Ref. Ares(2017)3269636 - 29/06/2017

Mobile Robots with Novel Environmental Sensors for Inspection of Disaster Sites with Low Visibility

Project start: January 1, 2015

Duration: 3.5 years

Deliverable 2.7

Gas Distribution and Airflow Mapping

Due date: month 30 (June 2017)

Lead beneficiary: ORU

Dissemination Level: Public

Main Authors:

Victor Hernandez Bennetts (ORU)

Erik Schaffernicht (ORU)

Han Fan (ORU)

Version History:

0.1: initial draft, vh, June, 2017

0.2: first revision, es, June 2017

Contents

Contents 3
A Introduction and purpose of this document 4
B Airflow Mapping 4
C Towards Online Gas Distribution Mapping for Emergency Scenarios 8
C.1 Airflow aware gas distribution mapping 9
C.2 Gas Identity aware gas distribution mapping 11
D Summary and Outlook 13
References 15
Annex 16

A Introduction and purpose of this document

The SmokeBot project will develop a new tele-operated robotic platform for supporting fire brigades for search and rescue missions with a special focus on operation environments with limited visibility. In disaster scenarios, where SmokeBot operates, the presence of hazardous chemicals is a risk factor to consider and thus, it is critical to have sensing modalities that allow to detect, identify and quantify gaseous substances present in the environment. These gas sensing capabilities are provided by an array of novel bandwidth sensors (T2.1) and an unsupervised gas identification algorithm (T2.3) that not only allows to identify previously seen target gases, but it also allows to detect the presence of unknown gas interferents. The information related to these gas detection events is presented in the form of gas distribution maps that show the spatial distribution and the identity of the detected gas concentrations.

This document describes the proposed gas distribution and airflow mapping algorithms and the achieved results in the period from M1-M30 and as well provides an overview of future developments on this task. The rest of this report is structure as follows: Section B presents the technical description and evaluation of the proposed airflow mapping algorithm. Section C presents the developments on gas distribution mapping. Section D presents a summary and outlook towards future developments.

B Airflow Mapping

For mobile robots that operate in complex environments, such as the Smokebot platform, airflow information can be of great importance. For example, airflow information is used to enrich the produced gas distribution maps, it can be used in the detection of gas leaks and it can allow identifying safe areas in a disaster scenario, where airflow patterns suggest the lack of presence of dangerous chemicals.

Commonly, airflow information is used in the form of mean values for wind speed and direction. However, this approach does not capture the complex spatio-temporal variations of turbulent airflow observed in indoor environments. As previously reported in [1], considerable fluctuations in wind speed and direction are observed in different scenarios. Fig. 1 shows an example of the wind variability indoors. The plot was produced from a set of wind measurements acquired with a mobile robot exploring an indoor scenario. Notice in 1.a that directional variations are consistently observed, even at neighboring measurement locations. Temporal variations in the wind flow are observed in the rose plot (Fig. 1.b) and histogram plot (Fig. 1.c) of the wind direction and speed respectively, computed from measurements acquired at a single location. As can be seen in Fig 1.b, the mean direction (red line), provides a rather inaccurate estimation of the wind direction. Similarly, it can be observed in the wind speed histogram of Figure 1.c that a mean wind speed is not a good indicator of the wind flow conditions.

Alternative to instantaneous airflow measurements or mean values, airflow information can be conveyed in the form of an airflow model. AirFlow Modelling (AFM) is the task of creating a spatio temporal representation of variables such as wind speed and wind direction. AFM has been extensively studied in meteorology and atmospheric research [2]. However, atmospheric AFM algorithms are hardly applicable in the context of Smokebot, mainly due to the difference in spatial scales. Meteorologist have mostly focused on AFM in macro and meso scales, with different temporal resolutions. The lower range of these scales is approximately 2 Km [3].

