn earworm). Both are insects whose female counterparts incorporate Z11–16:Ald as the major pheromone component. Two different species were used on the basis of their availability from our insect colonies. Because they are very similar species, significant differences in the EAGs obtained was not expected or observed. References documenting pheromonal components present in the female glands of moth species used in this study follow: *H. zea* [43] *H. virescens* [44].

The outdoor source consisted of a bundle of 100 plastic drinking straws whose surfaces had been coated with Z11–16:Ald by shaking in a solution of hexane (∼50 ml) and Z11–16:Ald (200 mg) and allowing the hexane solvent to evaporate. The source was oriented so that the straws were parallel to the average wind direction and placed at the height of the EAG sensor, 1.5 m.

2.3. Preliminary data processing

Raw EAG sensor data consist of voltage versus time records from the four-channel probe sampled at 100 Sa s^{-1}. These records are later loaded from disk and passed through a 0.8–5 Hz (−3 dB) finite impulse response (FIR) symmetric linear phase bandpass filter to remove noise and baseline drift in the EAG signal. Brownian-type noise and more thoroughly mixed background odors especially contribute to low frequency components of the EAG signal, which are removed by this filter. Feature extraction and subsequent detection of possible depolarizations were performed using the stage 1 classifier described in [40]. The false alarm rate was manipulated by adjusting the background prior probability. The representative results of the detection procedure while standing at location 2 (ca 10 m from the source) are shown in figure 2.

Output from the detector was then combined with data from the Gill WindSonic anemometer and location data from the GPS resulting in a sequence of random vectors with timing governed by the anemometer sampling interval, T. Each 1.0 s interval includes an average anemometric velocity, $\mathbf{u}(kT)$ a detector location $\mathbf{r}_D(kT)$, and a detection measurement $D(kT)$. Two-dimensional anemometric measurements (north/south and east/west) $ns\,(kT)$ and $ew\,(kT)$ were transformed into components $u_1(kT)$ and $u_2(kT)$, which point in the mean wind direction and the horizontal crosswind direction, respectively. If one or more 'hits' (strand of concentrated odor passes over the probe) occurs during an interval, the variable in the random process $D(kT)$ is assigned a value of 1 otherwise $D(kT)$ is assigned a value of 0. Although the GPS readings were recorded, locations calculated using the laser range finder were used to record the detection locations, $\mathbf{r}_D(kT)$. The recorded sequences are designated as $\mathbf{u}_k$, $\mathbf{r}_{Dk}$ and D_k throughout this manuscript.

3. Dispersion model

The goal is to estimate the source location utilizing the data sequences described in section 2.3. It is shown, under certain assumptions, in section 4 that to use Bayesian inference to solve for the source location, a probabilistic model for the dispersion of odor containing parcels of air from the source location to the anemometer/detector is necessary. This is described by the probability density $f_D\,(D_k\,|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}\,)$. This is the probability density of obtaining a detection (or no detection) at location $\mathbf{r}_{Dk}$ during the kth time interval from a parcel originating at the source location, $\mathbf{r}_S$, given the wind velocity measurement, $\mathbf{u}_k$, made at the detection location. The parameter vector $\boldsymbol{\theta}$ describes the size of the odor parcels and probability of false alarm, which are also unknown. The way this expression is obtained is described in sections 3.2 through 3.6. Note that the probability of detection is dependent on only one velocity measurement. This is due to the representation of the movement of air parcels with the Langevin equation for which only one velocity measurement gives all the information necessary for estimation of a parcel's location. Solving the Langevin equation results in an expression for the probability density that a parcel intersects location $\mathbf{r}_I$ at time t given its velocity was $\mathbf{u}_0$ at location $\mathbf{r}_0$ at time 0, or $f_{\mathbf{R}_I}\,(\mathbf{r}_I\,|\mathbf{u}_0, \mathbf{r}_0, t\,)$. The solution to the Langevin equation is dependent on atmospheric parameters whose estimation is described in section 3.4. In section 3.5,

an expression is developed for the probability of obtaining a *positive* $\mathrm{Pr}_P\,(P = 1\,|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta})$, rather than detection during some finite interval. Incorporation of detection errors to obtain the final expression for $f_D\,(D_k\,|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta})$ is then described in section 3.6.

3.1. Lagrangian coordinates

The model employed here is a simple implementation of a Lagrangian dispersion model. Lagrangian techniques are so named because they utilize Lagrangian coordinates, which label the fluid flow elements, rather than fixed locations in space. This is useful in our model because it is the position of the Lagrangian element that passes through the anemometer at some time that we would like to track. Following the notation of Pope [45], $\mathbf{r}^+\,(\mathbf{r}_0, t)$ refers to the Eulerian coordinates, the fluid element which is at the position $\mathbf{r}_0$ at time t_0, where $\mathbf{r}_0 = [r_{01},\ r_{02},\ r_{03}]$ are the Lagrangian coordinates, and $\mathbf{r}^+\,(\mathbf{r}_0, t_0) = \mathbf{r}_0$. This makes $\mathbf{r}^+$ the transformation from Lagrangian to Eulerian coordinates. Each element's velocity may be transformed from Lagrangian to Eulerian coordinates when the Eulerian velocity field is known according to equation (1):

$$\mathbf{u}^+(\mathbf{r}_0, t) = \mathbf{u}[\mathbf{r}^+(\mathbf{r}_0, t), t] \tag{1}$$

where $\mathbf{u}[\mathbf{r},\ t]$ is the (known) Eulerian vector field describing the velocity. The Eulerian velocity of the element initially at $\mathbf{r}_0$ may be integrated in order to track its position:

$$\mathbf{r}^+(\mathbf{r}_0, t) = \int_{t_0}^{t} \mathbf{u}^+(\mathbf{r}_0, \tau)\,\mathrm{d}\tau + \mathbf{r}_0. \tag{2}$$

3.2. Langevin equation

Among others, Hanna [46] provided experimental evidence that the power spectrum of the Lagrangian velocity in the convective boundary layer decays at a rate of ω^{-2}, which is consistent with the spectrum of an Ornstein–Uhlenbeck (OU) process [47] describing Brownian motion. Solutions to the Langevin equation, known as a stochastic differential equation (SDE), result in OU processes [45, 48]. As a result, Langevin equations were first used by Thomson [49, 50] in forward atmospheric models. A simplification employed here is that the flow is assumed to be homogeneous and uncorrelated in all three Cartesian directions. This is not an unrealistic assumption for the horizontal directions in a large open field. However, flow in the z direction is not homogeneous, due to the boundary at the ground. Further, the velocity in the z direction is not uncorrelated with horizontal velocities.

The Langevin equation is meant to apply to the mean-subtracted velocity process. The process of breaking velocity measurements into a mean part and a fluctuating part is known as the Reynolds decomposition. The mean wind speed can be obtained by averaging as a result of the spectral gap [51, 52] at around 1 h. In this case, the coordinates have been rotated so that the mean wind points in the direction of the mean of u_1. The u_1 component is separated into the sum of a fluctuating part, $\tilde{u}_1$, and a constant, mean wind speed $\bar{u}_1$:

$$u_1\,[\mathbf{r}_D, kT] = \bar{u}_1\,[\mathbf{r}_D, kT] + \tilde{u}_1\,[\mathbf{r}_D, kT]\,. \tag{3}$$

The Langevin equation below, formulated for atmospheric models, describes a Cartesian component of the motion of the Lagrangian fluid element $\mathbf{r}_0$ [45]:

$$\mathrm{d}\tilde{u}_j^+(\mathbf{r}_0, t) = -\frac{\tilde{u}_j^+(\mathbf{r}_0, t)}{T_{Lj}}\,\mathrm{d}t + \sigma_{uj}\sqrt{\frac{2}{T_{Lj}}}\,\mathrm{d}W_j(t) \tag{4}$$

where $\mathrm{d}W(t)$ may be thought of as a time-independent zero-mean Gaussian random variable with a variance of $\mathrm{d}t$. σ_{uj} is the standard deviation of the Lagrangian wind velocity in the direction indexed by j, and T_{Lj} is a time constant indicative of the amount of memory the process has, is known as the Lagrangian integral time scale.

The Langevin model, besides being a reasonable representation of reality, has the desirable property that it is Markov, meaning future states are dependent on only the present state of the process. This means future values of the velocity are estimated using only its present value. This simplifies our analysis because we wish to estimate the position of fluid elements that pass through the anemometer at some time.

