Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

On-line Dispersion Source Estimation using Adaptive Gaussian Mixture Filter

Marco F. Huber *

* *AGT International*
Darmstadt, Germany
e-mail: marco.huber@ieee.org

Abstract: The reconstruction of environmental events has gained increased interest in the recent years. In this paper, the focus is on estimating the location and strength of a gas release from distributed measurements. The estimation is formulated as Bayesian inverse problem, which utilizes a Gaussian plume forward model. A novel recursive estimation algorithm based on statistical linearization and Gaussian mixture densities with adaptive component number selection is used in order to allow accurate and computationally efficient source estimation at the same time. The proposed solution is compared against state-of-the-art methods via a simulations and a real-word experiment.

1. INTRODUCTION

If a hazardous gas has been released—either accidentally or deliberately—into atmosphere, it is of paramount importance to gain knowledge of this event at an early stage in order to increase the effectiveness of counter measures for protecting the public and for mitigating the harmful effects. By means of so-called *atmospheric dispersion models* (ADMs), it is possible to predict the concentration spread of the released gas in space and time. These models, however, merely provide reliable predictions if the characteristics of the gas source are known precisely. To determine or estimate the source characteristics it necessary to solve an *inverse problem*, where one has to infer the location and strength of the gas release from concentration measurements of spatially distributed sensors.

In general, solution methods of the source estimation problem can be classified into *forward* and *backward* methods (Rao [2007]). Forward methods employ an forward-running ADM multiple times in order to find an estimate of the source that best describes the given concentration measurements. Here, the mostly used techniques are based on Bayesian inference in combination with Monte Carlo sampling. Sequential Monte Carlo methods as described in Sohn et al. [2002] or Zhang and Wang [2013] employ a set of samples or particles that forms the posterior probability distribution of the source parameters. This distribution is updated by means of Bayes' rule whenever new concentration measurements from sensors are available. In contrast to this online procedure, Markov chain Monte Carlo (MCMC) methods process all acquired concentration measurements in a batch in order to determine the posterior distribution. For this purpose, samples are drawn from the posterior distribution by simulating a Markov chain that has the desired posterior distribution as its stationary distribution. Given a properly constructed Markov chain it can be shown that MCMC reaches the stationary distribution after a typically large number of sampling steps. Application of MCMC to source estimation can be found for instance in Senocak et al. [2008], Borysiewicz et al. [2012], and Hirst et al. [2013].

Backward methods instead perform only one model run in the reverse direction from the sensors to the source. Commonly used techniques are backtracking, where an inverse version of an ADM is utilized (see e.g. Hourdin and Talagrand [2006]), and variational methods, where a cost function between model predictions and concentration measurements is optimized (see e.g. Stockie [2011], Rudd et al. [2012]). The backward approach is preferred over forward methods, when the number of sources is larger than the number of sensors (Rao [2007]).

In this paper, a novel forward approach is proposed that aims at performing on-line source estimation, i.e., the current source estimate is updated at run-time whenever sensors provide new concentration measurements. For this purpose, a Gaussian plume dispersion model is employed that allows predicting the gas dispersion in closed form with low computational overhead. This forward model is employed in a recursive Bayesian inference framework to allow for uncertainties arising from modeling errors and sensor noise. The resulting statistical inverse problem, however, cannot be solved in closed form due to nonlinearities in the Gaussian plume model. To overcome this issue, the so-called *adaptive Gaussian mixture filter* (AGMF) proposed in Huber [2011] is employed. Here, the posterior distribution is approximated via a sum of Gaussians, which is known to be a universal function approximator (Maz'ya and Schmidt [1996]). To limit the computational demand but still perform accurate estimation, the number of Gaussian components is adapted at run-time depending on the nonlinearity of the dispersion model.

The paper is structured as follows: In the next section, a general ADM and its special case, the Gaussian plume model, are introduced. The recursive Bayesian estimation problem is stated in Section 3, while Section 4 describes the AGMF. A comparison of the proposed source estimation method with state-of-the-art is provided in Section 5, followed by a conclusion in Section 6.

2. PROBLEM FORMULATION

In this section, a physical model describing the spatial dispersion of a substance in atmosphere is derived. It models the transportation of a substance from an emitting source to regions of low concentration under consideration of various environmental conditions.

2.1 General Dispersion Model

In the following, $c(\underline{x},t)$ is the concentration of the substance at position $\underline{x} = [x,y,z]^{\mathrm{T}} \in \mathbb{R}^3$ and at time $t \geq 0$. The concentration follows the *advection-diffusion equation*

$$\frac{\partial c(\underline{x},t)}{\partial t} = \nabla \cdot \big(\mathbf{K}\nabla c(\underline{x},t) - \underline{v}\cdot c(\underline{x},t)\big) + s(\underline{x},t) \qquad (1)$$

with $\nabla \triangleq [\partial/\partial x, \partial/\partial y, \partial/\partial z]^{\mathrm{T}}$ (see e.g. Holzbecher [2012]). The term $\mathbf{K}\nabla c(\underline{x},t)$ describes the diffusion according to Fick's law with diffusion matrix $\mathbf{K}(\underline{x},t)$ and the term $\underline{v}\cdot c(\underline{x},t)$ represents linear advection due to wind with velocity $\underline{v}(\underline{x},t)$. Finally, $s(\underline{x},t)$ is a source or sink term.

Analytical solutions of (1), i.e., functions $c(\underline{x},t)$ satisfying the equation, exist merely for some special cases. In the following, one important special case utilized throughout the paper is introduced.

2.2 Gaussian Plume

In order to obtain a closed-form solution, it is necessary to make several assumptions:

(i) The substance is emitted at a constant *rate* $q > 0$ from a single point source at location $\underline{x}_s \triangleq [x_s, y_s, z_s]^{\mathrm{T}}$. Thus, the source term $s(\underline{x},t)$ in (1) becomes

$$s(\underline{x},t) = q \cdot \delta(x - x_s) \cdot \delta(y - y_s) \cdot \delta(z - z_s) \ ,$$

where $\delta(x - x_s)$ is the Dirac delta localized at x_s.