In micro-scales, simple AFM approaches include the use of trace gases, as presented in [4], or the use of sophisticated Computational Fluid Dynamics (CFD) simulations [5]. However, these methods are not directly applicable to mobile robotics. For example, CFD-based techniques require precise a-priori knowledge on boundary conditions and moreover, these techniques require high computational resources, which may not be available on-board mobile platforms.

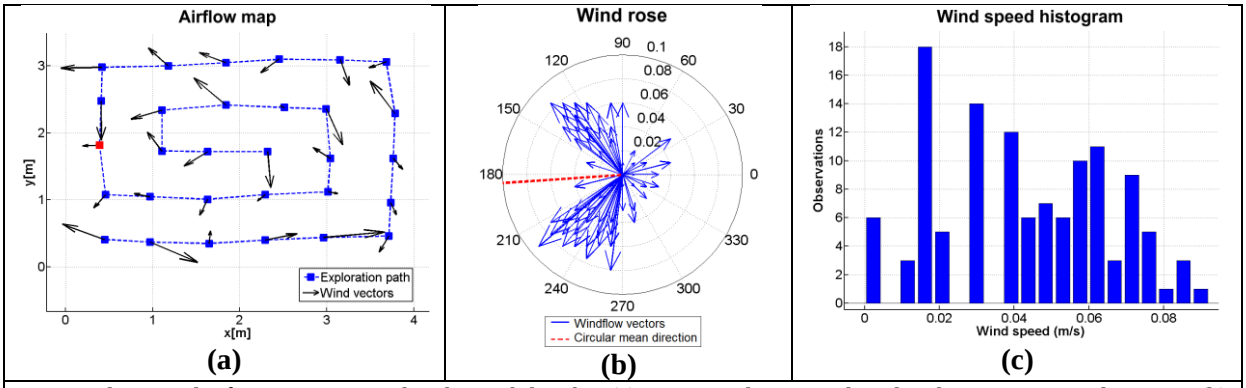

Fig 1: Indoor wind information acquired with a mobile robot (a) Mean wind vectors plotted at the measurement locations. (b) Wind rose plot of all the acquired wind measurements. (c) Histogram of the acquired wind speed measurements.

Data-driven AFM techniques are a viable option for mobile robotics platforms. To perform AFM, mobile robots should be able to acquire wind information while exploring the environment. Thus, in Smokebot, we developed we developed a data-driven AFM algorithm that uses localized wind speed and direction measurements, acquired at a sparse number of locations, as inputs. The algorithm's output is a 2-D model (i.e. a map) in which a discrete representation of the air flow distribution (i.e. a histogram) is modelled for each query location. The AFM algorithm developed in Smokebot is presented in the article "Probabilistic Air Flow Modelling Using Turbulent and Laminar Characteristics for Ground and Aerial Robots", published in RA-L as a journal publication and presented in the 2017 IEEE International Conference on Robotics and Automation (ICRA). The accepted publication is annexed to this deliverable.

The AFM process is divided into two stages, namely the estimation of the joint Probability Density Function (PDF) at the measurement locations, and an extrapolation process to infer joint PDFs at non-visited locations. In the first stage of the AFM process, joint PDFs of the wind speed and wind direction are iteratively computed, at distinct measurement locations, as data is being acquired. The joint PDFs are modelled as a discrete, two dimensional space that has a total of N states. Each state in the joint PDF is a tuple of wind direction and speed. These local PDFs incorporate knowledge about the sensors, and wind variability in the form of a sensor model and a transition matrix respectively. The sensor model is defined according to the sensor parameters specified by the manufacturer while the transition matrix is a frequentist model that assigns probabilities based on the observed transitions between states of the PDF. This means that higher posteriors are assigned to states that do not exhibit intermittent transitions but rather stable/frequent conditions over time. This allows to discriminate between dominant wind states and turbulence. The transition matrix does not make a priori assumptions and it is learned as data is recorded.