Although wind speed measurements are available in horizontal directions, no measurement is available in the vertical direction. Therefore, estimates of the position of a parcel that passes through the anemometer in the horizontal directions differ from that in the vertical direction. In the vertical direction, when only the position and not the velocity of the element is known, its mean displacement is simply the integral of the mean velocity (assuming $t_0 = 0$):

$$E\left[r_3^+(r_{03}, t)\middle|r_{03}\right] = \bar{u}_3^+ t + r_{03}. \tag{5}$$

In this case, $\bar{u}_3^+ = 0$. The variance of an element's displacement as a function of time for an OU process (i.e. the Langevin equation) without knowing its velocity was given by Taylor [53]:

$$E\left[\left(r_3^+(r_{03}, t) - E\left[r_3^+(r_{03}, t)\right]\right)^2|r_{03}\right] = 2\sigma_{u3}^2 T_{L3}^2\left[\frac{t}{T_{L3}} - \left(1 - \mathrm{e}^{-\frac{t}{T_{L3}}}\right)\right]. \tag{6}$$

In the horizontal direction, when both the position and velocity of the fluid element are known at $t_0 = 0$, the mean trajectory is given by [47]

$$E\left[r_j^+(r_{0j}, t)|r_{0j}, u_j^+(r_{0j}, 0)\right] = \tilde{u}_j^+(r_{0j}, t_0)T_{Lj}(1 - \mathrm{e}^{-t/T_{Lj}}) + \bar{u}_j^+ t. \tag{7}$$

The variance of the element's position when both its position and velocity are known at $t_0 = 0$ is given by [47]

$$E\left[\left(r_j^+(r_{0j}, t) - E\left[r_j^+(r_{0j}, t)\right]\right)^2\middle|r_{0j}, u_j^+(r_{0j}, 0)\right] = 2\sigma_{uj}^2 T_{Lj}^2\left[\frac{t}{T_{Lj}} - 2(1 - \mathrm{e}^{-t/T_{Lj}}) + \frac{1}{2}(1 - \mathrm{e}^{-2t/T_{Lj}})\right]. \tag{8}$$

Because the $\mathrm{d}W(t)$ is Gaussian, there is now enough information to describe the solution to our model of an odor parcel's movement. The probability density that an infinitesimally small parcel of air intersects the location $\mathbf{r}_I$

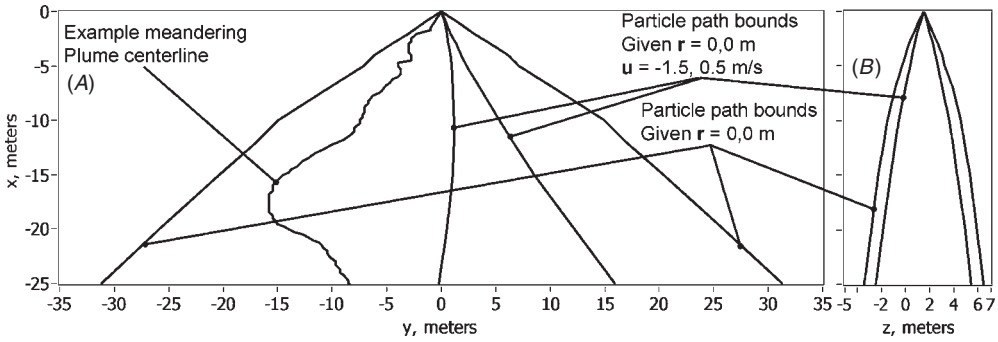

Figure 3. (*A*) Horizontal extent of average plume concentrations under two conditions (see the text for details). The wider plume shows that under our experimental conditions, the plume occupied a wide downwind area. The narrower plume is illustrative of the additional information provided by horizontal anemometric measurements under the model. A snapshot of what the centerline of a more coherent meandering plume within the envelope given velocity measurements is also shown, emphasizing the effect of the spatial correlation on particle dispersion. (*B*) The vertical extent of the plume under both conditions is shown and it is noted that the simple dispersion model extends below ground level.

t seconds after intersecting $\mathbf{r}_0$ ($\mathbf{r}^+(\mathbf{r}_0, t) = \mathbf{r}_I$) is given by the following equation:

$$f_{\mathbf{R}_I}(\mathbf{r}_I|\mathbf{u}_0, \mathbf{r}_0, t) = \prod_{j=1}^{3} \frac{1}{\sqrt{2\pi}\sigma_{Ij}(t)} \times \exp\left[-\frac{1}{2}\left(\frac{r_{Ij} - r_{0j} - \mu_{Ij}(t, u_j)}{\sigma_{Ij}(t)}\right)^2\right]. \quad (9)$$

The means and standard deviations are the Ornstein and Uhlenbeck conditional estimates in the jth direction for the horizontal components ($j = 1$ and 2), and the Taylor estimate in the z direction ($j = 3$), because no velocity measurement is available in that direction. Equations for the arguments are given by

$$\mu_{Ij}(t, u_{0j}) = \tilde{u}_{0j} T_{Lj}(1 - \mathrm{e}^{-t/T_{Lj}}) + \bar{u}_{0j} t \qquad j = 1, 2 \quad (9a)$$

$$\mu_{Ij}(t, u_{0j}) = \bar{u}_{0j} t = 0 \qquad j = 3 \quad (9b)$$

$$\sigma_{Ij}^2(t) = 2\sigma_{uj}^2 T_{Lj}^2 \left[\frac{t}{T_{Lj}} - 2(1 - \mathrm{e}^{-t/T_{Lj}}) + \frac{1}{2}(1 - \mathrm{e}^{-2t/T_{Lj}})\right] \qquad j = 1, 2 \quad (9c)$$

$$\sigma_{Ij}^2(t) = 2\sigma_{u3}^2 T_{L3}^2 \left[\frac{t}{T_{L3}} - (1 - \mathrm{e}^{-t/T_{L3}})\right] \qquad j = 3. \quad (9d)$$

Equation (9) estimates the position of a particle at time based entirely on its position and velocity at $t = 0$ and does not incorporate any information that may be available at other times. Note that in the vertical direction ($j = 3$), the exponent in equation (9) is practically always 0 in the vertical plane of the source, so that only the normalization constant is affected as time progresses.

The model given by equation (9) indicates that under the low wind speed weather conditions recorded during our experiment, velocity information was highly informative in determining the path of a particle, while a downwind area with a large angular range was available for making measurements. Consider a point source releasing particles continuously at a constant rate. Equation (9) may be integrated over time ($t = 0$ to ∞) to predict the resulting average concentration of particles in space given the conditions under which the particles were released. Or it may be assumed that the source has released a single particle at some unknown time with a constant prior probability. Then equation (9) may be integrated over (some limited) time to determine the probability density of finding that particle somewhere in space. With either interpretation, the time integral provides useful information about the bounds of the plume. Figure 3 was included to illustrate the bounds of particles following the Lagrangian paths modeled by this equation under the experimental weather conditions encountered. The time integral of equation (9) was numerically computed in several y–z planes to obtain approximate plume bounds under two conditions. The first condition is that the velocity of the particle as it passes through $\mathbf{r}_0 = (0, 0, h)$, where h is the height of the source, is unknown. Under this condition, the means and standard deviations are the same as assumed in (9*b*) and (9*d*) in all three directions. In the second condition it is assumed that the horizontal velocities are known thus the integral is computed with the assumptions that (9*a*–9*d*) hold. Under both conditions, the integral of the function in the y–z plane was very nearly constant as a function of x, indicating that most particles released are eventually carried by the mean wind into the next y–z plane. The isoprobability contour containing 68% of the density in each y–z plane was then computed. Figure 3 indicates the horizontal and vertical extents of these isoprobability contours in the x–y plane under both conditions. It should be noted that the modeled boundaries are *average* values under both conditions and do not represent the plume structure at some time. Movements of particles are correlated in space as well as time giving rise to meandering plumes within an average envelope given by the velocity process at the source. The centerline of an example of a meandering plume is also drawn in figure 3. Consideration of spatial correlation leads to the two-particle dispersion problem, on which initial work was simultaneously published by Brier [54] and Batchelor [55]. For a review, see [56].

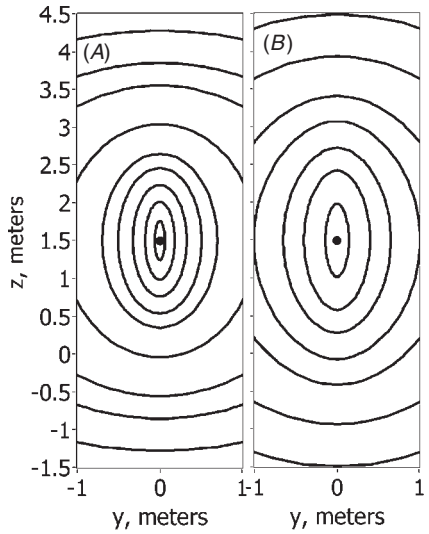

Figure 4. Isoprobability contours containing 68% of y–z plane densities computed at various values of x. (A) $u_1 = -1.5$ m s^{-1}, $u_2 = 0$ u_3 unknown. (B) u_1 unknown, $u_2 = 0$ m s^{-1}, u$_3$ unknown. Values of $-x$ are 1, 2, 3, 4, 5, 7, 10, 12 and 15 m. Dark circle in the middle is drawn to indicate the cross-sectional area of the actual source used.

Equation (9) applies to a point source during an infinitesimal time interval. The actual source used was 10 cm in diameter, which was unmodeled. Figure 4 has been included to illustrate the size of the source in comparison to the (68%) plume boundaries corresponding to the point source assumption. When the horizontal velocities take on the most likely values in panel A ($u_1 = -1.5$ m s^{-1}, $u_2 = 0$, u_3 unknown), the source size becomes insignificant subjectively at about 3 m downwind. But for a data collection location located along the x axis where a particle passes through the source with $u_2 = 0$ at any speed, the source size becomes insignificant at perhaps 2 m downwind. Because the data collection locations were all significantly further from the source than 2–3 m, it is doubtful that the Lagrangian dispersion model was significantly affected by the finite cross-sectional area of the source in this case. It should be noted, however, that the properties of the concentration variations passing over the detector were likely affected by the physical size of the source, likely resulting in wider concentration peaks than for a smaller more concentrated source of equal odorant emission rate.