(ii) The wind is constant with velocity $v \geq 0$ and the wind direction draws an angle ϕ with the x-axis so that $\underline{v} = v \cdot [\cos\phi, \sin\phi, 0]^{\mathrm{T}}$.

(iii) The diffusion is a function of the downwind distance only. Furthermore, it is assumed that the diffusion in wind direction can be neglected.

(iv) The terrain is flat and the ground cannot be penetrated by the substance.

(v) The solution is steady state, i.e., time independent.

Based on these assumptions and additional boundary conditions that force vanishing concentrations at infinite distance from the source and at upwind distances, (1) has the time-invariant solution

$$c(\underline{x}) = \frac{q}{2\pi \cdot v \cdot \sigma_y \sigma_z} \cdot \exp\!\left(-\frac{(1+2\sin(\phi)\cos(\phi)) \cdot (y-y_s)^2}{2\sigma_y^2}\right) \cdot \left[\exp\!\left(-\frac{(z-z_s)^2}{2\sigma_z^2}\right) + \exp\!\left(-\frac{(z+z_s)^2}{2\sigma_z^2}\right)\right] \ , \qquad (2)$$

which is the well-known *Gaussian plume* dispersion model (for a detailed derivation see Stockie [2011]). Here, σ_y and σ_z are the so-called standard deviations of the Gaussian concentration distribution. They are both functions of x, y, ϕ and they depend on the stability of the atmosphere.

The Gaussian plume model (2) is employed as it is widely used and suitable for describing short range substance releases. Furthermore, being an analytical model, it allows for an on-line and computationally light-weight estimation of the unknown parameters.

In this paper, the focus is on estimating the source rate q and location $\underline{x}_s$ from a set of spatially distributed concentration measurements. It is assumed that the measurements become available sequentially over time, i.e., batch or off-line estimation is impractical. Additional parameters like wind speed or direction are assumed to be known, as they can be provided reliably from external sources like weather stations. However, the approach proposed in this paper can be easily extended to estimate also these additional parameters.

3. RECURSIVE ESTIMATION

The Gaussian plume model forms an instance of a so-called *forward model*

$$\underline{z} = \underline{g}(\underline{\theta}) \ , \qquad (3)$$

where the output or observations $\underline{z}$ are defined based on physical transformations $\underline{g}(.)$ and model parameters $\underline{\theta}$. In the considered problem, $\underline{z}$ corresponds to a set of concentration measurements, $\underline{g}$ is the Gaussian plume model (2), and $\underline{\theta}^{\mathrm{T}} \triangleq \left[q, \underline{x}_s^{\mathrm{T}}\right]$ comprises the source rate as well as the source location.

As the goal is to determine the parameters $\underline{\theta}$, an *inverse problem* of (3) needs to be considered, where $\underline{\theta}$ is estimated given the observed concentrations $\underline{z}$. Inverse problems are typically difficult to solve: they are often ill-conditioned, ambiguities exist—the observations $\underline{z}$ can be explained by different parameters $\underline{\theta}$—, and an inverse transformation $\underline{g}^{-1}$ often is not available.

3.1 Bayesian Estimation

For solving inverse problems, deterministic or probabilistic approaches can be applied. In this paper, a probabilistic approach employing *Bayesian inference* is considered. In doing so, uncertainties arising for instance from sensor noise or modeling errors can be incorporated.

According to Bayes' theorem (see e.g. Särkkä [2013]), the so-called *posterior density* $p(\underline{\theta}|\underline{z})$ of $\underline{\theta}$ is calculated according to

$$p(\underline{\theta}|\underline{z}) = \frac{p(\underline{z}|\underline{\theta}) \cdot p(\underline{\theta})}{p(\underline{z})} \ , \qquad (4)$$

which is the conditional probability of the unknown model parameters given the measurements. This density function represents the solution of the inverse problem. In (4), $p(\underline{z}|\underline{\theta})$ is the *likelihood*, $p(\underline{\theta})$ is the *prior density*, and $p(\underline{z}) = \int p(\underline{z}|\underline{\theta}) \cdot p(\underline{\theta}) \, \mathrm{d}\underline{\theta}$ is a normalization constant.

By inspecting (4) it can be seen that all concentration measurements are processed at once. Under weak assumptions however, Bayes' theorem also allows an recursive calculation of the posterior distribution. This is especially useful, when the concentration measurements z_k are acquired over time at discrete time steps t_k with $k = 1, 2, \ldots$ and one is interested in constantly updating the posterior. Assuming that the concentration measurements z_k are conditionally

independent given the model parameters, (4) can be reformulated in a recursion

$$p(\underline{\theta}|z_{1:k}) = \frac{p(z_k|\underline{\theta}) \cdot p(\underline{\theta}|z_{1:k-1})}{p(z_k|z_{1:k-1})} \tag{5}$$

which commences from the prior $p(\underline{\theta})$ and where $z_{1:k} \triangleq (z_1, z_2, \ldots, z_k)$ is the collection of all measurements up to and including z_k.

3.2 Measurement Model

While the prior reflects the a priori knowledge of the user on the source, the likelihood $p(z_k|\underline{\theta})$ relates a concentration measurement to the unknown source parameters. In order to define the likelihood, one can make use of the Gaussian plume model derived above, assuming that this model represents the underlying dispersion mechanism. For this purpose, suppose that the k-th measurement z_k is acquired by a sensor at location $\underline{x}_r \triangleq [x_r, y_r, z_r]^{\mathrm{T}}$ at time t_k. The resulting measurement model is given by

$$z_k = c(\underline{x}_r; \underline{\theta}) + v_k \ , \quad v_k \sim p(v_k) \triangleq \mathcal{N}(0, \sigma_v^2) \ , \tag{6}$$

where $c(\underline{x}_r; \underline{\theta})$ is the true concentration value according to (2) and v_k is the sensor's noise, which is assumed to be zero-mean Gaussian with variance σ_v^2 and independent in time and space. The measurement model (6) can be turned into a statistical model according to

$$p(z_k|\underline{\theta}) = \int \underbrace{p(z_k|v_k, \underline{\theta})}_{\delta\left(z_k - c(\underline{x}_r;\underline{\theta}) - v_k\right)} \cdot p(v_k) \, \mathrm{d}v_k = \mathcal{N}\left(z_k; c(\underline{x}_r; \underline{\theta}), \sigma_v^2\right) \ , \tag{7}$$

where the second equality follows from the sifting property of the Dirac delta distribution $\delta(.)$. This completes the derivation of the Bayesian estimation formalism.

