The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

A Statistical Approach to Gas Distribution Modelling with Mobile Robots – The Kernel DM+V Algorithm

Achim J. Lilienthal, Matteo Reggente, Marco Trincavelli, Jose Luis Blanco and Javier Gonzalez

***Abstract*— Gas distribution modelling constitutes an ideal application area for mobile robots, which – as intelligent mobile gas sensors – offer several advantages compared to stationary sensor networks. In this paper we propose the Kernel DM+V algorithm to learn a statistical 2-d gas distribution model from a sequence of localized gas sensor measurements. The algorithm does not make strong assumptions about the sensing locations and can thus be applied on a mobile robot that is not primarily used for gas distribution monitoring, and also in the case of stationary measurements. Kernel DM+V treats distribution modelling as a density estimation problem. In contrast to most previous approaches, it models the variance in addition to the distribution mean. Estimating the predictive variance entails a significant improvement for gas distribution modelling since it allows to evaluate the model quality in terms of the data likelihood. This offers a solution to the problem of ground truth evaluation, which has always been a critical issue for gas distribution modelling. Estimating the predictive variance also provides the means to learn meta parameters and to suggest new measurement locations based on the current model. We derive the Kernel DM+V algorithm and present a method for learning the hyper-parameters. Based on real world data collected with a mobile robot we demonstrate the consistency of the obtained maps and present a quantitative comparison, in terms of the data likelihood of unseen samples, with an alternative approach that estimates the predictive variance.**

I. INTRODUCTION

Gas distribution modelling constitutes an ideal application area for mobile robotics. Acting as intelligent mobile gas sensors, gas-sensitive robots offer several advantages compared to stationary sensor networks. For stationary sensor networks it is a problem that the optimal sensor locations can vary with the environmental conditions. Mobile sensors can provide a distribution model with adaptive (and higher) resolution. Mobile robots that carry the sensors offer the required accurate localization and computational resources to create the distribution model on-line. Thus also the possibility to decide based on the current model which locations to observe next. Compared to human operators, mobile robots have the advantage to carry out the required repetetive measurement procedure without suffering from fatigue.

Gas distribution modelling with mobile robots at smaller scales has important applications in industry, science, and every-day life. Mobile robots equipped with gas sensors are deployed, for example, for pollution monitoring in public areas [1], or can be used for surveillance of industrial facilities producing harmful gases and inspection of contaminated areas within rescue missions.

Achim J. Lilienthal, Matteo Reggente and Marco Trincavelli are with the AASS Research Centre, Dept. of Technology, Örebro University, S-70182 Örebro, Sweden, achim@lilienthals.de, matteo.reggente@oru.se, mti@aass.oru.se

Jose Luis Blanco and Javier Gonzalez are with the Dept. of System Engineering and Automation, University of Malaga, 29071 Malaga, Spain, {jlblanco|jgonzalez}@ctima.uma.es

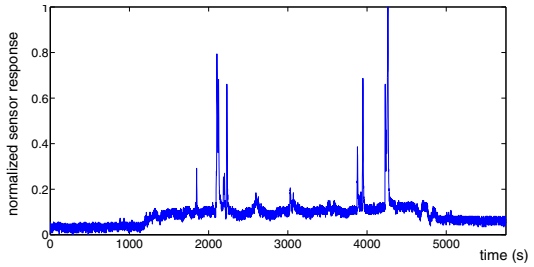

Fig. 1. Normalised raw response readings from an example trial. Corresponding gas distribution models are shown in Fig. 5.

Gas distribution modelling is the task of deriving a truthful representation of the observed gas distribution from a set of spatially and temporally distributed measurements of relevant variables, foremost gas concentration, but also wind and temperature, for example. In this paper we consider the case where only gas sensor measurements are available.

Building gas distribution models is a very challenging task. One main reason is that in many realistic scenarios gas is dispersed by turbulent advection. Turbulent flow creates packets of gas that follow chaotic trajectories [15]. This results in a concentration field that consists of fluctuating, intermittent patches of high concentration. Fig. 1 illustrates gas concentration measurements recorded with a mobile robot along a corridor containing a single gas source. It is important to note that the "noise" is dominated by the large fluctuations of the instantaneous gas distribution and not by electronic noise of the gas sensors.

While an exact description of turbulent flow remains an intractable problem, it is possible to describe turbulent gas distribution *on average* under certain assumptions [5]. We therefore aim at a modelling approach that represents the time-averaged distribution and the expected fluctuations without making strong assumptions about the environmental conditions. With respect to stationary and particularly mobile sensing applications we further want to avoid explicit assumptions about the sensing locations so that the algorithm can be applied on a mobile robot that is not primarily used for gas distribution monitoring, for example.

Many gas distribution models were developed for atmospheric dispersion [12]. RIMPUFF, for example, is a Gaussian puff model used to calculate the dispersion of airborne materials at the mesoscale under the condition of a moderate topography [17]. Such models cannot capture all the relevant aspects of gas propagation with a suffi-

cient level of detail. High resolution models are required particularly at small scales and in typical complex indoor and outdoor settings where critical gas concentrations often have a local character. In principle, CFD (Computational Fluid Dynamics) models can be applied, which try to solve the governing set of equations numerically. However, CFD models are computationally very expensive. They quickly become intractable with increasing resolution and depend sensitively on accurate knowledge of the state of the environment (boundary conditions), which is not available in practical situations.