3.3. The parcel and source model

Average velocity measurements were made over fixed and regular time intervals, during each of which many (as modeled by equation (9)) infinitesimal particles passed through a detector. Some of these particles may have intersected the source and some may not have. However, because the particles pass through the detector in close spatial proximity, movements of each of these processes are correlated, as a result of spatial correlation. Given the high spatial correlation of particles that are close to each other, those particles that pass through the detector that have not intersected the source but are in close proximity to odorant particles at the detector are more likely to have passed close by the source than those that are not in close spatial proximity to odorant particles at the detector. Thus when a detection is made, it is an indication that the particles, only whose average velocity is known, have likely passed through some region around the source, making the information contained in concentration relevant. Here we have taken the detection of any amount of odorant during an interval as an indication that the 'parcel' associated with that average velocity measurement has intersected the source point. However, in our detection scheme, detection times are associated with local concentration maxima, which means that parcels containing odorant that pass over the detector over two or more adjacent time intervals will be assigned to a single interval, effectively lumping odorant into a single infinitesimal element. This operation can tend to under-report positive detections, especially when short intervals are used.

To facilitate computation of the probabilities of obtaining positive detections, the object that is transported by the wind at unresolved scales is viewed as a coherent strand of points or parcel. Each element is assumed to share the same realization of the Langevin equation (sharing the same velocity measurement and Brownian motion, $W(t)$) shifted only in time. After the first point on the strand passes through a point in space, the rest follow through the same point for exactly one time interval, T. A single parcel is released from the source per time interval at regular intervals. Further, assuming that each parcel occupies the same volume it will stretch under high velocity conditions and compress under low velocity conditions. Each parcel has one plane (associated with detector peaks) that contains odorant. The frontal area of the parcel is used to model the source strength, but may also reflect the physical size of the source. For instance, a source with a smaller cross section and equal emission rate might be detected less often, especially near the source. The size of the source could, in more accurate models, become an important parameter. Although the meandering plume mixes and dilutes as it is transported by the wind, likely affecting its detectable cross-sectional area as a function of downwind distance, we have not initially modeled this effect. The primary reason for this is to reduce the number of source parameters. In our data from this experiment we have not found any clear trend in size of the parcels as a function of distance from the source.

It is also evident from recordings that odorant concentration peaks tend to be longer farther downwind. For instance from location 2 ($\sim$10 m from the source) peaks subjectively average about 0.25 s in width, while from location 10 ($\sim$20 m from the source) peaks are about 1 s in width. Thus it is expected that at close range, parcels are more heterogeneous than from farther distances.

Here, as in many Lagrangian dispersion models, each parcel is modeled as a new and independent realization of the Langevin equation. However, like the particles, the parcels being modeled are located in close spatial proximity, and movements of each of these processes are also correlated being function of weather conditions, including average wind speed. To consider spatial correlation, the many different combinations in the order in which the source parcels may cross the detector become more important. However, given

a sufficient number of measurements, the effect of spatial correlation should become less relevant. It is observed in our data (not shown) that detections are associated with time intervals where a high rate of intercept is predicted, but generally when detections are made, they exceed the predicted rate. At other times no detections are visible where the rate is predicted to be high. This is likely the result of the meandering plume moving within its predicted envelope, possibly averaging out over time. Another result of modeling each parcel independently is that any number of these continually released parcels may appear at the detector at one time.

3.4. Estimation of Lagrangian model parameters

Parameters to be estimated for the solution of (9) include the standard deviation of the Lagrangian wind velocity, σ_{uj}, and the integral time scale T_{Lj}, where the direction is enumerated by j. These parameters were estimated from anemometer data that were collected between 3:05 and 5:45 PM, i.e. while collecting data at locations 1–13. Because wind velocity measurements are available in the horizontal directions, horizontal estimates of these parameters differ from vertical estimates. In the horizontal directions, σ_{uj} is usually considered equal to the Eulerian value [57] given by

$$\sigma_{uj} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} \tilde{u}_j^2}. \tag{10}$$

Several authors have attempted to provide a method for estimating Lagrangian integral time scales, T_L, from Eulerian velocity measurements obtained at a fixed point in space. T_L is defined by the following equation [46]:

$$T_L = \int_0^\infty R_{uu}(t)\,\mathrm{d}t \tag{11}$$

where $R_{uu}(t)$ is the normalized autocorrelation function of the velocity of a Lagrangian element. The OU process being used to model parcel movement has an autocorrelation described by the following equation:

$$R_{uu}(t) = \mathrm{e}^{-\frac{|t|}{T_L}}. \tag{12}$$

Complicating the effort to estimate T_L is that the measured Eulerian velocity process at a fixed point in space is not an OU process and has long autocorrelation tails that are difficult to estimate from measurements. Hanna [46] has been frequently referenced by authors documenting Lagrangian dispersion models. Rather than integrating the Eulerian autocorrelation function using an equation analogous to (11), Hanna takes the time in account at which the Eulerian autocorrelation function reaches the value of e^{-1} for the Eulerian time scale. A common method [57] is to use a constant, β, to relate the Eulerian time scale obtained in this way to the Lagrangian time scale:

$$T_{Lj}/T_{Ej} = \beta_j \tag{13}$$

where

$$\beta_j = 0.44\frac{\bar{u}_1}{\sigma_{uj}} \tag{14}$$

and $\bar{u}_1$ is the mean wind speed.

The following approach was used to estimate the Lagrangian time constant and wind velocity standard deviation in the z ($j = 3$) direction without velocity measurements. The formulae used here are taken from other forward Lagrangian dispersion models [58, 59]. In unstable conditions, when the Obukhov length, L, is negative, the Lagrangian time constant in the z direction can be approximated by

$$T_L = \frac{0.5 r_3}{\sigma_{u3}}\left(1 - 6\frac{r_3}{L}\right)^{\frac{1}{4}}. \tag{15}$$

Further, in the direction of the mean wind,

$$\sigma_{u1} = u_*\sqrt{4 + 0.6\left(\frac{z_i}{-L}\right)^{2/3}} \tag{16}$$

where u_* is known as the friction velocity (for a description see [51]) and z_i is the boundary layer thickness. The friction velocity may be measured directly, or estimated from horizontal velocity measurements at several vertical locations. The friction velocity usually ranges from about 0.1 to 0.4 m s^{-1}. Here we have assumed a value of 0.2 m s^{-1}. The boundary layer thickness can vary quite a bit. Our assumption of 1000 m is typical [51] and is the same as that assumed by the authors in [58]. In any case, the time constant is insensitive to the boundary layer thickness. The value for L obtained from (16) is substituted into (17):

$$\sigma_{u3} = 1.25 u_*\left[1 + 3\frac{r_3}{-L}\right]^{1/3}. \tag{17}$$

The result is substituted into (15). Setting the height of the anemometer, $r_3 = 1.5$ m results in a value of 3.1 s for T_L and 0.27 m s^{-1} for σ_{u3}.

3.5. Probability of obtaining a positive for a single parcel during a finite interval

Equation (9) is valid when the velocity is sampled during an infinitesimal time interval for an infinitesimally small parcel. However, it is necessary for the maximum likelihood procedure to obtain an expression for the probability of obtaining detection during the kth time interval rather than an infinitesimal interval. In addition, the (average) velocity measurement is made at the detector, rather than at the source, so the particle path being estimated occurs prior to the time the velocity measurement is made. Fortunately, we may reverse the direction of time in equation (9) to predict where the parcel came from rather than where it is going. We begin by obtaining an expression for obtaining a positive from one parcel and then make an extension to multiple parcels in section 3.6.

In the absence of detection errors, the actual state of a true positive is indicated by the Boolean variable P. If the source releases a single (lth) parcel from $\mathbf{r}_S$ starting at time $t_{k-l} = t_i$, $\mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}, iT)$ is the probability of obtaining a positive at location $\mathbf{r}_D$ where the average velocity measurement is $\mathbf{u}_k$ during the kth measurement interval. To compute this probability using equation (9), the direction of time may be reversed. The probability of one parcel leaving the detector point $\mathbf{r}_D$, at time t_k with 'instantaneous' (actually

average) velocity $\mathbf{u}_k$, and intersecting the source $\mathbf{r}_S$ (rather than the detector) during the $i = k - l$th interval may be estimated.

An exact solution to this problem would require a full solution to a SDE with more complex boundary conditions for each location and flight time. Note that the probability of an intercept of a parcel can be modified by knowledge of its position during times other than the endpoints of the interval 0 to t. For instance, knowing that in the prior interval no detection occurred would provide information about where no parcels (probably) were during that time. The amount of information this knowledge lends to find the source parameters is considered negligible.