Fig. 2 shows an example run to compute a local PDF from 1500 randomly generated measurements. The measurements simulate high turbulence conditions and a stable, dominant air flow direction centred at (250°, 0.15 m/s) and with standard deviations of 25° in wind direction and 0.08 m/s in

wind speed. A total of 70% of the data points correspond to randomly generated noise. Notice that, compared to a simple histogram, the proposed method clearly identifies the dominant air flow direction by assigning higher probabilities to those states in the PDF that show lower variability.

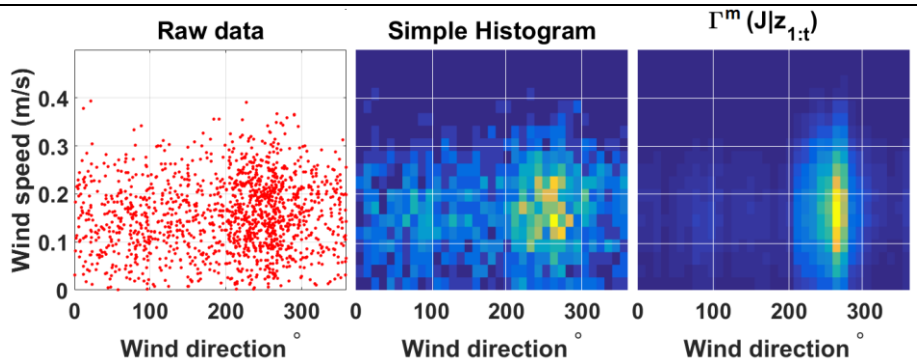

Fig. 2: A joint PDF computed at a measurement location using the method developed in Smokebot. Notice the comparison between a simple histogram and the proposed method (denoted by Γ^m), which can diferentiate between turbulence and dominating wind streams.

In the second stage of the AFM approach, the PDFs at non visited locations are predicted using the PDFs computed at the measurement location. In the proposed algorithm, we introduced a novel extrapolation approach named the LT algorithm, which is specifically tailored for the task of AFM. The **LT** algorithm models airflow as a weighted combination of laminar flow components (**L**) and turbulent flow components (**T**).

The laminar component is modeled as a laminar, unimodal distribution which is computed as the mean values of wind direction and speed from a set of artificial samples drawn from the local PDFs at measurement locations. The mean values are computed using the Nadaraya-Watson estimator with a Radial Basis Function (RBF) Kernel.

The **LT** algorithm assumes that turbulent flow is given by multimodal distributions. However, mean values (as computed in the **L** component) convey little to no information regarding multimodal (e.g. turbulent) wind flow distributions and moreover, the circular mean is undetermined if wind vectors are opposed to each other. Thus, a turbulent component (**T**) is modeled as a multimodal distribution, approximated by a linear combination of the local PDFs weighted by a RBF kernel. To weight between the **L** and the **T** components in the extrapolation process, a turbulence indicator is used. Since turbulent flow is assumed to have a multimodal distribution the circular variance (CV) of the **T** component is used as a basis of a turbulence indicator. Technical details about the implementation of the LT algorithm are explained in the annex.

The proposed AFM algorithm was evaluated using a Unmanned Aerial Vehicle (UAV) exploring outdoors. While the robot used for validation purposes drastically differs from the Smokebot platform, it operates in an environment often dominated turbulence, which is similar to the application scenarios of Smokebot. Figs 3.a and 3.b exemplifies the airflow conditions measured with the UAV. Both figures were produced with data from a single exploration run. In Fig. 3.a, the exploration path (dashed red line) of the robot is shown. Measurements were collected with the robot hovering at pre-defined positions (red markers). The blue arrows denote the instantaneous wind measurements where the arrow's direction denotes wind direction and the arrow's length denote wind speed. Figure 3.b

shows a circular histogram, produced from all the collected measurements. Notice that neither of the plots suggest a laminar flow but instead, highly turbulent conditions can be observed.