We propose an alternative approach to gas distribution modelling, that is to create a statistical model of the observed distribution, treating the sensor measurements as random variables. In this paper, we introduce the Kernel DM+V algorithm to learn a statistical two-dimensional distribution model from a sequence of localized sensor measurements. In the presented experiments, gas sensors were carried by a mobile robot, which has a number of advantages (mentioned above). However, the proposed algorithm addresses also the more general case including stationary sensors, since stationary sensors are just a special case of mobile sensors. The Kernel DM+V algorithm is non-parametric in the sense that it makes no assumptions about a particular functional form of the model, which includes that it does not assume certain environmental conditions such as a uniform airflow, for example. The learned model is represented as a pair of 2-d grid maps, one representing the distribution mean and the other one the corresponding predictive variance per grid cell. Instead of estimating the variance of the distribution mean, as could be obtained from a Bayesian solution to gas modelling, we carry out two parallel estimation processes, one for the mean and the other one for the variance. In contrast to the covariance *of the mean*, which only decreases as new observations are processed, our estimate of the variance will adapt to the real variability of gas readings at each location.

To measure the quality of a gas distribution model we cannot conduct a straightforward ground truth evaluation since it is not generally possible "to take a snapshot" of the instanteneous concentration field. This is actually another reason that makes gas distribution modelling difficult in practice. Gas sensors provide information about a small spatial region only since the measurements require direct interaction between the sensor surface and the molecules of the target chemical. As a consequence, it is usually impossible to independently measure the concentration field at the same time and the same height as with the sensors mounted on the robot. The fact that Kernel DM+V models the variance in addition to the mean makes it possible to evaluate distribution models by calculating the data likelihood of unseen measurements. In this paper, we use this standard criterion to compare how well future measurements are predicted by different distribution models. In addition, we observe whether the obtained model is "reasonable" in that it is consistent and complies with the observed environmental conditions and the known gas source location. Thus, a gas distribution model is considered *truthful* if it explains new observations well and allows to identify hidden parameters such as the location of the source of gas, for example.

By its capacity for model evaluation, estimating the predictive variance also provides the means to learn meta parameters. Apart from its importance for model evaluation, the estimation of the predictive variance entails further significant advantages [7], which we only mention here but cannot demonstrate in the paper due to lack of space: First, the data likelihood can be used to determine when the model should be updated or re-initialised. Second, models that include the variance much better fit the particular structure of gas distributions, which exhibit strong fluctuations with considerable spatial variations. Third, the predictive variance is often used in methods that suggest new measurement locations based on the current model (sensor planning). Fourth, referring to an "exotic" but fascinating possibility, the predictive variance is required to integrate gas distribution predictions into probabilistic localization methods [2].

This paper is organized as follows. After a discussion of related work in Section II, we describe the proposed Kernel DM+V algorithm in Section III-B. Then, the experiments are detailed in Sec. IV. In Sec. V we analyse the relative importance of the hyper-parameters and discuss how they can be learned. Finally, we present the experimental evaluation of our work based on gas sensor measurements collected with a mobile robot (Sec. VI) in an unmodified environment and conclude the paper in Sec. VI with a summary and suggestions for future work.

II. RELATED WORK

This section gives an overview of work in the area of gas distribution mapping at small scales with a particular focus on methods that have been developed for mobile robots.

A. Model-based Approaches

Model-based approaches infer the parameters of an analytical gas distribution model from the measurements. As discussed above, the application of complex numerical models based on fluid dynamics simulations is not feasible in practical situations. Simpler analytical models, as in [6], for example, often rest on rather unrealistic assumptions and are of course only applicable for situations in which the model assumptions hold. Approaches based on an analytical model also rely on well-calibrated gas sensors, an established understanding of the sensor-environment interaction and often require knowledge about the source intensity.

B. Statistical Approaches Without Predictive Variance

A common approach to creating a representation of a time-averaged concentration field is to acquire measurements using a fixed grid of gas sensors over a prolonged period of time, and to map average [6] or peak [14] concentrations obtained to the given grid approximation of the environment. Consecutive measurements with a single sensor were used in [3]. To make predictions at locations different from the measurement points bi-cubic or triangle-based cubic interpolation was applied. A problem with such interpolation

methods is that there is no means of "averaging out" instantaneous response fluctuations. Response values that were measured very close to each other appear independently in the gas distribution map and thus the representation tends to get more and more jagged while new measurements are added. This can be seen in the top right of Fig. 5, where an example of a gas distribution map resulting from using trilinear interpolation is shown. This map has to be compared to the maps in the middle row in the same figure, which show equivalent distribution mean maps obtained with the Kernel DM+V algorithm.