The following paragraph uses simplifying heuristics to provide a tenable approximation to the more complex SDE. There is, at the beginning of a time interval (in forward time), a probability that an odorant containing parcel of finite volume is already over the point detector. However, because the detection method assigns odorant times to concentration peak times of infinitesimal length, the probability of this being reported is zero. As a result, the probability of an intersection occurring during any time interval is approximated as the expected amount of probability density that has been traversed by the frontal area of the odorant-containing plane of the parcel during the time interval. This assumption can be expressed using the following heuristic formulation:

$$\frac{d\,\mathrm{Pr}_P}{\mathrm{d}t} \cong \frac{A_P}{E\left(\left\|\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t}\right\|\right)} f_{\mathbf{R}_I}(\mathbf{r}_S|\mathbf{u}, \mathbf{r}_D, t) E\left(\left\|\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t}\right\|\right). \tag{18}$$

The velocity $\mathrm{d}\mathbf{s}/\mathrm{d}t$, a random quantity, is the velocity the parcel travels given $\mathbf{r}_S$ $\mathbf{u}$, $\mathbf{r}_D$ and t. A_P denotes the frontal area of the parcel at 1 m s^{-1} and the total parcel volume is given by $A_P T$. Fortunately, because the frontal area of the parcel is assumed inversely proportional to its speed to maintain constant volume, $\mathrm{d}\mathbf{s}/\mathrm{d}t$ need not be explicitly calculated. Also, implicit in this equation is the assumption that $E[\mathrm{d}\mathbf{s}]$ never 'backtracks' itself—which is equivalent to assuming that an element of the parcel will never cross the detector twice. Because an element of the parcel may not cross the detector more than once, the only allowed intersections (for a point on the strand) are those that occur during a single infinitesimal interval and not at any other time. Under that simplification, only a simple integration is necessary to find P. Further, if A_P is small enough, then integration over the area need not be carried out. The probability of having an intersection from $t = iT$ to $t = iT + T$ is given by

$$\mathrm{Pr}_P(P = 1|\mathbf{r}_S, \mathbf{r}_{Dk}, \mathbf{u}_k, \boldsymbol{\theta}, iT) \cong A_P \int_{iT}^{iT+T} f_{\mathbf{R}_I}(\mathbf{r}_S|\mathbf{u}_k, \mathbf{r}_{Dk}, t)\,\mathrm{d}t. \tag{19}$$

3.6. Probability of obtaining a positive for multiple parcels

Assume that the source has been releasing parcels since $t = -\infty$. Because the parcels move independently each governed by an independent Brownian motion, more than one parcel may appear over the detector at one time. The probability of having *at least one* parcel in the vicinity of the detector originating from $\mathbf{r}_S$ during the kth sampling interval is

$$\mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}) = 1 - \prod_{i=1}^{\infty}\left(1 - \mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}, iT)\right). \tag{20}$$

Equation (20) can be approximated using the following equation because the summations in equation (21) are always much less than 1 sufficiently far from the detector:

$$\mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}) \cong 1 - \exp\left(-\sum_{i=1}^{\infty} \mathrm{Pr}_P\left(P = 1\,|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}, iT\right)\right). \tag{21}$$

Near the detector, point approximation of the area integrals over the densities may result in probability values that are greater than one. Equation (21) ensures that the total probability does not exceed one during the optimization procedure. Equation (21) can be re-written as an integral using equation (19):

$$\mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}) \cong 1 - \exp\left(-A_P \int_0^{\infty} f_{\mathbf{R}_I}\left(\mathbf{r}_{Dk}\,|\mathbf{u}_k, \mathbf{r}_S, t\right)\mathrm{d}t\right). \tag{22}$$

In general it is not necessary to compute the time integral in (22) using T as the time increment or even to use a Riemann sum.

3.7. Incorporation of detection errors

The probability of detection, rather than a true positive, given the velocity measurement is desired. Detection within an interval is indicated by the Boolean variable D. Using Bayes' theorem,

$$f_D(D|\mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta}) = \int f_D(D|P, \mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta}) \times f_P(P|\mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta})\,\mathrm{d}P. \tag{23}$$

The probability of obtaining a true positive (21) can be used to provide a continuous probability density for P necessary to compute the integral in (23):

$$f_P\left(P\,|\mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta}\right) = \delta(P - 1)\,\mathrm{Pr}_P(P = 1|\mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta}) + \delta\left(P\right)\left(1 - \mathrm{Pr}_P\left(P = 1\,|\mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta}\right)\right). \tag{24}$$

In general, $f_D(D|P, \mathbf{u}, \mathbf{r}_S, \mathbf{r}_D, \boldsymbol{\theta})$, or the probability density of a parcel being detected or not detected given a true positive is a function of $\mathbf{u}$, $\mathbf{r}_S$ and $\mathbf{r}_D$. However for simplicity, the function is defined using constant values. When the true positive rate of the detector is p_D, the false positive rate is p_{FA}, the false negative rate is $1 - p_D$, and the true negative rate is $1 - p_{\mathrm{FA}}$. Positives do not occur as modeled – as discrete uniform impulses. A true positive may be defined in different ways, such as when 1 or more molecules reach a neural receptor during any time interval. As noted earlier, the higher the concentration, the tighter the region around the source the particles that passed through the detector may have intersected. Thus, practically, a concentration threshold could be assigned. Here, very similar results are achieved as long

as the probability of detection is significantly greater than the false alarm rate. As a result, we have simply assumed that the probability of detection is 1. The detection probability density can be written as

$$f_D(D|P,\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta}) = \delta[P-1][\delta(D-1)p_D + \delta(D)(1-p_D)] + \delta[P][\delta(D-1)p_{\mathrm{FA}} + \delta(D)(1-p_{\mathrm{FA}})] \tag{25}$$

where the Kronecker delta function is given by $\delta[.]$ and the Dirac delta function is given by $\delta(.)$. After substitution of (24) and (25) into (23) and integrating, the following result is obtained for substitution into the likelihood function whose development will be described in section 4:

$$f_D(D|\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta}) = \delta(D-1) \times [p_{\mathrm{FA}} + (p_D - p_{\mathrm{FA}})\,\mathrm{Pr}_P(P=1|\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta})] + \delta(D) \times [(1-p_{\mathrm{FA}}) - (p_D - p_{\mathrm{FA}})\,\mathrm{Pr}_P(P=1|\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta})]. \tag{26}$$

Interestingly, (26) may be used to estimate the information in the random process $D(kT)$ conditioned on the model and velocity measurements (not to be confused with conditional entropy). The entropy H, in bits, for a random variable, x, that can take on one of the two values associated with probability p is [60]:

$$H = -p\log_2(p) - (1-p)\log_2(1-p). \tag{27}$$

Further, the entropy of independent variables is simply their sum.

4. Bayesian inference of source location

A probability density of the source coordinates obtained from the time-independent measurements may be obtained by Bayesian inference [61] using the following formulation:

$$f_{\mathbf{R}_S,\Theta}(\mathbf{r}_S,\boldsymbol{\theta}\,|\mathbf{u}_1\ldots\mathbf{u}_N, D_1\ldots D_N, \mathbf{r}_{D1}\ldots\mathbf{r}_{DN}) = \frac{f_{\mathbf{R}_S}(\mathbf{r}_S,\boldsymbol{\theta})\prod_{k=1}^{N} f_{\mathbf{R}_D,\mathbf{U},D}(\mathbf{r}_{Dk},\mathbf{u}_k,D_k|\mathbf{r}_S,\boldsymbol{\theta})}{\int_{\Omega_\theta}\int_{V_S} f_{\mathbf{R}_S}(\mathbf{r}_S,\boldsymbol{\theta})\prod_{k=1}^{N} f_{\mathbf{R}_D,\mathbf{U},D}(\mathbf{r}_{Dk},\mathbf{u}_k,D_k|\mathbf{r}_S,\boldsymbol{\theta})\,\mathrm{d}\mathbf{r}_S\,\mathrm{d}\boldsymbol{\theta}}, \tag{28}$$

where the number of measurement intervals is N. Integration in the denominator would be carried out over the region where the source could be and over the parameter space. To obtain an analytical expression for the estimate of the source location and parameters, the maximum *a posteriori* method is used, where

$$(\hat{\mathbf{r}}_S,\hat{\boldsymbol{\theta}}) = \arg\max_{\mathbf{r}_S,\boldsymbol{\theta}} f_{\mathbf{R}_S,\Theta}(\mathbf{r}_S,\boldsymbol{\theta}|\mathbf{u}_1\ldots\mathbf{u}_N, D_1\cdots D_N, \mathbf{r}_{D1}\ldots\mathbf{r}_{DN}). \tag{29}$$

It is assumed that prior information about the source coordinates and model parameters is vague enough compared to the information obtained from the measurements that it can be considered uniform, but that the domain is not infinite in size. Then it suffices to specify the prior density as some constant:

$$f_{\mathbf{R}_S}(\mathbf{r}_S,\boldsymbol{\theta}) = \text{const.} \tag{30}$$

A similar and related assumption is that the data collection locations, $\mathbf{r}_{D1},\ldots,\mathbf{r}_{DN}$, and velocity, $\mathbf{u}$, are independent of $\mathbf{r}_S$ and $\boldsymbol{\theta}$. This means that whatever the data collection locations and wind velocity measurements are, information about the source location and model parameters that can be inferred from those data alone is sufficiently vague that it may be neglected. Normally, this should be the case, because one would not know whether a source is detectable or in what direction relative to the wind velocity density it lies prior to collecting data. Using these assumptions and applying Bayes' theorem,

$$f_{\mathbf{R}_D,\mathbf{U},D}(\mathbf{r}_D,\mathbf{u},D|\mathbf{r}_S,\boldsymbol{\theta}) = f_{\mathbf{R}_D,\mathbf{U}}(\mathbf{r}_D,\mathbf{u}|\mathbf{r}_S,\boldsymbol{\theta})\,f_D(D|\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta}.) = f_{\mathbf{R}_D,\mathbf{U}}(\mathbf{r}_D,\mathbf{u})\,f_D(D|\mathbf{u},\mathbf{r}_S,\mathbf{r}_D,\boldsymbol{\theta}). \tag{31}$$

After substitution of (30) and (31) into (28) and simplifying, the following posterior density function is obtained:

$$f_{\mathbf{R}_S,\Theta}(\mathbf{r}_S,\boldsymbol{\theta}\,|\mathbf{u}_1\ldots\mathbf{u}_N, D_1\ldots D_N, \mathbf{r}_{D1}\ldots\mathbf{r}_{DN}) = \frac{\prod_{k=1}^{N} f_D(D_k|\mathbf{u}_k,\mathbf{r}_S,\mathbf{r}_{Dk},\boldsymbol{\theta})}{\int_{\Omega_\theta}\int_{V_S}\prod_{k=1}^{N} f_D(D_k|\mathbf{u}_k,\mathbf{r}_S,\mathbf{r}_{Dk},\boldsymbol{\theta})\,\mathrm{d}\mathbf{r}_S\,\mathrm{d}\boldsymbol{\theta}}. \tag{32}$$

Equation (26) may be substituted into (32) to solve for the posterior density of the source parameters. Maximization of this function gives maximum likelihood estimates for the source parameters that are sought.