It is worth mentioning that in case of multiple sources, the concentration measurement can be written as

$$z_k = \sum_{i=1}^{N} c_i(\underline{x}_r; \underline{\theta}_i) + v_k \ , \tag{8}$$

where the superposition of the concentration values $c_i(\underline{x}_r; \underline{\theta}_i)$ of the sources $i = 1, \ldots, N$ is exploited (see Stockie [2011]). The likelihood for multiple sources can be derived analogously as described above.

3.3 Approximate Estimation

Unfortunately, the Bayesian formalism in (5) is merely of limited practical use as a closed-form solution of the recursion is not available in general. Except for some special cases like the linear Gaussian one, an approximate solution is inevitable. This holds also for the considered source estimation problem due to the nonlinear Gaussian plume model (2) that forms the likelihood. Typical approximations of nonlinear Bayesian estimation problems rely on Monte Carlo methods like particle filters (see e.g. Arulampalam et al. [2002]). However, the obtained results are not reproducible and scaling for high dimensional parameters is an issue. Alternatives are Gaussian filters like the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) (see e.g. Särkkä [2013]), where the posterior is approximated by means of a Gaussian distribution. These approaches scale well with high dimensions, but the approximation can be poor in case of strong nonlinearities. In this paper, an approximate Bayesian estimator named *adaptive Gaussian mixture filter* (AGMF) is employed that provides Gaussian mixture posteriors. As a result, significantly better approximations compared to Gaussian filters are obtained with an at the same time low computational demand and good scalability.

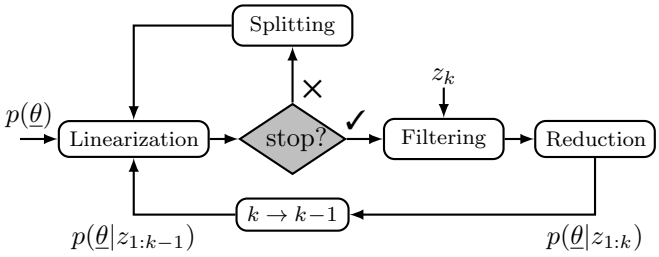

Fig. 1. Flow chart of the adaptive Gaussian mixture filter.

4. ADAPTIVE GAUSSIAN MIXTURE FILTER

In this section, a brief introduction of the AGMF proposed by Huber [2011] is given. It provides a Gaussian mixture representation of the posterior $p(\underline{\theta}|z_{1:k})$ according to

$$p(\underline{\theta}|z_{1:k}) = \sum_{i=1}^{L} \omega_{k,i} \cdot \mathcal{N}\left(\underline{\theta}; \hat{\underline{\theta}}_{k,i}, \mathbf{C}_{k,i}\right) \ ,$$

where L is the number of mixture components, $\omega_{k,i}$ are non-negative weights summing up to one, and $\hat{\underline{\theta}}_{k,i}$, $\mathbf{C}_{k,i}$ are the mean and covariance matrix, respectively, of the i-th Gaussian component. By substituting this mixture representation into the Bayesian recursion (5), one has to evaluate two terms: the product of the Gaussian mixture with the likelihood (7) and the integration required for the normalization constant $p(z_k|z_{1:k-1})$. In order to obtain the desired quantities, the AGMF performs four steps as depicted in Fig. 1: linearization, splitting, filtering, and reduction. Each step is explained in the following, where the time index k is omitted for improved readability.

4.1 Statistical Linearization

The product of likelihood and mixture boils down to multiple products between the likelihood and each Gaussian component of the mixture. These individual products also appear in standard nonlinear Gaussian filters like the EKF or UKF. The AGMF utilizes the same technique also used in the UKF in order to provide an approximate solution of the individual products. By means of so-called *statistical linearization*, the nonlinearity inducted by the Gaussian plume model is transformed into a linear one, i.e., the nonlinear model (6) is approximated by means of the linear model

$$z \approx \mathbf{H}_i \cdot \underline{\theta} + b_i + v \ . \tag{9}$$

The terms $\mathbf{H}_i$ and b_i are obtained via statistical linear regression, where the Gaussian plume model $c(.)$ is evaluated at a set of weighted regression points $\mathcal{L} \triangleq \left\{\alpha_j^{(i)}, \underline{\theta}_j^{(i)}\right\}_{j=0\ldots N}$ with non-negative weights $\alpha_j^{(i)}$. The regression points are drawn *deterministically* from the i-th Gaussian component in such a way that sample mean and covariance of $\mathcal{L}$ coincide with the mean $\hat{\underline{\theta}}_i$ and covariance $\mathbf{C}_i$. The solution of the linear regression yields

$$\mathbf{H}_i = \left(\mathbf{C}^{\theta z}\right)^{\mathrm{T}} \mathbf{C}_i^{-1} \quad \text{and} \quad b_i = \hat{z} - \mathbf{H}_i \cdot \hat{\underline{\theta}}_i \tag{10}$$

for the required terms in (9), where $\mathcal{L}$ is used to approximate the mean and cross-covariance of z according to

$$\hat{z} \approx \sum_{j=0}^{N} \alpha_j^{(i)} \cdot z_j^{(i)} \ , \quad \mathbf{C}^{\theta z} \approx \sum_{j=0}^{N} \alpha_j^{(i)} \cdot \left(\underline{\theta}_j^{(i)} - \hat{\underline{\theta}}\right) \cdot \left(z_j^{(i)} - \hat{z}\right)^{\mathrm{T}} ,$$

respectively, with $z_j^{(i)} = c\left(.; \underline{\theta}_j^{(i)}\right)$ for $j = 0 \dots N$. The terms in (10) minimize the square of the error

$$e_i \triangleq c(.; \underline{\theta}) - \mathbf{H}_i \cdot \underline{\theta} - b_i \tag{11}$$

in such a way that the error has zero mean and variance

$$\sigma_{e,i}^2 = \sigma_{z,i}^2 - \mathbf{H}_i \mathbf{C}_i \mathbf{H}_i^{\mathrm{T}} - \sigma_v^2 \ , \tag{12}$$

where $\sigma_{z,i}^2 \approx \sum_j \alpha_j^{(i)} \cdot \left(z_j^{(i)} - \hat{z}\right)\left(z_j^{(i)} - \hat{z}\right)^{\mathrm{T}} + \sigma_v^2$. The error variance (12) gives a good indication of the linearization error as it is merely zero iff $c(.)$ is affine. The Gaussian plume model approaches an affine function with an increasing distance between the sensor location $\underline{x}_r$ and the gas source.