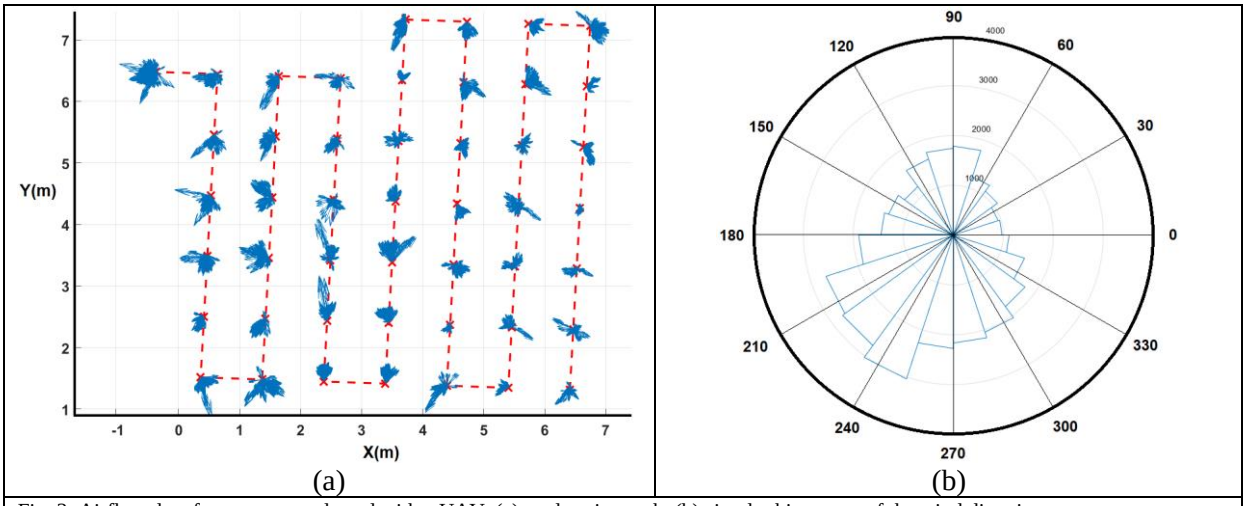

Fig. 3: Airflow data from a run conducted with a UAV. (a) exploration path. (b) circular histogram of the wind direction.

Using 11 exploration runs, where the UAV hovered over several measurement locations, a Leave One Out Cross Validation (LOOCV) process was conducted. A local PDF was computed at the left out data and the predictive capabilities of the generated air flow maps (generated with data from the remaining measurement location) were evaluated using the Kullback Leibler Divergence (KLD). The sensitivity to parameter selection was evaluated by testing different combinations of the two functional parameters of the **LT** extrapolation algorithm. Considering all parameter combinations, the 11 runs and the number of waypoints per run, around 15000 models were trained and tested. The obtained results were compared against laminar only models (**L**) and turbulence only models. As can be seen in Fig. 4, the LT algorithm outperforms both the L-only models and the T-only models proving in this way that a model that considers both laminar and turbulence components has better prediction capabilities, regardless of the selection of functional parameters, when compared against simpler extrapolation methods.

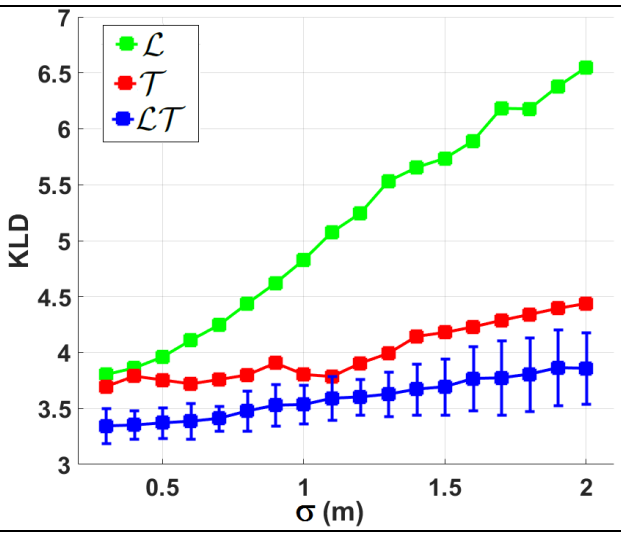

Fig 4: Performace plot of the LT algorithm (blue) for different functional parameters. Notice that the LT algorithm outperforms Laminar-

only (green) and Turbulence-only (red) extrapolation methods regardless of the selected parameters.