4.1. Maximization of a likelihood function

Although better optimization methods are available, as a starting point, the parameters r_1, r_2, A_P and p_{FA} that maximize the likelihood function are found by a finite-difference Newton's method [62]. While expressions are available for the first derivatives, expressions for the second derivatives are tedious to obtain, and are therefore approximated. As usual, to maximize the likelihood function, the log is taken first before differentiation:

$$\frac{\partial}{\partial(\mathbf{r}_S,\boldsymbol{\theta})}\ln[f_{\mathbf{r}_S}(\mathbf{r}_S,\boldsymbol{\theta}|\mathbf{u}_1\ldots\mathbf{u}_N, D_1\cdots D_N, \mathbf{r}_{D1}\ldots\mathbf{r}_{DN})] = \mathbf{0}. \tag{33}$$

Substitution of (32) into (33) results in the following summation:

$$\sum_{k=1}^{N}\frac{\frac{\partial}{\partial(\mathbf{r}_S,\boldsymbol{\theta})}f_D(D_k|\mathbf{u}_k,\mathbf{r}_S,\mathbf{r}_{Dk},\boldsymbol{\theta})}{f_D(D_k|\mathbf{u}_k,\mathbf{r}_S,\mathbf{r}_{Dk},\boldsymbol{\theta})} = \mathbf{0}. \tag{34}$$

To evaluate the partial derivatives with respect to the source location, substitution of (26) and (21) into (34) and differentiating with respect to $\mathbf{r}_S$ has the following result:

$$\sum_{k=1}^{N}\frac{D_k(1-\mathrm{Pr}_{Pk})\sum_{i=1}^{t_{\max}/\Delta t}\left[\frac{(r_{Sj}-r_{Dkj}-\mu_{1j}(t_{fi},u_{kj}))}{\sigma_{1j}(t_{fi})^2}\mathrm{Pr}_{Pik}\right]}{\frac{p_{\mathrm{FA}}}{(p_D-p_{\mathrm{FA}})}+\mathrm{Pr}_{Pk}} - \sum_{k=1}^{N}\frac{(1-D_k)(1-\mathrm{Pr}_{Pk})\sum_{i=1}^{t_{\max}/\Delta t}\left[\frac{(r_{Sj}-r_{Dkj}-\mu_{1j}(t_{fi},u_{kj}))}{\sigma_{1j}(t_{fi})^2}\mathrm{Pr}_{Pik}\right]}{\frac{(1-p_{\mathrm{FA}})}{(p_D-p_{\mathrm{FA}})}-\mathrm{Pr}_{Pk}} = 0 \tag{35}$$

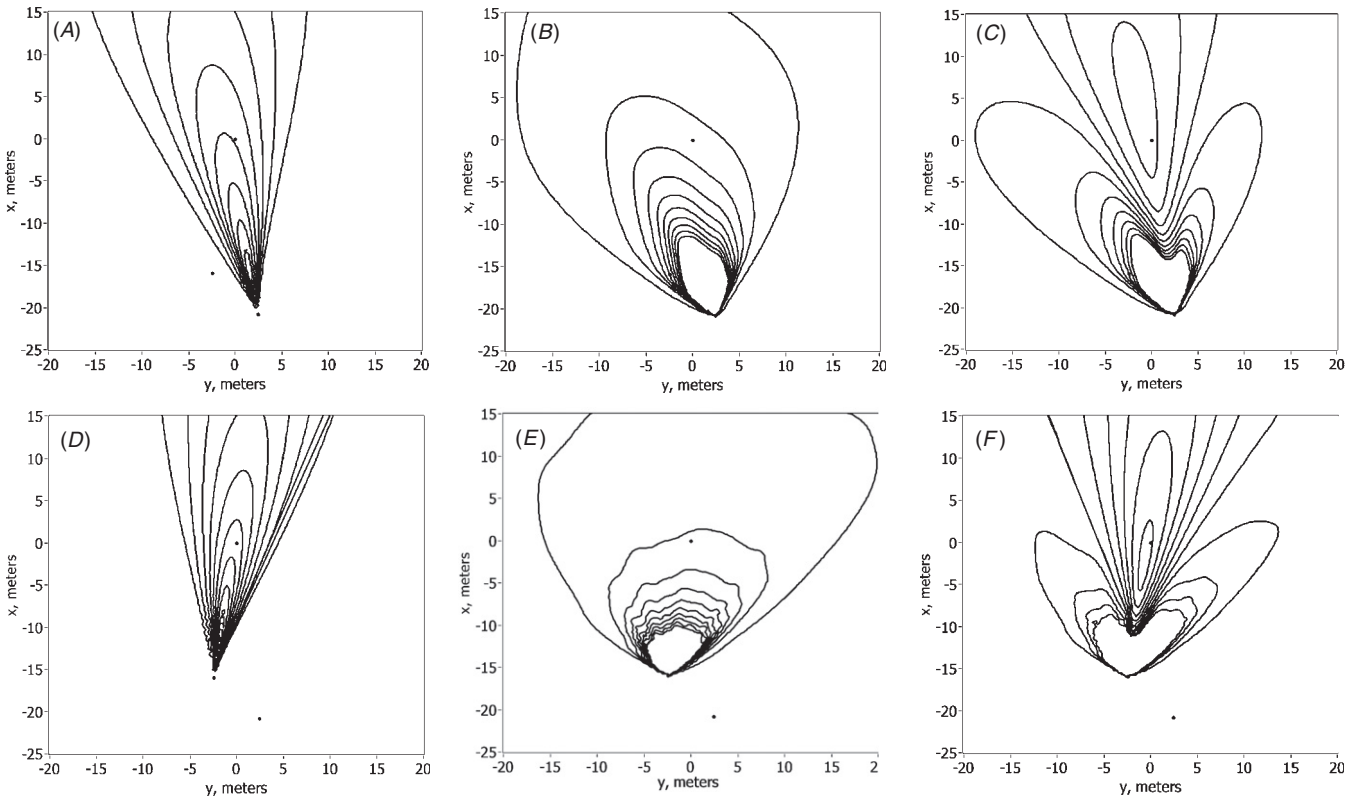

Figure 5. Conditional ($\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$) iso-probability contours of different components of log-likelihood functions resulting from computations made from just two locations. (*A*) Detections from location 10. (*B*) Non-detections from location 10. (*C*) Detections and non-detections from location 10. (*D*) Detections from location 9. (*E*) Non-detections from location 9. (*F*) Detections and non-detections from location 9.

where

$$\begin{aligned}\mathrm{Pr}_{Pik} &= \mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}, \mathrm{i}\Delta t) \\ &= A_P \Delta t f_{\mathbf{R}_I}(\mathbf{r}_S|\mathbf{u}_k, \mathbf{r}_{Dk}, \mathrm{i}\Delta t)\end{aligned}$$

and

$$\begin{aligned}\mathrm{Pr}_{Pk} &= \mathrm{Pr}_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}) \\ &= 1 - \exp\left(-A_P \Delta t \sum_{i=1}^{t_{\max}/\Delta t} f_{\mathbf{R}_I}(\mathbf{r}_S|\mathbf{u}_k, \mathbf{r}_{Dk}, i\Delta t)\right).\end{aligned}$$

In this implementation, simple Riemann sums are used to calculate (35) in increments of $\Delta t = 0.25$ s. Indices are repeated here for convenience and clarity. i enumerates the flight times, j enumerates the three Cartesian directions, and k enumerates the acquisition time intervals. Similarly, taking the derivative with respect to A_P of the log-likelihood function, (34), has the following result:

$$\sum_{k=1}^{N} \frac{D_k \sum_{i=1}^{t_{\max}/\Delta t} [\mathrm{Pr}_{Pik}](1 - \mathrm{Pr}_{Pk})}{\frac{p_{\mathrm{FA}}}{(p_D - p_{\mathrm{FA}})} + \mathrm{Pr}_{Pk}} - \sum_{k=1}^{N} \frac{(1 - D_k) \sum_{i=1}^{t_{\max}/\Delta t} [\mathrm{Pr}_{Pik}](1 - \mathrm{Pr}_{Pk})}{\frac{(1-p_{\mathrm{FA}})}{(p_D - p_{\mathrm{FA}})} - \mathrm{Pr}_{Pk}} = 0. \quad (36)$$

Last, taking the derivative with respect to p_{FA} of the log-likelihood function, (34), gives the last simultaneous equation:

$$\sum_{k=1}^{N} \frac{D_k (1 - \mathrm{Pr}_{Pk})}{\frac{p_{\mathrm{FA}}}{(p_D - p_{\mathrm{FA}})} + \mathrm{Pr}_{Pk}} - \sum_{k=1}^{N} \frac{(1 - D_k)(1 - \mathrm{Pr}_{Pk})}{\frac{(1-p_{\mathrm{FA}})}{(p_D - p_{\mathrm{FA}})} - \mathrm{Pr}_{Pk}} = 0. \quad (37)$$

It can be seen that the first term in equations (35), (36) and (37) is attenuated whenever the first term in the denominator exceeds the second term in the denominator. This roughly defines a sampling region, outside of which, positive detections 'do not count' as much. The region shrinks as the false alarm rate increases.