Histogram methods reflect the spatial correlation of concentration measurements to some degree by the quantization into histogram bins. The 2-d histogram proposed in [4] accumulates the number of "odor hits" received in an area assigned to the histogram bins. Odor hits are counted whenever the response of a gas sensor exceeds a defined threshold. Disadvantages of this method include the dependency on bin size and selected threshold, that a perfectly even coverage of the inspected area is required, and that only binary measurements are used and so useful information is discarded.

Kernel extrapolation distribution mapping ("Kernel DM") can be seen as an extension of histogram methods. The concentration field is represented in the form of a grid map. Spatial integration is carried out by convolving sensor readings and modelling the information content of the point measurements with a Gaussian kernel [8].

C. Statistical Approaches With Predictive Variance

All the methods discussed so far model the average or the peak gas concentration but not the concentration fluctuations. The Kernel DM+V algorithm proposed in this paper also models the observed distribution variance per grid cell.

Another method that predicts the mean concentration and the corresponding variance uses Gaussian process mixture (GPM) models [16]. It treats gas distribution modelling as a regression problem. Two components of the GPM represent the rather smooth "background signal" and areas of high concentration. The components of the mixture model and a gating function, that decides to which component a data point belongs, are learned using Expectation Maximization (EM). In contrast to the Kernel DM+V approach, the model is represented directly using the training data. Because it requires the inversion of matrices that grow with the number of training samples n, the computational complexity of learning the GPM is $\mathcal{O}(n^3)$. This is addressed in [16] by adaptive sub-sampling of the observations to obtain a sparse training set. The sparsification of the training data is integrated into the EM-based learning procedure. Similarly to the Kernel DM+V approach, the dependancy between nearby locations is modelled in the GPM approach by a radially symmetric, squared exponential covariance function.

III. KERNEL DM+V

In this section, we introduce the basic ideas of the Kernel DM+V algorithm and develop the underlying equations in a step-by-step manner.

A. Preliminary Remarks and Assumptions

The general gas distribution modelling problem addressed here is to learn a predictive model of a measurement z at the query location $\mathbf{x}$

$$p(z|\mathbf{x}, \mathbf{x}_{1:n}, z_{1:n}), \tag{1}$$

given a set of measurements $z_{1:n}$ taken at locations $\mathbf{x}_{1:n}$. All the approaches reported in Sec. II and also the Kernel DM+V method proposed in this paper, are used to learn a two dimensional spatial model that represents time-constant structures in the gas distribution. While the statistical approaches are not generally restricted to represent a 2-d distribution, the assumption that the model is learned from measurements, which are generated by a time-constant random process, will not generally be valid. However, this assumption is often made in indoor environments [18] and suggestions how to handle this issue are given in Section VIII. The sample index $i \in [1, n]$ in Eq. (1) corresponds to a time t_i when the measurement was performed. Due to the assumption that the samples are generated by an underlying time-constant random process the measurement time does not have to be considered explicitly. The sample index i is only used to identify individual samples. Please note that in order to avoid calibration issues, which occur because the metal oxide gas sensors used in our experiments rarely reach the equilibrium state when exposed to the quickly fluctuating gas distribution, we model the sensor response r directly. In order to compensate for drift issues and individual variations between different gas sensors, "raw" response values R_i are normalised to $r \in [0, 1]$ as

$$r_i = \frac{R_i - \min(\{R_i\})}{\max(\{R_i\}) - \min(\{R_i\})}. \tag{2}$$

A further assumption we make is that the response is caused by a single target gas, i.e. we do not consider problems related to interferents or simultaneous mapping of multiple odours. In principle the proposed method can be extended to the case of multiple odour sources as described in [11]. In this paper, we also assume perfect knowledge about the position $\mathbf{x}_i$ of a sensor at the time of the measurement. To account for the uncertainty about the sensor position, the method in [10] can be used.

B. The Kernel DM+V Algorithm

Inspired by the Parzen window method [13], Kernel DM+V treats distribution modelling as a density estimation problem. As a non-parametric estimation approach, it makes no assumptions about the particular functional form of the modelled gas distribution. Gas sensor measurements are interpreted as noisy samples from the distribution we wish to estimate given the set of samples $z_{1:n} = r_{1:n}$. In contrast to the estimation of probability density functions we do not sample from the gas distribution directly when creating the gas distribution map. It is therefore necessary to make the assumption that the trajectory of the sensors roughly covers the available space. A perfectly even coverage, however, is not necessary.