Fig. 5.a and Fig 5.b show air flow maps computed with data from two different UAV runs. For simplicity, the maps show the expected wind speed and direction at each query location. The expected air flow map for the experiment performed on 2011-07-07 (Fig. 5.a) shows stable flow conditions and a well-defined wind direction while in the case of the experiments conducted on 2011-07-12 (Fig. 5.b), low wind velocities, eddies and vortices are predicted at different locations in the explored environment.

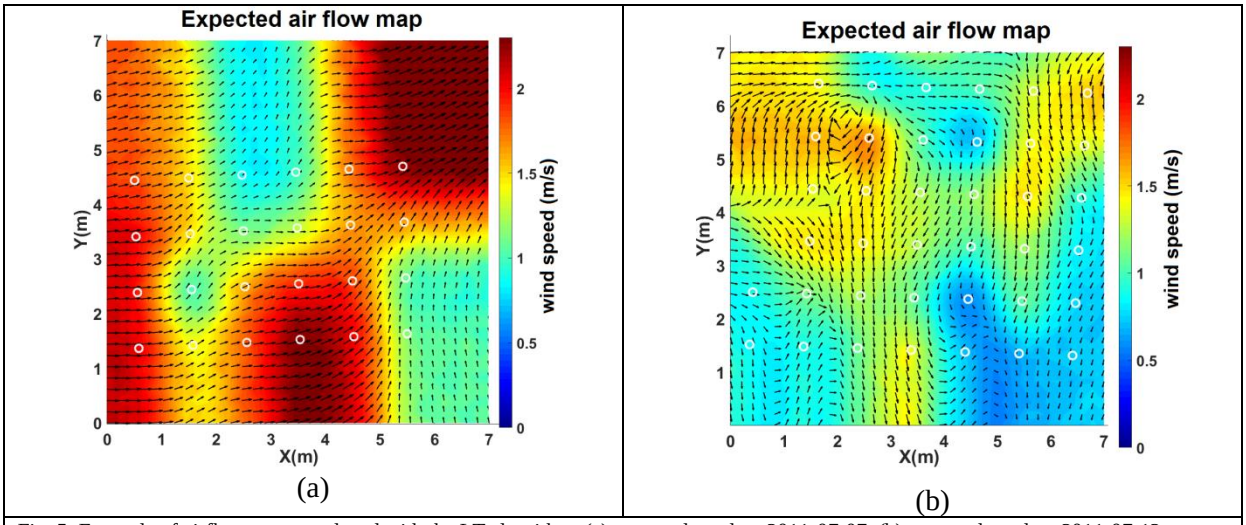

Fig. 5: Example of airflow maps produced with the LT algorithm. (a) run conducted on 2011-07-07. (b) run conducted on 2011-07-12

C Towards Online Gas Distribution Mapping for Emergency Scenarios

As previously stated, a gas distribution map can convey useful information to firefighters brigades by providing a spatial representation that shows how a given set of airborne chemicals of interest distribute over the emergency area. In the particular case of Smokebot, a gas distribution map can be seen as a direct input to related tasks and work packages in the project, such as T2.5.- Gas source localization and WP4. Situation Analysis. The computation of such gas distribution maps is a high level process that requires the integration of different relevant information such as:

- The identity of the detected gases (Tasks T2.1, T2.2 and T2.3)
- The geometrical information of the environment (WP1)
- Relevant environmental information such as airflow models (T2.4)

In the case of an emergency scenario, these gas distribution maps are computed from information acquired mostly as the robotic platform traverses the area of interest and while it is possible that a-priori information might be available (e.g. emergency map that shows the physical layout of the area under surveillance), this information is likely to change in an scenario where rubble and collapse walls might be present.