4.2 Splitting

As statistical linearization (9) has to be performed for every component of the Gaussian mixture $p(\underline{\theta}|z_{1:k-1})$, it is called *local linearization* in the following. It typically provides a better estimation performance compared to a single *global* linearization (see Alspach and Sorenson [1972]). A further performance improvement can be achieved by increasing the number of mixture components. It was proven in Ali-Löytty [2009] that Gaussian mixture filters relying on a local linearization converge towards the optimal estimate when increasing the component number.

The AGMF adds additional components to the mixture by *splitting* existing ones, i.e., an existing component is replaced by several new components that share some statistics with the original component and that have lower weights and covariances. To select a component for splitting, AGMF takes the local linearization error induced by each component into account. Here, the statistical linearization described above already provides a measure via the error variance σ_e^2 in (12). Besides the linearization error, also the importance of a mixture component, which is given by its weight ω_i, is crucial. Both ingredients are combined in the so-called *selection criterion*

$$i^* = \arg \min_{i=1\dots L} \{s_i\} \ , \ s_i \triangleq \omega_i^{\gamma} \cdot \left(1 - \exp(-\sigma_{e,i}^2)\right)^{1-\gamma} \tag{13}$$

in order to select component i^* for splitting. Here, the term $1 - \exp(-\sigma_{e,i}^2)$ normalizes the error $\sigma_{e,i}^2$ of the i-th component to the interval $[0, 1]$. The criterion considers the component weight and linearization error through a geometric interpolation with parameter $\gamma \in [0, 1]$, where for $\gamma = 0$ only the linearization error is the determining factor and for $\gamma = 1$ it is the weight.

In order to trade the reduction of the linearization error off against controlling the computational load and the growth of the number of components, the Gaussian component selected for splitting is replaced by two Gaussians only in each splitting round. As splitting is performed recursively by the AGMF, the newly introduced components can be split again in the next rounds if the linearization error is still too high.

Let $\omega \cdot \mathcal{N}(\underline{\theta}; \hat{\underline{\theta}}, \mathbf{C})$ be the component considered for splitting. It is replaced by two components according to

$$\omega \cdot \mathcal{N}(\underline{\theta}; \hat{\underline{\theta}}, \mathbf{C}) \approx \sum_{n=1}^{2} \omega_n \cdot \mathcal{N}(\underline{\theta}; \hat{\underline{\theta}}_n, \mathbf{C}_n) \ ,$$

with parameters

$$\begin{aligned} &\omega_1 = \omega_2 = \tfrac{\omega}{2} \ , \\ &\hat{\underline{\theta}}_1 = \hat{\underline{\theta}} + \sqrt{\lambda} \cdot \mu \cdot \underline{\epsilon} \ , \quad \hat{\underline{\theta}}_2 = \hat{\underline{\theta}} - \sqrt{\lambda} \cdot \mu \cdot \underline{\epsilon} \ , \\ &\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C} - \lambda \cdot \mu \cdot \underline{\epsilon}\,\underline{\epsilon}^{\mathrm{T}} \ , \end{aligned} \tag{14}$$

where $\mu \in [-1, 1]$ is a free parameter. The parametrization in (14) ensures moment preservation, i.e., the original Gaussian component and its split counterpart have the same mean and covariance. Furthermore, λ and $\underline{\epsilon}$ in (14) are a particular eigenvalue and eigenvector, respectively, of $\mathbf{C}$. Splitting is merely performed along eigenvectors of $\mathbf{C}$, which is computationally cheap and numerically stable compared to arbitrary splitting directions. Among all possible eigenvectors, the one is chosen that currently induces the highest error according to (11), i.e., splitting is performed along the eigenvector where the Gaussian plume model posses the strongest nonlinearity.

As indicated in Fig. 1, in every splitting round a *stopping criterion* is evaluated. Splitting stops, if at least one of the three following thresholds is reached:

Error threshold: The value s_i in the selection criterion (13) drops below $s_{\max} \in [0, 1]$ for *every* component.

Component threshold: The number of mixture components excels $L_{\max}$.

Deviation threshold: The deviation between the original Gaussian mixture $p(\underline{\theta})$ and the mixture obtained via splitting $\tilde{p}(\underline{\theta})$ excels $d_{\max} \in [0, 1]$.

The deviation considered for the latter threshold is determined by means of the normalized integral squared distance measure

$$D\left(p(\underline{\theta}), \tilde{p}(\underline{\theta})\right) = \frac{\int \left(p(\underline{\theta}) - \tilde{p}(\underline{\theta})\right)^2 \mathrm{d}\underline{\theta}}{\int p(\underline{\theta})^2 \,\mathrm{d}\underline{\theta} + \int \tilde{p}(\underline{\theta})^2 \,\mathrm{d}\underline{\theta}} \ \in [0, 1] \ .$$

Since splitting always introduces an approximation error to the original mixture, continuously monitoring the deviation limits this error.