Kernel DM+V uses a uni-variate Gaussian weighting function $\mathcal{N}$ to represent the importance of measurement r_i obtained at location $\mathbf{x}_i$ to model the gas distribution at grid cell k. First, two temporary grid maps are computed: $\Omega^{(k)}$ by integrating importance weights and $R^{(k)}$ by integrating weighted readings as

$$\begin{aligned} \Omega^{(k)} &= \textstyle\sum_{i=1}^{n} \mathcal{N}(|\mathbf{x}_i - \mathbf{x}^{(k)}|, \sigma), \\ R^{(k)} &= \textstyle\sum_{i=1}^{n} \mathcal{N}(|\mathbf{x}_i - \mathbf{x}^{(k)}|, \sigma) \cdot r_i. \end{aligned} \tag{3}$$

Here, $\mathbf{x}^{(k)}$ denotes the center of cell k and the kernel width σ is a parameter of the algorithm. The integrated weights $\Omega^{(k)}$ are used for normalisation of the weighted readings $R^{(k)}$, thus even coverage is not necessary. The integrated weights $\Omega^{(k)}$ also provide a confidence measure for the estimate at cell k. A high value means that the estimate is based on a large number of readings recorded close to the center of the respective grid cell. A low value, on the other hand, means that few readings nearby the cell center are available and that therefore a prediction has to be made using sensor readings taken at a rather large distance. We formalize this by introducing a confidence map $\alpha^{(k)}$ computed as

$$\alpha^{(k)} = 1 - e^{-(\Omega^{(k)})^2/\sigma_\Omega^2}. \tag{4}$$

Confidence values $\alpha^{(k)}$ are normalized to the interval $[0, 1)$. The confidence map $\alpha^{(k)}$ depends on the trajectory of the sensors, the size of grid cells c, the width of the kernel σ and the scaling parameter σ_Ω. This map is used to compute the mean concentration estimate $r^{(k)}$ as

$$r^{(k)} = \alpha^{(k)} \frac{R^{(k)}}{\Omega^{(k)}} + \{1 - \alpha^{(k)}\} r_0 \tag{5}$$

where r_0 represents an estimate of the mean concentration for cells for which we do not have sufficient information from nearby readings, indicated by a low value of $\alpha(k)$. We set r_0 to be the average over all sensor readings.

As it was mentioned above, we want to estimate the real variability of gas readings at each location instead of the *covariance of the mean* and therefore carry out a parallel estimation process. Similarly to the distribution mean map, Eq. (5), the variance map $v^{(k)}$ is computed from *variance contributions* integrated in a further temporary map $V^{(k)}$

$$\begin{aligned} V^{(k)} &= \textstyle\sum_{i=1}^{n} \mathcal{N}(|\mathbf{x}_i - \mathbf{x}^{(k)}|, \sigma)(r_i - r^{(k(i))})^2, \\ v^{(k)} &= \alpha^{(k)} \tfrac{V^{(k)}}{\Omega^{(k)}} + \{1 - \alpha^{(k)}\} v_0 \end{aligned} \tag{6}$$

where $k(i)$ is the cell closest to the measurement point $\mathbf{x}_i$, and thus $r^{(k(i))}$ is the mean prediction of the model for cell k. The estimate v_0 of the distribution variance in regions far from measurement points is computed as the average over all variance contributions.

Fig. 2 shows an example of a weight map $\Omega^{(k)}$ (top row) and the corresponding confidence map $\alpha^{(k)}$ (bottom row). For narrow kernels, and large values of σ_Ω (left column) one can see the trajectory of the gas sensor carried by the robot, indicating that the predictions from extrapolation will only be considered trustworthy close to actual measurement locations. For wider kernels or smaller values of σ_Ω (right column) the area for which predictions based on extrapolation are made is larger.

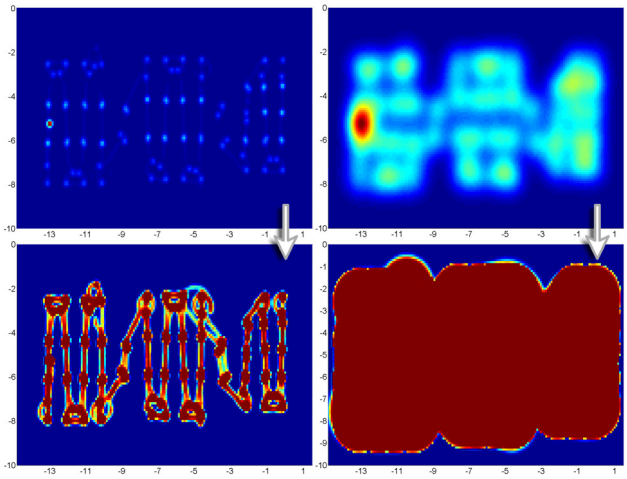

Fig. 2. Weight map $\Omega^{(k)}$ (top row) and the corresponding confidence map $\alpha^{(k)}$ (bottom row) obtained using the parameters $\sigma = 0.10\,m$, $\sigma_\Omega = 5.0 \cdot \mathcal{N}(0, 0.10) \approx 20.0$ (left column) and $\sigma = 0.50\,m$, $\sigma_\Omega = 1.0 \cdot \mathcal{N}(0, 0.50) \approx 0.8$ (right column) on a grid with cell size $c = 0.05\,m$.