Thus, in Smokebot, we propose that gas distribution mapping is carried out in a highly dynamic way, with new information (e.g. obstacle layouts, new airborne chemicals identified) added as the robot explores the environment, with previously computed gas distribution map segments updated according to new information. Thus, the overall result is a map that, in practice, is a snapshot of the gas distributions present while the robot explored the disaster scenario. This result is in stark contrast, when compared to other applications, such as methane emission surveillance [6] or indoor pollution monitoring [7], where a long term model of the gas distribution is required to observe the evolution of the gas emissions over high spatial granularities (day, weeks, months).

In the remainder of this section we explore how these different sources of information can be included in the computation of gas distribution maps. First, we explore the effect of the integration of airflow maps. Then, we explore the first steps towards a gas distribution mapping algorithm that dynamically integrates the identities of airborne contaminants as gas concentration measurements are acquired. In the final subsection, we explore how geometrical information of the environment can be integrated in the computation of the gas distribution maps

C.1 Airflow aware gas distribution mapping

In this section, we introduce the algorithm implemented for the Smoke platform to integrate airflow information in the computation of gas distribution maps. Airflow dependent gas distribution maps was extensively studied by Reggente and co-authors in their work presented in [8, 9], where they proposed the Kernel DM+V/W. Reggente et al. integrate wind information into the computation of the gas distribution maps by using either instantaneous measurements or mean values for the wind speed and direction. As we previously explained, these approaches do not capture the complexities of the airflow conditions. In Smokebot, we extend the work of Reggente et al. by using airflow local PDFs instead of single airflow values. We proposed a gas distribution mapping algorithm called the Kernel DM+V/Γ, which is based on the Nadaraya-Watson estimator and, similar to Reggente's algorithm, it shrinks/stretches the semi-major and semi-minor axis of the RBF kernel with a bandwidth σ, under the constraint that the area of the ellipses remain constant. Thus, the semi major axes (a) and the semi minor axes (b) are modified as follows:

$$a = \sigma + \gamma|\vec{v}| \quad (1)$$

$$b = \frac{\sigma}{1 + \frac{\gamma|\vec{v}|}{\sigma}} \quad (2)$$

In addition, the kernel's covariance matrix is rotated so that the semimajor axis is aligned with the wind direction. Once the kernel has been modified according to the wind vector, the extrapolation of the gas concentration measurements is carried out as explained in [10]. Following the same principle proposed by Reggente, the Kernel DM+V/ Γ considers each state (wind speed/direction) in the discrete airflow PDF Γ and computes an ellipsoid $\mathcal{N}_j$ ($|x_i - x^k|$,σ), stretched and rotated according to the wind vector (speed/direction) for each state *j* in Γ. Thus, in the Kernel DM+V/ Γ algorithm, gas concentrations are extrapolated using a multimodal kernel Ω (x_i ,x^k)which is a summation of elliptical kernels (rotated and stretched according to states j in Γ), weighted by the states' joint probabilities

(p(j)) according to the PDF Γ. Given n gas measurements c_i, acquired at locations x_i, the extrapolated gas concentration measurement c^k at location x^k is given as follows:

$$\Omega(x_i, x^k) = \sum_{j=1}^{J} p(j)\, \mathrm{N}_j(|\boldsymbol{x_i} - \boldsymbol{x^k}|, \boldsymbol{\sigma}) \tag{3}$$

$$c^k = \frac{\sum_{i=1}^{n} \Omega(x_i, x^k)\boldsymbol{c_i}}{\sum_{i=1}^{n} \Omega(x_i, x^k)} \tag{4}$$

Fig. 6 show examples of multimodal kernels Ω (right) for different airflow PDFs (left). In the figures, the white Notice how the kernel Ω influences neighbouring locations according to the probabilities given by the PDF. It is also worth notice that multimodal airflow PDFs influences neighbouring locations in the directions of the dominant airflow streams. Notice as well that PDFs that predict high confidences for specific wind directions at high wind speeds translate into kernels Ω that influences several locations but only in the direction of the dominant airflow stream.