4.3 Filtering

Let $\omega_i^s \cdot \mathcal{N}(\underline{\theta}; \hat{\underline{\theta}}_i^s, \mathbf{C}_i^s)$ be the Gaussians resulting from splitting with $i = 1 \dots L^s$ and $L^s \in [L, L_{\max}] \subset \mathbb{N}$. Given the concentration measurement z_k, the recursive Bayesian update according to (5) boils down to a bank of Kalman filter updates thanks to the locally linearized models (9). Thus, the update of each (prior) Gaussian component gives rise to the parameters of the corresponding component of the posterior Gaussian mixture $p(\underline{\theta}|z_{1:k})$ according to

$$\begin{aligned} \omega_i &= c \cdot \omega_i^s \cdot \mathcal{N}(z_k; \hat{z}_i, \sigma_{z,i}^2) \ , \\ \hat{\underline{\theta}}_i &= \hat{\underline{\theta}}_i^s + \mathbf{K}_i \cdot (z_k - \hat{z}_i) \ , \\ \mathbf{C}_i &= \mathbf{C}_i^s - \mathbf{K}_i \mathbf{H}_i \mathbf{C}_i^s \ , \end{aligned} \tag{15}$$

with predicted measurement $\hat{z}_i = \mathbf{H}_i \cdot \hat{\underline{\theta}}_i^s + b_i$, $\mathbf{H}_i$ and b_i according to (10), Kalman gain $\mathbf{K}_i = {\mathbf{C}_i^s \mathbf{H}_i^{\mathrm{T}}}/{\sigma_{z,i}^2}$, and innovation variance $\sigma_{z,i}^2 = \mathbf{H}_i \mathbf{C}_i^s \mathbf{H}_i^{\mathrm{T}} + \sigma_v^2 + \sigma_{e,i}^2$ with $\sigma_{e,i}^2$ being the linearization error variance (12). In the calculation of the weight ω_i in (15), $c = {1}/{\sum_i \omega_i^s \cdot \mathcal{N}(z_k; \hat{z}_i, \sigma_{z,i}^2)}$ is a normalization constant.

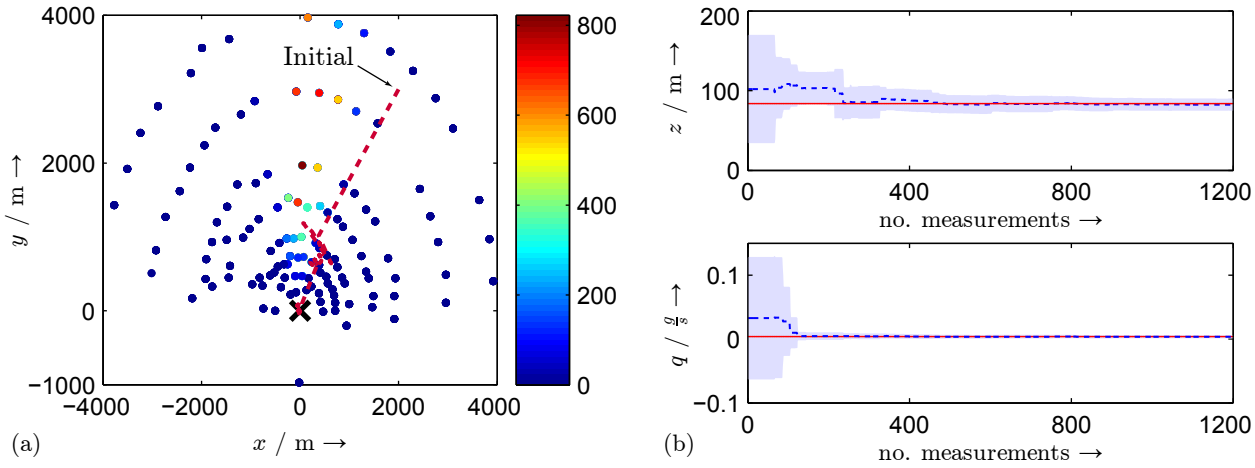

Fig. 2. Source estimate by AGMF based on the data from the Indianapolis field study. (a) The red dashed line marks the trajectory of the estimated source location $[x, y]^{\mathrm{T}}$, whereas the true location of the source is marked by the black cross. Circular markers denote the sensor locations colored with the measured concentration in ppt. (b) Estimate of source height z and emission rate q with increasing number of measurements. The shaded area denotes the 3-sigma confidence region and the red line indicates the true value.

4.4 Reduction

As the number of components of $p(\underline{\theta}|z_{1:k})$ grows due to splitting, it is necessary to bound this growth in order to reduce the computational and memory demand of subsequent estimation steps. For this purpose, one can exploit the redundancy and similarity of Gaussian components. Furthermore, many components will have negligible weights and thus, they can be removed without introducing significant errors. To reduce a Gaussian mixture, many algorithms have been proposed in the recent years (see e.g. West [1993], Runnalls [2007], Huber and Hanebeck [2008]). Most of these algorithms require a *reduction threshold* $L_{\mathrm{red}} \ll L^s$ to which the number of components of the given Gaussian mixture has to be reduced. The reduction to L_{red} components closes the calculation of the posterior Gaussian mixture in (5) for time step k.

5. RESULTS

The estimation performance of the AGMF is assessed in the following by means of real data in Section 5.2 and by means of a synthetic example in Section 5.3. However, as the AGMF is a generic estimator allowing specific parametrization depending on the estimation task, the selected setup is described first.

5.1 AGMF Parametrization

The three processing steps linearization, splitting, and reduction allow specialized parametrization. For the statistical linearization, the selection of the regression points is based on the scaled unscented transform (Julier and Uhlmann [2004]). Hence, the points and their weights are given by

$$\begin{aligned} \underline{\theta}_0 &= \hat{\underline{\theta}}\,, & \alpha_0 &= \tfrac{\lambda^2-d}{\lambda^2}\,, & \\ \underline{\theta}_j &= \hat{\underline{\theta}} + \lambda \cdot \underline{P}_j\,, & \alpha_j &= \tfrac{1}{2\lambda^2}\,, & j &= 1 \dots d\,, \\ \underline{\theta}_{d+j} &= \hat{\underline{\theta}} - \lambda \cdot \underline{P}_j\,, & \alpha_{d+j} &= \alpha_j\,, & j &= 1 \dots d\,, \end{aligned}$$

with d being the dimension of $\underline{\theta}$, $\lambda \triangleq \nu \cdot \sqrt{d+\kappa}$ being a scaling factor, and $\underline{P}_j$ being the j-th column of the matrix $\mathbf{P} = \sqrt{\mathbf{C}}$, where the matrix square root is calculated via the Cholesky decomposition. The free parameters in λ are chosen to be $\nu = 1$, $\beta = 2$, and $\kappa = 0.5$ in the following experiments, which leads to equal weights $\alpha^{(j)}$ for all $j = 0 \dots 2d$.