The complexity of computing the maps in Eq. (3) is generally $\mathcal{O}[n \cdot (\frac{D}{c})^2]$ where n is the number of training samples, D is the dimension of the environment and c is the cell size. In practice, we do not have to evaluate the Gaussian weighting function $\mathcal{N}$ for all cells and limit the region for which the weights are computed to a circle of radius $4\,\sigma$ around the measurement location. Therefore the effective computational complexity is $\mathcal{O}[n \cdot (\frac{\sigma}{c})^2]$. The complexity of computing the maps $\alpha^{(k)}$, $r^{(k)}$, and $v^{(k)}$ in Eqs. 4, 5, and 6 is $\mathcal{O}[(\frac{D}{c})^2]$ and computing $V^{(k)}$ in 6 requires one pass through the data ($\mathcal{O}[n]$), thus the overall complexity is $\mathcal{O}[n \cdot (\frac{\sigma}{c})^2]$.

IV. EXPERIMENTS

We carried out gas distribution mapping experiments in which a robot followed a predefined sweeping trajectory covering the area of interest. Measurements were recorded at a frequency of 1 Hz. Along its path, the robot was stopped at a pre-defined set of grid points to carry out measurements on the spot for 10 s (outdoors) and 30 s (indoors). In this way we can investigate how the proposed algorithm deals with the case of a moving sensor (by using only the measurements taken between the stops) or a situation similar to that of a stationary sensor network (using only the measurements from the stopped robot). The spacing between the grid points was set to values between 0.5 m to 2.0 m depending on the available space. The sweeping motion was performed twice in opposite directions and the robot was driven at a maximum speed of 5 cm/s in between the stops. The gas source was a small cup filled with ethanol.

Apart from a SICK laser range scanner used for pose correction, the robot was equipped with an electronic nose and an anemometer (not used to compute gas distribution maps). The electronic nose comprises six Figaro gas sensors (2×TGS 2600, TGS 2602, TGS 2611, TGS 2620, TGS 4161)

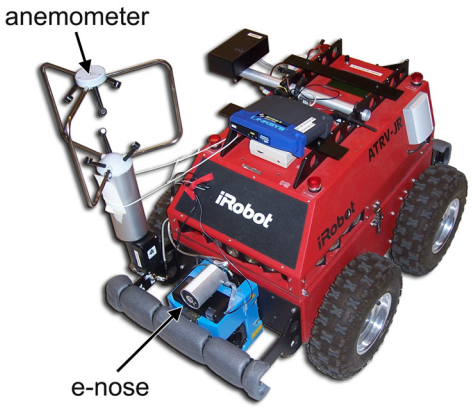

Fig. 3. The pollution monitoring robot "Rasmus" equipped with a SICK laser scanner for pose correction, an electronic nose and an anemometer.

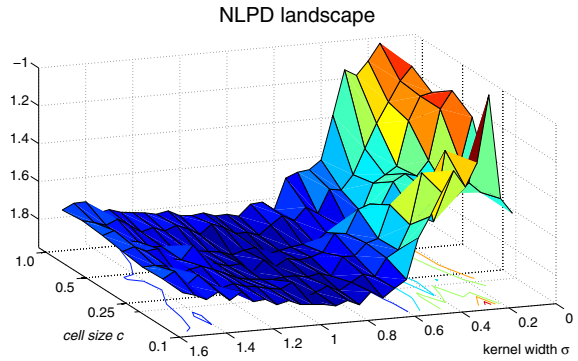

Fig. 4. NLPD landscape for the data shown in Fig. 1 measured in the *3-room* environment depending on the cell size c and the kernel width σ.

enclosed in an aluminum tube. This tube is horizontally mounted at the front side of the robot at a height of 34 cm, see Fig. 3. The electronic nose is actively ventilated through a fan that creates a constant airflow towards the gas sensors. This lowers the effect of external airflow and the movement of the robot on the sensor response and guarantees a continuous exchange of gas in situations with very low external airflow. In this work, we address the problem of modeling the distribution from a single gas source. With respect to this task, the response of the different sensors in the electronic nose is highly redundant and we therefore compute the gas distribution maps from the response of a single sensor (TGS 2620).

Three environments with different properties have been selected for the gas distribution mapping experiments. Experiments were carried out in an enclosed indoor area that consists of three rooms separated by slightly protruding walls in between them (*3-rooms*). Here, the area covered by the path of the robot was $\approx 14 \times 6\,m^2$. There is very little exchange of air with the "outer world" in this environment. The gas source was placed in the middle of the central room and all three rooms were monitored. The second location was a part of a *corridor* with open ends and a high ceiling. The area covered by the trajectory of the robot is $\approx 14 \times 2\,m^2$. The gas source was placed on the floor in the middle of the investigated corridor segment. Finally, an *outdoor* scenario was considered. Here, the experiments were carried out in an $8 \times 8\,m^2$ region that is part of a much bigger open area. The gas source was placed in the middle of this area.

V. PARAMETER SELECTION

The Kernel DM+V algorithm depends on three parameters: the kernel width σ, which governs the amount of extrapolation on individual readings (and the complexity of the model); the cell size c that determines the resolution at which different predictions can be made; and the scaling parameter σ_Ω, which defines a soft threshold between values of $\Omega^{(k)}$ that are considered "high" (where $\alpha^{(k)}$ is close to 1) and "low" ($\alpha^{(k)}$ is close to 0). Smaller values of σ_Ω entail a lower threshold on $\Omega^{(k)}$, i.e. an increasing tendency to trust the distribution estimate obtained from extrapolation on local measurements.