5. Example illustrating how the likelihood function locates the source

The likelihood function, (32), can be separated as a product of density functions by location. It can also be further separated into components that are due to detections and non-detections (see equation (26)) This separation helps to visualize how the method works, which we believe is largely through triangulation. Figures 5 and 6 are included to illustrate components that contribute to the total likelihood function when the inference is conducted from just two points, locations 9 and 10 (see figure 8). This analysis was computed using the most likely value of the parcel area and false alarm rate obtained from data at both locations. In each panel, contours of a log-likelihood function are plotted. Ten isocontours were drawn with the first at a level 2.3 below the maximum and then contour levels were changed in increments of −7 after that. In panels A and D, detections seem to mainly provide information about the direction to the source. In panels B and E, the largest contour has the highest value, descending as the contours shrink. In panel B, it can be seen that the source is

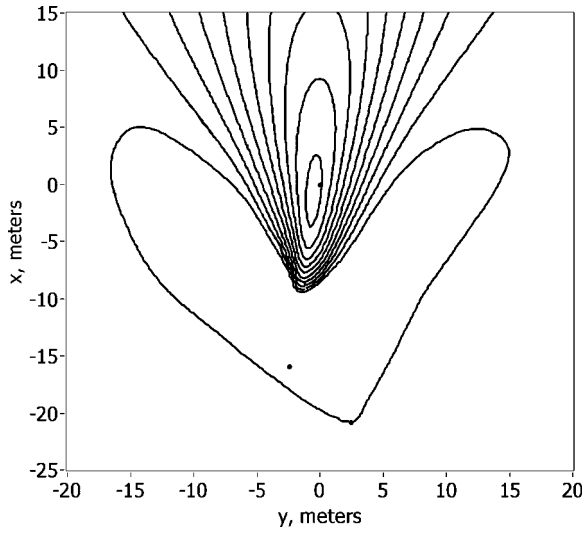

Figure 6. Conditional ($\theta = \hat{\theta}$) iso-probability contours of the log-likelihood function resulting from computations made from just two locations.

very unlikely to be near the upwind side of location 10 and also less likely to be to the left of the source. When both detections and non-detections obtained at location 10 are combined in panel C, the non-detections served to push the estimate of the distance to the source away from the detection location and the direction to the source perhaps slightly to the right. When the parcel area is increased, non-detections will generally push the estimate farther upwind from the detection location and vice-versa for a reduction in the parcel area. Similarly, the likelihood functions from location 9 shown in D and E combine to produce an estimate of the source posterior density function in F. When all six components of the likelihood function are combined, the log-likelihood function in figure 6 results. The most likely value for the parcel area is that which enables the parcels associated with detections to arrive at both detection locations from the same location.

6. Assumptions

We have compiled a list of assumptions so that the reader may refer to it. These follow.

- Wind direction is sufficiently variable that the chances of plume interception are high and triangulation may be used to locate the source.
- Prior information about source parameters is vague enough to be approximated as ignorant.
- Wind velocity and detection locations are unrelated to the source parameters.
- Source releases one parcel per time interval.
- Parcel movement is governed by the Langevin equation.
- Parcel movements in orthogonal directions are independent.
- Variance energy above sampling frequency is negligible.
- Parcels each occupy a constant volume.
- Parcels do not re-cross the detector.
- Probability of detection is 1.
- Probability of false alarm is constant.

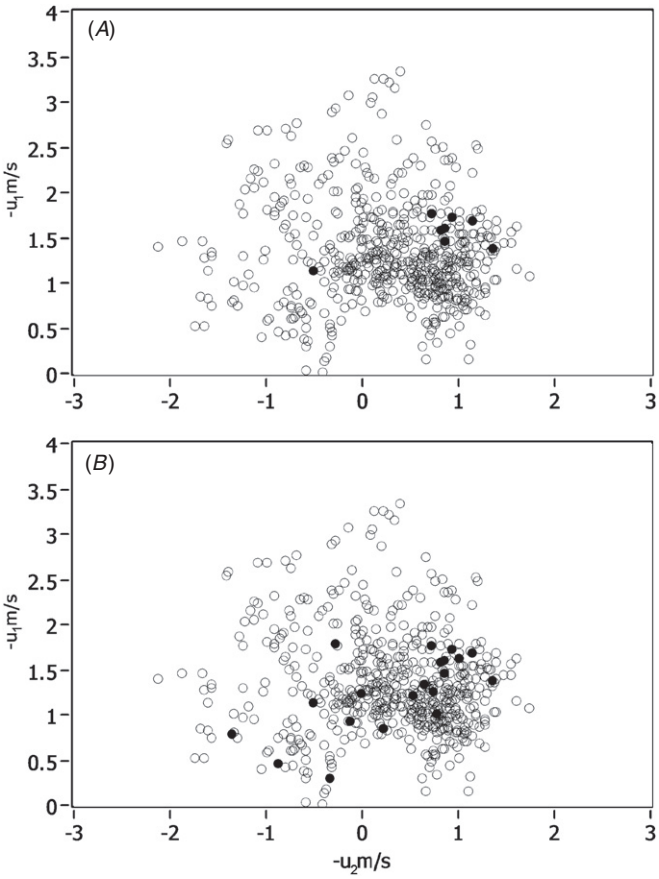

Figure 7. Anemometer data (hits and non-hits) recorded at location 5. Anemometer velocities associated with hits are marked with larger black dots, while velocities associated with non-hits are open circles.

- Parcels are driven by independent Brownian motions, and thus correlation due to spatial proximity is neglected.
- Ignorance of parcel history based on previous measurements is assumed.

7. Results

Data were collected and processed as described in section 2. The EAG data were run through the stage 1 detector using two background prior probability values, providing two analyses—one with a relatively high false alarm rate and one with a low false alarm rate. One might observe that the preliminary data processing reduces the EAG to a set of only two symbols received only once per second. For example, the computed average information rate (using (26) and (27)) during the time interval displayed in figure 2 (low p_{FA}) is 0.30 bits s^{-1}. The average information rate for the entire recording is 0.085 bits s^{-1}. This is in contrast to the estimated channel capacity of the EAG which is in the tens of bits per second. Thus quite a bit of information is being discarded that could potentially be of use in the determination of source parameters.

Figure 7 is included to illustrate both the nature of the anemometer data and the data from the resulting analyses. This figure shows the wind velocity recorded while standing at location 5 (see figure 8). The open circles record the velocity

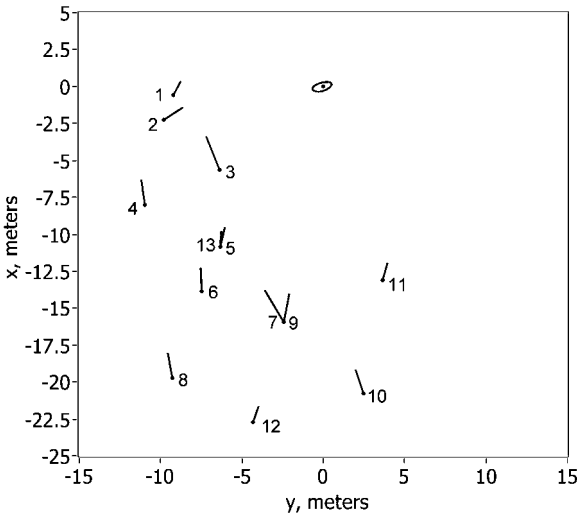

Figure 8. Top view showing topological distribution of the source (0,0) and the recording locations. An iso-probability density is drawn representing the conditional ($\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$) result of the low P_{FA} 'all locations included' analysis for the scale. Contour encompasses 68% (based on the assumed Gaussian density). Mean wind is directed along the x axis. The mean anemometer vector in meters per second while standing at each location (looking into the wind) is drawn at each location.

Table 1. Detection event counts.

Position	Time (s)	Low p_{FA} hits	High p_{FA} hits
1	767	15	35
2	568	40	51
3	518	0	10
4	812	0	10
5	608	8	20
6	514	0	5
7	639	1	5
8	659	6	14
9	513	17	31
10	702	13	20
11	583	4	4
12	528	2	5
13	512	3	4

measured during non-detection times (non-hits), while the dark circles mark the wind velocity measured during positive detection intervals (hits). In panel A, it can be seen that the wind was blowing from the source during almost all of the hit times. Panel B has more hits, but more of them were recorded when wind was not coming from the source and are likely false positives.

The numbers of positive detections resulting from these analyses while standing at each location are summarized in table 1. So that the reader may make sense of the hit counts in table 1 the average anemometer velocity vector is drawn at each location in figure 8. For instance one might expect similar results to be obtained at locations 7 and 9, but it can be seen that the average direction of the wind was significantly different during the collection times.

Note that the most numerous hits occurred while standing at locations 1, 2 and 9. The angles to the source from locations 1 and 2 are almost 90° from the angle from location 9, possibly due to synoptic scale changes. These locations allow the source location to be accurately estimated through triangulation. Therefore to test the robustness of the method, analyses were conducted without the data recorded from these locations resulting in four analysis types (see table 3).