Both the mixing parameter γ in the selection criterion (13) and the displacement parameter μ in (14) are set to be 0.5. The thresholds for stopping component splitting are chosen to be $s_{\max} = 0$ and $d_{\max} = 1$, which actually disables these thresholds. Only the threshold $L_{\max}$ is active, but set differently depending on the considered experiment.

For reducing the Gaussian mixture after performing the filtering step, the reduction algorithm proposed by Runnalls [2007] is employed as it provides a good trade-off between computational demand and reduction error. The reduction threshold L_{red} is set to be $L_{\max}/8$.

5.2 Indianapolis Field Study

In the first experiment, it is demonstrated how the proposed source estimation solution performs on a real data set. For this purpose, the data acquired during the EPRI Indianapolis field study is considered, where a SF_6 tracer gas was released from a $z_s = 83.8$ m stack at a power plant in Indianapolis, Indiana, USA. Data was recorded by 160 ground-level sensors over 19 days in September and October 1985 for 8 to 9 hours every day. Details about the field study and the data can be found in Hanna et al. [1997].

In Fig. 2(a), the locations of the sensors and the sensors' concentration measurements are depicted for the 19$^{\mathrm{th}}$ September 1985. The source is located at the origin and the emission rate of the tracer gas is $q = 0.0041\,\mathrm{g/s}$. Information about wind speed, wind direction, and atmospheric stability was made available by meteorological

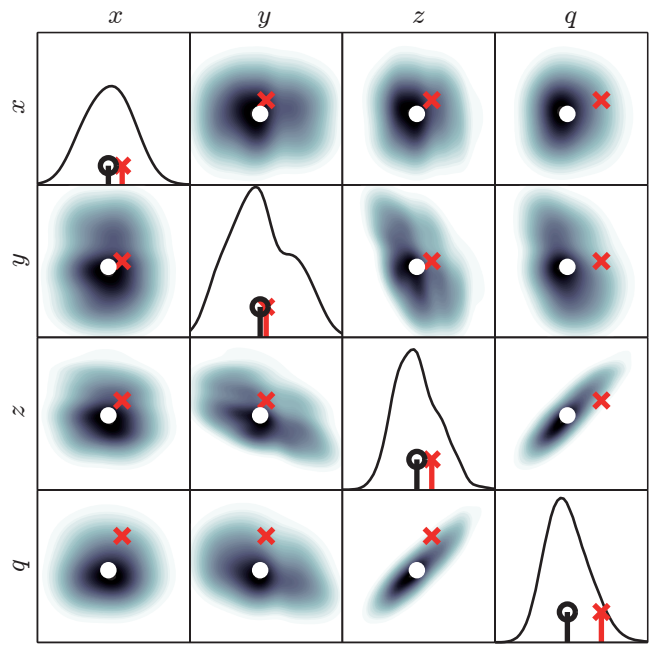

Fig. 3. Bivariate posterior densities of the AGMF source estimate. The diagonal plots are the univariate marginal densities. Red crosses indicate the true value, while white and black circles, respectively, denote the mean of the respective density.

observations. The initial estimate of the source at time step $k = 0$ is given by a single Gaussian with mean vector $\hat{\underline{\theta}}_0 = [2000, 3000, 100, 0.033]^{\mathrm{T}}$ and covariance matrix $\mathbf{C}_0 = \mathrm{diag}(10^6, 10^6, 500, 0.001)$. Fig. 2(a) and (b) show the convergence of the source estimate towards the true source location over time and with increasing number of concentration measurements, respectively. It is important to note that many sensor measurements (typically 60%-70%) provide a concentration measurement of almost zero as most of the sensors are outside the gas plume, as can be seem in Fig. 2(a). This explains the discontinuous convergence of the estimate and reduction of the variance in Fig. 2(b).

The posterior density $p(\underline{\theta}|z_{1:k})$ after all $k = 1200$ measurements is depicted in Fig. 3. It can be seen that the mean of the estimate is close to the true source parameters. Slight deviation from the ground truth is only observed for the emission rate, but the true parameters are still within the high confidence region of the estimate. Thus, the proposed estimator is not overconfident.

5.3 Simulation

In contrast to the previous experiment, a synthetic example is considered here allowing the comparison with state-of-the-art source estimation methods via Monte Carlo simulations. In this example, the locations and emission rates of two sources have to be estimated. Thus, the measurement model (8) for $N = 2$ applies here. The employed simulation parameters are listed in Table 1. The standard deviations σ_y, σ_z of the Gaussian plume model are assumed to take the form

$$\sigma(x) = a \cdot x \cdot (1 + b \cdot x)^{-c} \tag{16}$$

Table 1. Simulation parameters, where $\mathcal{U}(a, b)$ is a uniform distribution over the interval $[a, b]$.

Parameter	Distribution / Value
location 1st source	$x_{s,1},\ y_{s,1} \sim \mathcal{U}(0\,\mathrm{m}, 100\,\mathrm{m})$
height 1st source	$z_{s,1} \sim \mathcal{N}(4\,\mathrm{m}, 1\,\mathrm{m}^2)$
emission rate 1st source	$q_1 \sim \mathcal{U}(0.005\,\mathrm{g/s}, 0.006\,\mathrm{g/s})$
location 2nd source	$x_{s,1},\ y_{s,2} \sim \mathcal{U}(0\,\mathrm{m}, 100\,\mathrm{m})$
height 2nd source	$z_{s,2} \sim \mathcal{N}(6\,\mathrm{m}, 1\,\mathrm{m}^2)$
emission rate 2nd source	$q_2 \sim \mathcal{U}(0.0075\,\mathrm{g/s}, 0.0085\,\mathrm{g/s})$
wind speed	$u \sim \mathcal{U}(1\,\mathrm{m/s}, 2\,\mathrm{m/s})$
wind direction	$\phi \sim \mathcal{U}(-\pi/4\,\mathrm{rad}, \pi/4\,\mathrm{rad})$
standard deviation σ_y	$a = 0.08,\ b = 0.0001,\ c = 0.5$
standard deviation σ_z	$a = 0.06,\ b = 0.0015,\ c = 0.5$

according to Briggs [1973]. The constants in (16) depend on the atmospheric stability, for which class D is assumed—corresponds to "neutral" in accordance to the Pasquill-Gifford classification scheme. As described in Carrascal et al. [1993], the corresponding values of the constants for class D are as listed in the last two rows of Table 1.