In order to evaluate how well the map estimate captures the true properties of the gas distribution, we divide the sample set $\mathcal{D}$ into disjoint sets $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$. Optimal values of the model parameters are determined by cross-validation on $\mathcal{D}_{train}$. An obvious way to measure how well unseen measurements are predicted by the distribution model is to compute the average prediction error. Due to the large fluctuations of the instantaneous gas distribution, however, this measure of model quality is not particularly suitable for gas distribution modelling. A gas distribution model should represent the time-averaged concentration *and* the expected fluctuations. These properties are both captured by the average negative log predictive density (NLPD), which is a standard criterion to evaluate distribution models. Under the assumption of a Gaussian posterior $p(r_i|\mathbf{x}_i)$, the NLPD of unseen measurements $\mathcal{D} = \{r_1, ..., r_n\}$ acquired at locations $\{x_1, ..., x_n\}$ is computed as

$$NLPD = -\frac{1}{n}\sum_{i\in\mathcal{D}} \log\{p(r_i|\mathbf{x}_i)\} = \frac{1}{2n}\sum_{i\in\mathcal{D}} \left\{\log \hat{v}(x_i) + \frac{(r_i - \hat{r}(x_i))^2}{\hat{v}(x_i)}\right\} + \log(2\pi). \quad (7)$$

The estimations $\hat{v}(x_i)$ and $\hat{r}(x_i)$ are obtained from the corresponding cells in the maps in Eqs. 5 and 6 as $\hat{v}(x_i) = v^{(k(i))}$, $\hat{r}(x_i) = r^{(k(i))}$. Since the goal is to maximize the likelihood of the data points, we search for parameters that minimise the NLPD in Eq. (7).

Regarding the selection of the model parameters, we first note that the exact value of σ_Ω is not critical as long as it is of the right scale. We have observed that for all practical aspects it can be tied to the selected kernel width σ using $\sigma_\Omega = \mathcal{N}(0, \sigma)$. Fig. 4 shows a typical NLPD landscape, which was obtained from the data shown in Fig. 1 by cross-validation using 5% of the data for training and the remaining 95% for testing. One can first observe that the dependency on the grid cell size is rather weak, with a tendency to favour smaller cell sizes as expected. We can further see that the minimum with respect to the kernel size σ is shallow to the end of large kernel sizes and steep towards small kernel sizes

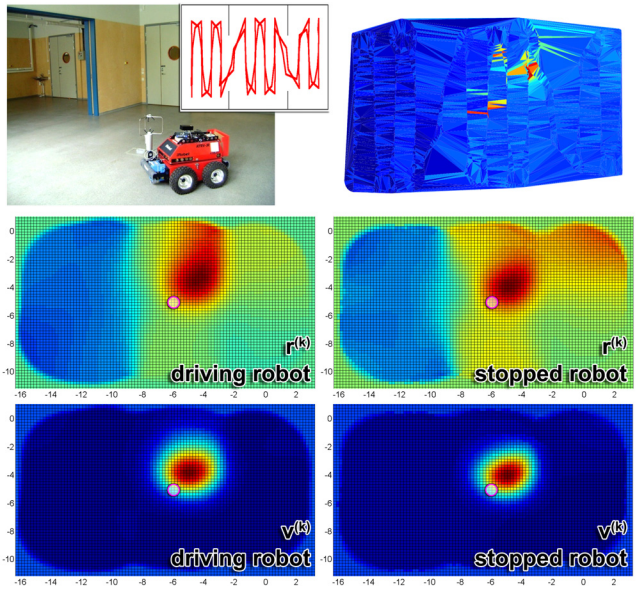

Fig. 5. Top row, left: trajectory of the robot and a picture of the experimental environment *3-room*. Top row, right: gas distribution map created by trilinear interpolation (using the Matlab function *trisurf*). The second row shows the mean distribution map $r^{(k)}$ and the third row the variance distribution map $v^{(k)}$, obtained from measurements where the robot was driving (left) and standing, respectively (right). The location of the gas source is indicated by a circle. Parameters were selected as described in the text by cross-validation on 5% of the data. The optimal values of the kernel width σ were found to be 1.05 m (driving robot) and 1.0 m (stopped robot).

where overfitting occurs. Remarkably, for a fixed cell size, the relative course of the NLPD depending on the kernel size σ does not change substantially. In the example in Fig. 4, the minimum occurs between 0.75 m and 0.95 m and the optimal value was found to be $\sigma = 0.90$ m. This enables us to reduce the parameter selection problem from three (σ, c, σ_Ω) to one dimension (σ). In order to make the learning process even faster, we search for the optimal value of σ using a relatively large cell size and recompute the map with the optimal value of σ found at a larger resolution. With a cell size of 0.8 m, for example, the computation of a map from 5% of the data (approx. 300 data points in the example shown in Fig. 4) takes less than 0.1 s in an un-optimized Matlab implementation.