Model parameters were assigned using the methods documented in table 2. Following this, the parameters ($\mathbf{r}_S$, $\boldsymbol{\theta}$) that maximized the likelihood function were computed under the four conditions. The resulting parameter estimates are tabulated in table 3.

Note that the p_{FA} estimate increases as expected when the false alarm rate in the EAG data is increased. The false alarm rates obtained are also consistent with the number that may be surmised from the anemometer data in figure 7. At the high rate, the number of expected false alarms at location 5 (10.1 min) would be about 8 to 10 and at the low false alarm rate it would be 0.8 to 1.5.

The parcel frontal area values have a limited geometric interpretation. In a limited sense, they can be interpreted as the frontal area released per second by the source. Because information about the length of the odorant detections is discarded, nothing can be said about the plume volume. If the average length of odorant peaks was known, we would expect the volumetric flow rate could be roughly estimated, and thus an estimate of the cross-sectional area of the plume. Or, if a concentration threshold was applied (rather than peak detection) and the anemometer interval reduced so that odorant peaks were sufficiently resolved, approaching a constant detection/non-detection ratio, a volumetric interpretation could apply that would represent an average value of the meandering plume volumetric flow rate. When locations 1, 2 and 9 are excluded, A_P decreases because the locations with the most hits have been intentionally removed, biasing this parameter to a smaller value. The parcel area increases when the high false alarm rate EAG data is used under low p_{FA} conditions where the source location estimates are similar. This is likely the result of an increased probability of detection, p_D, which has a constant assumed value of 1. However it is evident that the probability of detection increases in the anemometer data in figure 7. When the likelihood function is approximated (not shown) assuming $p_{\mathrm{FA}} \ll p_D$ and $\Pr_P(P = 1|\mathbf{u}_k, \mathbf{r}_S, \mathbf{r}_{Dk}, \boldsymbol{\theta}) \ll 1$, (equation (22)), the $p_D A_P$ product is constant. This effect is not observed in the high p_{FA} results because the source location estimates differ significantly.

Each maximization iteration takes approximately 3 s on 132 min of data on a 1.6 GHz Intel T5200. When maxima are attainable from the start point, five iterations suffice. It is expected that a simple change such as implementing an efficient integration algorithm would speed up the operation by perhaps a factor of 5–10, so that maxima could be obtained in about 2 or 3 s.

For illustrative purposes, conditional posterior probability densities for the source location were calculated and iso-probability density contours were drawn. The conditions imposed were that the parameter vector $\boldsymbol{\theta}(A_P$ and $p_{\mathrm{FA}})$ took on its most likely value, tabulated in table 1. Figure 7 includes

Table 2. Model parameters.

Symbol	Description	Source	Value
r_3	Height above ground	Approximate	1.5 m
u^*	Friction velocity	Approximate	0.2 m s^{-1}
z_i	Boundary layer height	Approximate	1000 m
$t_{\max}$	Integration limit	Assigned	25 s
p_D	Probability of detection	Assigned	1
T	Sampling interval	Assigned	1.00 s
Δt	Numerical integration interval	Assigned	0.25 s
$\bar{u}_1$	Mean wind speed	Measured	−1.50 m s^{-1}
σ_{u1}	Standard deviation	See equation (10)	0.70 m s^{-1}
σ_{u2}	Standard deviation	See equation (10)	0.92 m s^{-1}
σ_{u3}	Standard deviation	See equation (17)	0.27 m s^{-1}
T_{L1}	Lagrangian time constant	See equations (13), (14)	115 s
T_{L2}	Lagrangian time constant	See equations (13), (14)	95 s
T_{L3}	Lagrangian time constant	See equation (15)	3.1 s

Table 3. Maximum likelihood results.

Analysis type	Positions excluded	A_P (m^2)	$p_{\mathrm{FA}} \times 10^{-3}$	False alarm rate (min^{-1})	Source x (m)	Source y (m)
Low p_{FA}	None	2.38	2.4	0.14	−0.03	−0.20
Low p_{FA}	1, 2, 9	1.28	1.1	0.08	0.25	−0.04
High p_{FA}	None	2.73	14	0.85	0.06	−0.32
High p_{FA}	1, 2, 9	1.16	11	0.98	−1.46	−0.83

the result from the low p_{FA} analysis including all of the data. The iso-probability contour drawn includes 68% of the density (or approximately 1 standard deviation in a one-dimensional Gaussian).

Figure 9 was included to visualize the conditional posterior densities of all four analyses. In panel A, each iso-probability contour includes 68% of the density, where the level is based on the assumption that the distribution is a two-dimensional Gaussian function. Panel B includes 95% of the probability density under the same assumption.

Computation of a likelihood function value takes about 0.45 s. Computation of conditional posterior densities (60 × 60) used to construct figures 8 and 9 each took about 27 min. If marginal densities were to be computed ($\boldsymbol{\theta}$ unknown) the amount of time to obtain the posterior density function could take days or weeks. Although robotic maneuvers are excluded from this paper, it can be pointed out that tiny male moths lacking such computing power are able to locate females through optomotor–anemotactic maneuvers in which moths use the ground as an optical reference for upwind-directed movement when intermittent pheromone is detected [63] in just minutes.

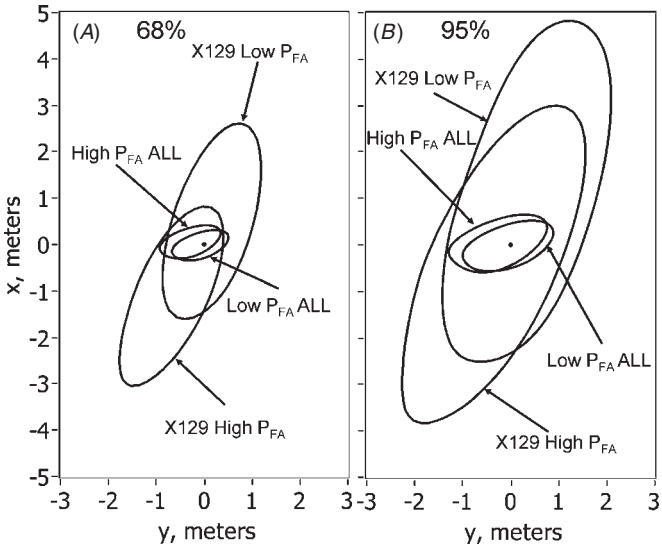

Figure 9. Conditional ($\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$) posterior density iso-probability contours. Mean wind is directed along the x axis. The source was located at (0,0). The labels 'ALL' refer to analyses that include all locations. The labels 'X129' refer to analyses that excluded locations 1, 2 and 9. (*A*) Contours encompass 68% (based on the assumed Gaussian density) of the probability density. (*B*) Contours encompass 95% (based on the assumed Gaussian density) of the probability density

8. Discussion

We have introduced a Lagrangian dispersion model for the solution to the odor source localization problem when local velocity measurements near a fast detector are known. The model was used in conjunction with Bayesian inference to estimate the source's location. Results from an experiment were also presented. The experimental setup involved a pheromonal odor source at 1.5 m over flat mowed grass, relatively distant from obstructions. Pheromone was detected at downwind distances of up to 23 m from the source using live-moth EAG preparations. Results were successful, but further improvements are likely to be possible. Improved measurement capability could include the addition of inertial measurements for automated probe location measurements, addition of a third direction to anemometric measurements and/or anemometric arrays. Better modeling of scalar dispersion, including relative dispersion, concentration

information, and inhomogeneous flow in the z direction may enable multiple direct downwind measurements to be used to estimate distance to the source. This would enable a maneuvering planner for a robotic approach to optimize its position downwind more often, balancing unlikely but valuable off-angle measurements with direct downwind measurements as well as enabling source location estimates to be made under weather conditions where triangulation is difficult.

Acknowledgments

The authors gratefully acknowledge our grant support from the Office of Naval Research Counter-IED program and from the Defense Threat Reduction Agency. We would like to thank Bryan Banks for supplying the test insects.

References

[1] Keats A, Yee E and Lien F-S 2007 Bayesian inference for source determination with applications to a complex urban environment *Atmos. Environ.* **41** 465–79

[2] Yee E, Lien F-S, Keats A and D'Amours R 2008 Bayesian inversion of concentration data: source reconstruction in the adjoint representation of atmospheric diffusion *J. Wind Eng. Indust. Aerodyn.* **96** 1805–16

[3] Harvey D J, Lu T-F and Keller M A 2008 Comparing insect-inspired chemical plume tracking algorithms using a mobile robot *IEEE Trans. Robot.* **24** 307–17

[4] Ishida H, Nakayama G, Nakamoto T and Moriizumi T 2005 Controlling a gas/Odor plume-tracking robot based on transient responses of gas sensors *IEEE Sens. J.* **5** 537–45

[5] Lytridis C, Kadar E E and Virk G S 2006 A systematic approach to the problem of odour source localisation *Auton. Robots* **20** 261–76

[6] Kowadlo G and Russell R A 2008 Robot odor localization: a taxonomy and survey *Int. J. Robot. Res.* **27** 869–94

[7] Kuwana Y, Shimoyama I and Miura H 1995 Steering control of a mobile robot using insect antennae *Proc. 1995 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems. (Pittsburgh, PA, USA, 1995)* vol 2 pp 530–5

[8] Kuwana Y, Nagasawa S, Shimoyama I and Kanzaki R 1999 Synthesis of the pheromone-oriented behaviour of silkworm moths by a mobile robot with moth antennae as pheromone sensors *Biosens. Bioelectron.* **14** 195–202