For comparison, the following estimators are considered:

- AGMF_{16} Proposed source estimator with $L_{\max} = 16$.
- AGMF_{32} Proposed source estimator with $L_{\max} = 32$.
- GMF_{16} Gaussian mixture estimator by Simandl and Duník [2005], i.e., AGMF without adaptation. The number of components remains constant at 16.
- GMF_{32} Same estimator as GMF_{16} but with 32 mixture components.
- UKF Unscented Kalman filter proposed by Julier and Uhlmann [2004], i.e., only a single Gaussian represents the posterior.
- GN Backward off-line method utilizing Gauß-Newton optimization proposed by Rudd et al. [2012].
- MCMC Forward off-line method utilizing Markov chain Monte Carlo sampling. The number of sampling steps is chosen in such a way that the runtime of MCMC is similar to AGMF_{32}.

The initial estimate of each estimator at time step $k = 0$ is Gaussian with mean vector $\hat{\underline{\theta}}_0$ drawn randomly from $\mathcal{N}(\underline{\theta}; \mathbf{C})$ with

$$\underline{\theta} = [x_{s,1}, y_{s,1}, z_{s,1}, q_1, x_{s,2}, y_{s,2}, z_{s,2}, q_2]^{\mathrm{T}} ,$$
$$\mathbf{C} = \mathrm{diag}\left(50^2, 50^2, 10^2, 0.0025^2, 50^2, 50^2, 10^2, 0.0025^2\right) .$$

The initial covariance matrix is set to be $\mathbf{C}_0 = \mathbf{C}$.

Three different sensor noise levels for σ are considered: $10\,\mathrm{mg/m^3}$ (strong noise), $5\,\mathrm{mg/m^3}$ (medium noise), and $1\,\mathrm{mg/m^3}$ (low noise). For each noise level, 100 Monte Carlo simulation runs are performed. In each run, 600 concentration measurements are acquired at locations drawn randomly from $\mathcal{U}(0\,\mathrm{m}, 200\,\mathrm{m}) \times \mathcal{U}(0\,\mathrm{m}, 200\,\mathrm{m})$. In Table 2, average values of the distance between source estimates and true source locations for both sources as well as the average absolute deviation between estimated and true emission rate are listed. For low and medium noise, the proposed AGMF provides the lowest source distance, where allowing more components for splitting leads to a slightly better estimation performance. Regarding the emission rate, GN

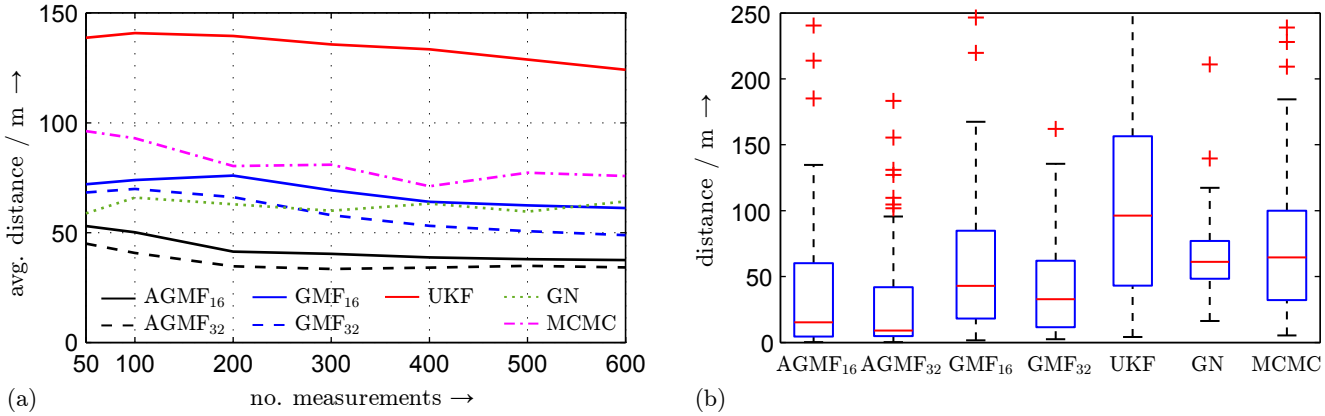

Fig. 4. Distances between estimated source position and true source positions for the low noise case. (a) Average distance depending on the number of measurements. For the batch methods GN and MCMC, for each number of measurements separate estimation runs have been performed. (b) Median (red line), lower and upper quantiles (blue box), and spread (black lines) of the distances for 600 measurements. Red crosses indicate outliers.

Table 2. Average distance and average source rate deviation in case of two sources.

	Strong Noise		Med. Noise		Low Noise	
	dist	rate	dist	rate	dist	rate
$AGMF_{16}$	49.9	29.0	37.1	31.9	37.5	27.3
$AGMF_{32}$	45.6	28.5	**34.6**	32.1	**34.2**	28.1
GMF_{16}	58.0	27.2	42.6	30.9	61.2	36.8
GMF_{32}	52.2	28.2	41.6	31.1	48.8	33.7
UKF	69.9	29.1	74.5	34.7	124.1	59.8
GN	63.8	**21.8**	63.5	**21.3**	64.2	**21.6**
MCMC	**34.3**	25.5	43.0	29.9	75.8	36.3

is the most accurate estimator. In case of strong noise, MCMC is the best estimator with respect to the source location. MCMC however, is an off-line estimator, i.e., estimates become available after processing all measurements in a batch, while the AGMF continuously provides estimates at runtime.

The superior estimation performance of the AGMF compared to the other estimators in case of low noise does not vary with the number of concentration measurements as shown in Fig. 4(a). From 50 to 600 measurements AGMF always provides the smallest distance between estimated and true source location. A similar result is obtained for the other noise levels, except that MCMC always provides a better estimate for the strong noise case.