VI. RESULTS

Using the parameter selection method described in Sec. V, we obtain the maps shown in Fig. 5 for the example experiment that was used for illustration so far in this paper (see Fig. 1, Fig. 2 and Fig. 4). The left column shows the mean and variance map computed from measurements recorded while the robot was driving. The maps in the right column were created from the measurements collected while the robot was stopped. The similarity between the maps is a remarkable indication of the consistency of the maps created with the Kernel DM+V algorithm considering that the underlying data sets are disjunct. The observed gas distribution was found to be restricted to two of the three rooms although these rooms where connected by large alleyways, a finding

TABLE I
NLPD COMPARISON OF KERNEL DM+V AND THE GPM METHOD.

Dataset	NLPD, GPM	NPLD, Kernel DM+V
3-rooms	-1.54	-1.44
corridor	-1.60	-1.81
outdoor	-1.80	-1.75

that is supported by the wind measurements, which indicate the existence of an effective wall created by stable airflows in the room. This experiment provides an illustrative example of the use of mobile robots in a rescue mission where they could report that the leftmost room is secure for human rescue staff while the concentration levels in the other rooms are too high.

For one experiment in each of the three environments described in Sec. IV we compared the Kernel DM+V algorithm to the Gaussian process mixture approach [16] ("GPM"). We do this by dividing the data into a training set, comprising all the measurements collected during the first full sweep. This data set is used to learn the gas distribution model, which is then evaluated on a separate test set that comprises the remaining measurements collected during the second sweep in opposite direction. The GPM approach has a mechanism to select training data and we used exactly the same data to learn the parameter σ for the Kernel DM+V algorithm.

Although the training data were selected for the GPM method and GPM learns three parameters compared to one parameter in the case of Kernel DM+V, the results shown in Tab. (I) are comparable. This gives Kernel DM+V an edge because of its ability to scale better to larger training samples (having complexity $\mathcal{O}[n \cdot (\frac{\sigma}{c})^2]$ compared to $\mathcal{O}[n^3]$ in the case of GPM) and the fact that the learning procedure is simpler since it is not necessary to simultaneously learn model components and a gating function that switches between these components.

VII. SUMMARY

Gas distribution modelling constitutes an ideal application area for mobile robots. As intelligent mobile gas sensors, robots can provide a higher and adaptive resolution than a stationary sensor network. They further offer accurate localization and the computational resources to create the distribution model on-line. Thus also the possibility to decide based on the current model which locations to observe next.

In this paper we propose the Kernel DM+V algorithm to learn a statistical 2-d distribution model from a sequence of localized gas sensor measurements. We discuss the hyper-parameters and present a method to learn them from the data. In the presented experiments the sensors were carried by a mobile robot, which was stopped at pre-defined measurement points to demonstrate that the algorithm can be applied to a stationary sensor network without modifications. The Kernel DM+V algorithm is non-parametric, i.e. it makes no assumptions about a particular functional form of the distribution model. In contrast to most previous gas distribution modelling approaches, Kernel DM+V estimates the observed variance in addition to the distribution mean. Estimating the predictive variance entails a significant improvement for gas distribution modelling since it among other benefits allows

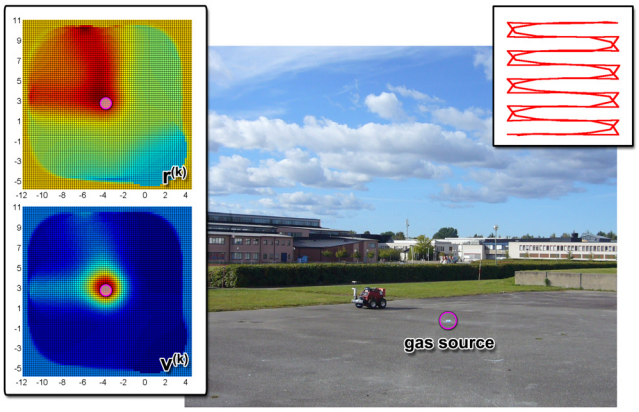

Fig. 6. Trajectory of the robot (right) and the maps $r^{(k)}$ and $v^{(k)}$ (right) plotted partly over a picture of the experimental environment *outdoor*. The location of the gas source is indicated by a circle. Parameters were selected as described in the text by cross-validation on 5% of the data. The optimal value of the kernel width σ was found to be 1.20 m.

to evaluate the model quality in terms of the data likelihood and provides the means to learn meta parameters.

The results, based on real world data collected with a mobile robot, demonstrate the consistency of maps obtained from stationary and mobile sensors. We also present a quantitative comparison with a recently presented alternative approach [16] using Gaussian process mixture models ("GPM"), which is, to the best of our knowledge, currently the only other statistical gas distribution modelling approach that estimates a predictive variance. The comparable performance in terms of the data likelihood demonstrates that the two approaches produce maps, which predict unseen samples similarly well, giving Kernel DM+V an edge due to its simpler learning procedure and better scaling properties in the case of larger training samples.

VIII. FUTURE WORK

Confirming results of a previous study [9], we made the observation in the experiments presented in this paper that the maximum in the variance map often provides a more accurate and better localized prediction of the location of the gas source than the maximum in the distribution mean map, see Fig. 6. In order to statistically support this observation, we need to carry out and evaluate further experiments.