[9] Pyk P, Badia S B i, Bernardet U, Knusel P, Carlsson M, Gu J, Chanie E, Hansson B S, Pearce T C and Verschure P F M J 2006 An artificial moth: chemical source localization using a robot based neuronal model of moth optomoter anemotactic search *Auton. Robots* **20** 197–213

[10] Bailey J K, Willis M A and Quinn R D 2005 A multi-sensory robot for testing biologically-inspired odor plume tracking strategies *Proc. 2005 IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics (Monterey, CA, USA, 2005)*

[11] Balkovsky E and Shraiman B I 2002 Olfactory search at high Reynolds number *Proc. Natl Acad. Sci.* **99** 12589–93

[12] Pang S and Farrell J A 2006 Chemical plume source localization *IEEE Trans. Syst. Man Cypern.* B **36** 1068–80

[13] Vergassola M, Villermaux E and Shraiman B I 2007 'Infotaxis' as a strategy for searching without gradients *Nature* **445** 406–9

[14] Farrell J A 2003 Plume mapping via hidden Markov methods *IEEE Trans. Syst. Man Cybern.* B **33** 850–63

[15] Yee E, Kosteniuk P R, Chandler G M, Biltoft C A and Bowers J F 1993 Statistical characteristics of concentration fluctuations in dispersing plumes in the atmospheric surface layer *Bound.-Layer Meteorol.* **65** 69–109

[16] Guo S, Yang R, Zhang H, Weng W and Fan W 2009 Source identification for unsteady atmospheric dispersion of hazardous materials using Markov Chain Monte Carlo method *Int. J. Heat Mass Transfer* **52** 3955–62

[17] Abida R and Bocquet M 2009 Targeting of observations for accidental atmospheric release monitoring *Atmos. Environ.* **43** 6312–27

[18] Zhang Y, Schervish M J, Acar E U and Choset H M 2001 Probabilistic methods for robotic landmine search *Proc. SPIE* **18** 243–54

[19] Jakuba M V and Yoerger D R 2008 Autonomous search for hydrothermal vent fields with occupancy grid maps *Proc. of ACRA '08, 2008*

[20] Baldocci D 1992 A Lagrangian random-walk model for simulating water vapor, CO_2 and sensible heat flux densities and scalar profiles over and within a soybean canopy *Bound.-Layer Meteorol.* **61** 113–44

[21] Dosio A, Arellano J V-G D and Holtslag A A M 2005 Relating Eulerian and Lagrangian statistics for the turbulent dispersion in the atmospheric convective boundary layer *J. Atmos. Sci.* **62** 1175–91

[22] Flesch T K, Wilson J D and Yee E 1995 Backward time Lagrangian dispersion models and their applications to estimate gaseous emissions *J. Appl. Meteorol.* **34** 1320–32

[23] Schneider D 1957 Electrophysiologische Untersuchungen von Chemo-und Mechanorezeptoren der Antennae des Seidenspinner Bombix mori L *Z. Vgl. Physiol.* **40** 8–41

[24] Justice K A, Carde R T and French A S 2005 Dynamic properties of antennal responses to pheromone in two moth species *J. Neurophiol.* **93** 2233–9

[25] Schuckel J, Meisner S, Torkkeli P H and French A S 2008 Dynamic properties of *Drosophila* olfactory electroantennograms *J. Comp. Physiol.* A **194** 483–9

[26] Kaissling K-E 1995 Single unit and electroantennogram recordings in insect olfactory organs *Experimental Cell Biology of Taste and Olfaction* ed A I Spielman and J G Brand (Boca Raton, FL: CRC Press) p 361

[27] Patte F, Etcheto M, Marfaing P and Laffort P 1989 Electroantennogram stimulus-response curves for 59 odourants in the honey bee, *Apis mellifica J. Insect Physiol.* **35** 667–75

[28] French A S and Meisner S 2007 A new method for wide frequency range dynamic olfactory stimulation and characterization *Chem. Senses* **32** 681–8

[29] Matthes J, Groll L and Keller H B 2005 Source localization by spatially distributed electronic noses for advection and diffusion *IEEE Trans. Signal Proc.* **53** 1711–9

[30] Cai J and Levy D C 2007 Using stationary electronic noses network to locate dynamic odour source position *Proc. IEEE Int. Conf. on Integration Technology (Shenzhen, China)* pp 793–8

[31] Persaud K and Dodd G 1982 Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose *Nature* **299** 352–5

[32] Pearce T C, Schiffman S S, Nagle H T and Gardner J W 2002 *Handbook of Machine Olfaction* (Weinheim: Wiley-VCH)

[33] Arshak K, Lyons G M, Cunniffe C, Harris J and Clifford S 2003 A review of digital data acquisition hardware and software for a portable electronic nose *Sensor Rev.* **23** 332–44

[34] Arshak K, Moore E, Lyons G M, Harris J and Clifford S 2004 A review of gas sensors employed in electronic nose applications *Sensor Rev.* **24** 181–98

[35] Rock F, Barsan N and Weimar U 2008 Electronic nose: current status and future trends *Chem. Rev.* **108** 705–25

[36] Wibe A 2004 How the choice of method influence on the results in electrophysiological studies of insect olfaction *J. Insect Physiol.* **50** 497–503
[37] Park K C, Ochieng S A, Zhu J and Baker T C 2002 Odor discrimination using insect electroantennogram responses from an insect antennal array *Chem. Senses* **27** 343–52
[38] Park K C and Hardie J 1998 An improved aphid electroantennogram *J. Insect Physiol.* **44** 919–28
[39] Ziegler C, Gopel W, Hammerle H, Hatt H, Jung G, Laxhuber L, Schmidt H L, Schutz S, Vogtle F and Zell A 1998 Bioelectronic noses: a status report: part II *Biosens. Bioelectron.* **13** 539–71
[40] Myrick A J, Park K-C, Hetling J R and Baker T C 2008 Real-time odor discrimination using a bioelectronic sensor array based on the insect electroantennogram *Bioinspiration Biomimetics* **3** 046006
[41] Myrick A J, Park K C, Hetling J R and Baker T C 2009 Detection and discrimination of mixed odor strands in overlapping plumes using an insect-antenna-based chemosensor system *J. Chem. Ecol.* **35** 118–30
[42] Turner D B 1964 A diffusion model for an urban area *J. Appl. Meteorol.* **3** 83–91
[43] Pope M M, Gaston L K and Baker T C 1984 Composition, quantification, and periodicity of sex pheromone volatiles from individual *Heliothis zea* females *J. Insect Physiol.* **30** 943–5
[44] Klun J A, Plimmer J R, Bierl-Leonhardt B A, Sparks A N and Chapman O L 1979 Trace chemicals: the essence of sexual communication systems in *Heliothis* species *Science* **204** 1328–30
[45] Pope S B 1994 Lagrangian PDF methods for turbulent flows *Annu. Rev. Fluid Mech.* **26** 23–63
[46] Hanna S R 1981 Lagrangian and Eulerian time-scale relations in the daytime boundary layer *J. Appl. Meteorol.* **20** 242–9
[47] Uhlenbeck G E and Ornstein L S 1930 On the theory of the Brownian motion *Phys. Rev.* **36** 823–41
[48] Gardiner C W 1985 *Handbook of Stochastic Methods* 2nd edn (New York: Springer)
[49] Thomson D J 1984 Random walk modelling of diffusion in inhomogenous turbulence *Q. J. R. Meteorol. Soc.* **110** 1107–20
[50] Thomson D J 1987 Criteria for the selection of stochastic models of particle trajectories in turbulent flows *J. Fluid Mech.* **180** 529–56
[51] Stull R B 1988 *An Introduction to Boundary Layer Meteorology* (Dordrecht: Kluwer)
[52] van der Hoven I 1957 Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour *J. Meteorol.* **14** 160–4
[53] Taylor G I 1921 Diffusion by continuous movements *Proc. London Math. Soc.* **20** 196–212
[54] Brier G W 1950 The statistical theory of turbulence and the problem of diffusion in the atmosphere *J. Meteorol.* **7** 283–90
[55] Batchelor G K 1950 The application of the similarity theory of turbulence to atmospheric diffusion *Q. J. R. Meteorol. Soc.* **76** 133
[56] Sawford B 2001 Turbulent relative dispersion *Annu. Rev. Fluid Mech.* **33** 289–317
[57] Degrazia G A, Anfossi D, Carvalho J C, Mangia C, Tirabassi T and Velho H F C 2000 Turbulence parameterisation for PBL dispersion models in all stability conditions *Atmos. Environ.* **34** 3575–83
[58] Aylor D E and Flesch T K 2001 Estimating spore release rates using a Lagrangian stochastic simulation model *J. Appl. Meteorol.* **40** 1196–208
[59] Wilson J D 1982 Estimation of the rate of gaseous mass transfer from a surface source plot to the atmosphere *Atmos. Environ.* **16** 1861–7
[60] Shannon C E 1948 A mathematical theory of communication *Bell Syst. Tech. J.* **27** 379–423, 623–56
[61] Barnett V 1982 *Comparative Statistical Inference* 2nd edn (New York: Wiley)
[62] Fletcher R 2000 *Practical Methods of Optimization* 2nd edn (Chichester, UK: Wiley)
[63] Vickers N J and Baker T C 1994 Visual feedback in the control of pheromone mediated flight of *Heliothis virescens* males (*Lepidoptera: Noctuidae*) *J. Insect Behav.* **7** 605–32