Fig. 4(b) shows that the majority of the estimates provided by the AGMF is better than the average listed in Table 2. Some outliers significantly lower the average performance. These outliers result from random initializations, where the initial estimated source location is far outside the sensor area $[0\,\text{m}, 200\,\text{m}] \times [0\,\text{m}, 200\,\text{m}]$. Due to the potentially large distance between estimate and measurement location, the Gaussian plume model becomes almost linear, which is also indicated by the linearization error measure (12). As a consequence no splitting is performed and thus, the AGMF degrades to a simple UKF. By means of allowing also splitting towards the sensor location would resolve this issue and would result in a better performance.

Of all on-line estimators, AGMF performs best. This clearly shows that a single Gaussian and also a fixed Gaussian mixture representation are not sufficient. By means of the adaptation via splitting, a much better source estimation is possible. It is worth mentioning that it is not necessary to utilize the maximum number of components $L_{\max}$ at all time steps in order to provide a meaningful representation of the posterior density. As shown in Table 3, the average number of components per time step is significantly below $L_{\max}$, whereas the lower the measurement noise the lower the required number of components. Thus, the proposed AGMF can carefully control the demand of splits.

Table 3. Average and standard deviation of number of mixture components per time step.

	Strong Noise	Med. Noise	Low Noise
$AGMF_{16}$	12.46 ± 8.07	9.77 ± 8.93	8.59 ± 8.96
$AGMF_{32}$	25.13 ± 16.02	19.64 ± 17.85	16.53 ± 17.86

6. CONCLUSION

The state-of-the-art in dispersion source estimation mainly focuses on MCMC methods. While this allows accurate estimates, only batch processing is possible. The proposed adaptive Gaussian mixture filter shows that a similar and in some cases an even better estimation performance can be obtained with an on-line estimator. Therefore, it is of paramount importance to not rely on a single Gaussian or a fixed Gaussian mixture representation. By means of splitting that adapts the Gaussian mixture to the nonlinearity of the Gaussian plume dispersion model, a significant improvement can be achieved.

REFERENCES

Simo Ali-Löytty. *Gaussian Mixture Filters in Hybrid Positioning*. PhD thesis, Tampere University of Technology, Tampere, Finland, August 2009.

Daniel L. Alspach and Harold W. Sorenson. Nonlinear Bayesian Estimation using Gaussian Sum Approximation. *IEEE Transactions on Automatic Control*, 17(4): 439–448, August 1972.

M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A Tutorial on Particle Filters

for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, February 2002.

Mieczyslaw Borysiewicz, Anna Wawrzynczak, and Piotr Kopka. Bayesian-Based Methods for the Estimation of the Unknown Model's Parameters in the Case of the Localization of the Atmospheric Contamination Source. *Foundations of Computing and Decision Sciences*, 37(4): 253–270, 2012. doi: DOI: 10.2478/v10209-011-0014-9.

G. A. Briggs. Diffusion estimation for small emissions, Atmospheric Turbulence and Diffusion Laboratory Contribution. Technical Report 79, National Oceanic and Atmospheric Administration, Oak Ridge, 1973.

M. D. Carrascal, M. Puigcerver, and P. Puig. Sensitivity of Gaussian plume model to dispersion specifications. In *Theoretical and Applied Climatology*, volume 48, pages 147–157. Springer, 1993.

Steven Hanna, Joseph Chang, and Helge R. Olesen. *Indianapolis Tracer Data and Meteorological Data*, May 1997.

Bill Hirst, Philip Jonathan, Fernando G. del Cueto, David Randell, and Oliver Kosut. Locating and quantifying gas emission sources using remotely obtained concentration data. *Atmospheric Environment*, 74:141–158, August 2013.

Ekkehard Holzbecher. *Environmental Modeling*. Springer, 2nd edition, 2012.

Frédéric Hourdin and Olivier Talagrand. Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parametrization of subgrid-scale transport. *Quarterly Journal of the Royal Meteorological Society*, 132 (615):567–583, January 2006.

Marco F. Huber. Adaptive Gaussian Mixture Filter Based on Statistical Linearization. In *Proceedings of the 14th International Conference on Information Fusion (Fusion)*, Chicago, Illinois, July 2011.

Marco F. Huber and Uwe D. Hanebeck. Progressive Gaussian Mixture Reduction. In *Proceedings of the 11th International Conference on Information Fusion (Fusion)*, Cologne, Germany, July 2008.

Simon J. Julier and Jeffrey K. Uhlmann. Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

Vladimir Maz'ya and Gunther Schmidt. On approximate approximations using gaussian kernels. *IMA Journal of Numerical Analysis*, 16:13–29, 1996.

K. Shankar Rao. Source estimation methods for atmospheric dispersion. *Atmospheric Environment*, 41:6964–6973, 2007.

Alison Rudd, Alan G. Robins, Jason J. Lepley, and Stephen E. Belcher. An Inverse Method for Determining Source Characteristics for Emergency Response Applications. *Boundary-Layer Meteorology*, 144(1):1–20, July 2012.

Andrew R. Runnalls. Kullback-Leibler Approach to Gaussian Mixture Reduction. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3):989–999, July 2007.

Simo Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

Inanc Senocak, Nicolas W. Hengartner, Margaret B. Short, and W. Brent Daniel. Stochastic Event Reconstruction of Atmospheric Contaminant Dispersion Using Bayesian Inference. *Atmospheric Environment*, 42(33):7718–7727, October 2008.

Miroslav Simandl and Jindrich Duník. Sigma Point Gaussian Sum Filter Design Using Square Root Unscented Filters. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 2005.

Michael D. Sohn, Pamela Reynolds, Navtej Singh, and Ashok J. Gadgil. Rapidly Locating and Characterizing Pollutant Releases in Buildings. *Journal of the Air & Waste Management Association*, 52(12):1422–1432, 2002.

John M. Stockie. The Mathematics of Atmospheric Dispersion Modelling. *SIAM Review*, 53(2):349–372, 2011.

Mike West. Approximating Posterior Distributions by Mixtures. *Journal of the Royal Statistical Society: Series B*, 55(2):409–422, 1993.

Yong Zhang and Li Wang. Particle Filtering Method for Source Localization in Wireless Sensor Network. In *Advanced Technology in Teaching: Selected papers from the 2012 International Conference on Teaching and Computational Science (ICTCS 2012)*, volume 163, pages 517–523. Springer, 2013.