For a fixed location (i.e. a particular grid cell), the Kernel DM+V algorithm assumes a normal distribution. This assumption does not generally hold, particularly for gas distributions, which often exhibit rather a multimodal distribution at one measurement point. Future work will therefore investigate the possibility to relax the assumption of a normal distribution towards other parametric or non-parametric distributions.

An important assumption of the proposed gas distribution modelling method (and of all the statistical gas distribution modelling methods mentioned in Sec. II) is that the model is learned from measurements, which are generated by a time-constant random process. Possible extensions of the Kernel DM+V algorithm that allow to cope with dynamic situations include adding a "lazy update" mechanism. The quality of the current gas distribution map could be continuously evaluated on new sensor readings and updated only if the data likelihood drops significantly. Such an update could be compared to a map that is computed from fewer, more recent samples only. By eventually selecting this map, the representation could follow slow changes of the gas distribution. Such a mechanism would also help providing an answer to the question how many measurements are needed in a given environment to provide a truthful representation.

IX. ACKNOWLEDGMENTS

This work has partly been supported by the EC under contract number FP6-045299-Dustbot: Networked and Cooperating Robots for Urban Hygiene, and FP7-224318-DIADEM: Distributed Information Acquisition and Decision-Making for Environmental Management.

REFERENCES

[1] DustBot Consortium. DustBot - Networked and Cooperating Robots for Urban Hygiene. http://www.dustbot.org.

[2] F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte Carlo Localization for Mobile Robots. In *Proc. ICRA*, pages 1322–1328, 1999.

[3] P. Pyk et al. An Artificial Moth: Chemical Source Localization Using a Robot Based Neuronal Model of Moth Optomotor Anemotactic Search. *Auton Robot*, 20:197–213, 2006.

[4] A.T. Hayes, A. Martinoli, and R.M. Goodman. Distributed Odor Source Localization. *IEEE Sensors Journal, Special Issue on Electronic Nose Technologies*, 2(3):260–273, 2002. June.

[5] J. O. Hinze. *Turbulence*. McGraw-Hill, New York, 1975.

[6] H. Ishida, T. Nakamoto, and T. Moriizumi. Remote Sensing of Gas/Odor Source Location and Concentration Distribution Using Mobile System. *Sensors and Actuators B*, 49:52–57, 1998.

[7] A. J. Lilienthal, S. Asadi, and M. Reggente. Estimating Predictive Variance for Statistical Gas Distribution Modelling. In *AIP: Olfaction and Electronic Nose - Proc. ISOEN*, volume 1137, pages 65–68, 2009.

[8] A. J. Lilienthal and T. Duckett. Building Gas Concentration Gridmaps with a Mobile Robot. *Robotics and Autonomous Systems*, 48(1):3–16, August 2004.

[9] A. J. Lilienthal, T. Duckett, H. Ishida, and F. Werner. Indicators of Gas Source Proximity using Metal Oxide Sensors in a Turbulent Environment. In *Proc. IEEE / RAS-EMBS Biorob*, Pisa, Italy, 2006.

[10] A. J. Lilienthal, A. Loutfi, J. L. Blanco, C. Galindo, and J. Gonzalez. A Rao-Blackwellisation Approach to GDM-SLAM Integrating SLAM and Gas Distribution Mapping. In *Proc. ECMR*, pages 126–131, 2007.

[11] A. Loutfi, S. Coradeschi, A. J. Lilienthal, and J. Gonzalez. Gas Distribution Mapping of Multiple Odour Sources using a Mobile Robot. *Robotica*, June 4 2008. Online.

[12] H. R. Olesen, P. Løfstrøm, R. Berkowicz, and M. Ketzel. Regulatory Odour Model Development: Survey of Modelling Tools and Datasets with Focus on Building Effects. Technical Report 541, NERI, Denmark, 2005.

[13] E. Parzen. On the estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.

[14] A.H. Purnamadjaja and R.A. Russell. Congregation Behaviour in a Robot Swarm Using Pheromone Communication. In *Proc. of the Australian Conf. on Robotics and Automation*, 2005.

[15] B. Shraiman and E. Siggia. Scalar Turbulence. *Nature*, 405:639–646, 8 June 2000. Review Article.

[16] C. Stachniss, C. Plagemann, and Achim J. Lilienthal. Gas Distribution Modeling using Sparse Gaussian Process Mixtures. *Autonomous Robots*, 26(2-3):187–202, April 2009.

[17] S. Thykier-Nielsen, S. Deme, and T. Mikkelsen. Description of the Atmospheric Dispersion Module RIMPUFF. Technical Report RODOS(WG2)-TN(98)-02, Risø National Laboratory, Roskilde, Denmark, 1999.

[18] M. R. Wandel, A. J. Lilienthal, T. Duckett, U. Weimar, and A. Zell. Gas Distribution in Unventilated Indoor Environments Inspected by a Mobile Robot. In *Proc. ICAR 2003*, pages 507–512, 2003.