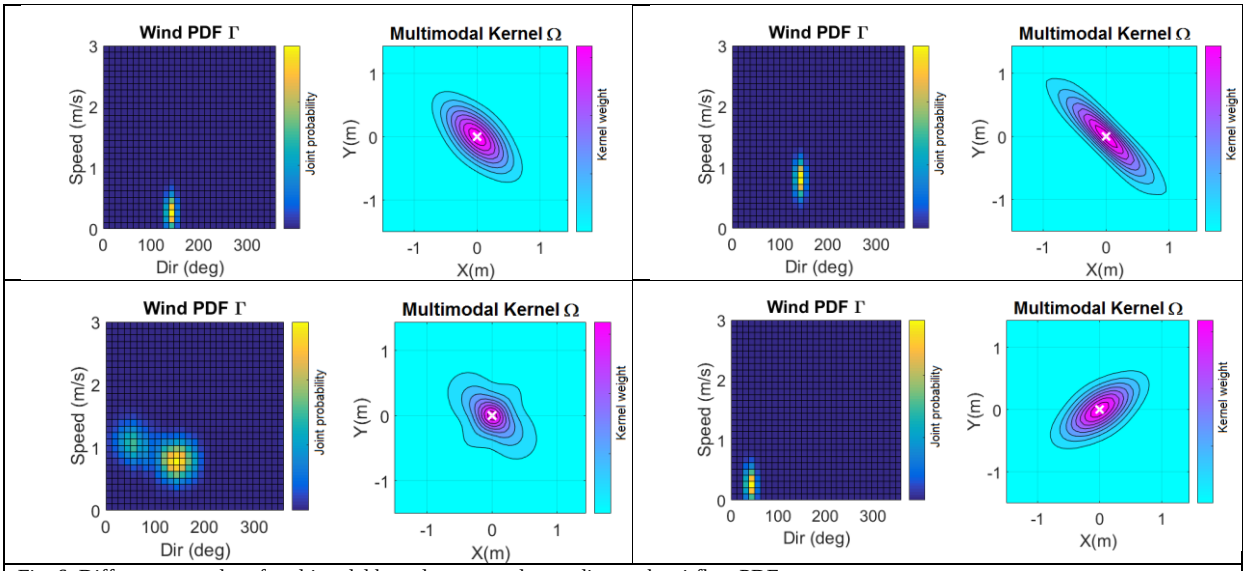

Fig. 6: Different examples of multimodal kernels generated according to the airflow PDF.

Figures 7.a to 7.c show examples of gas distribution maps generated without wind information (fig 7.a), with wind information using the Kernel DM+V/ W (Fig. 7.b), using mean wind vectors, and with the Kernel DM+V/ Γ (Fig. 7.c). In the absence of wind information gas concentrations are extrapolated in an isotropic way. This may lead to unrealistic results since airflow plays a key role in gas dispersion by transporting gas patches and thus, gas is expected to disperse in irregular shapes. In the case of wind-dependent gas distribution mapping, both algorithms depend on parameter γ, which controls the stretching and shrinking of the semi axis of the kernel (equations 1 and 2). Both figures were produced using a γ = 2.5. While both algorithms show the effect of the wind in the gas distribution, the kernel Kernel DM+V/ W is highly sensitive to the selection of parameter γ. Notice

that gas plumes are overly stretched in the direction of the wind (Fig. 7.b). On the other hand, the Kernel DM+V/ Γ does not show these narrow, high concentration plumes but instead, gas concentrations are extrapolated according to several dominating airflows instead a single one, which leads to a less narrow plume of irregular shape.

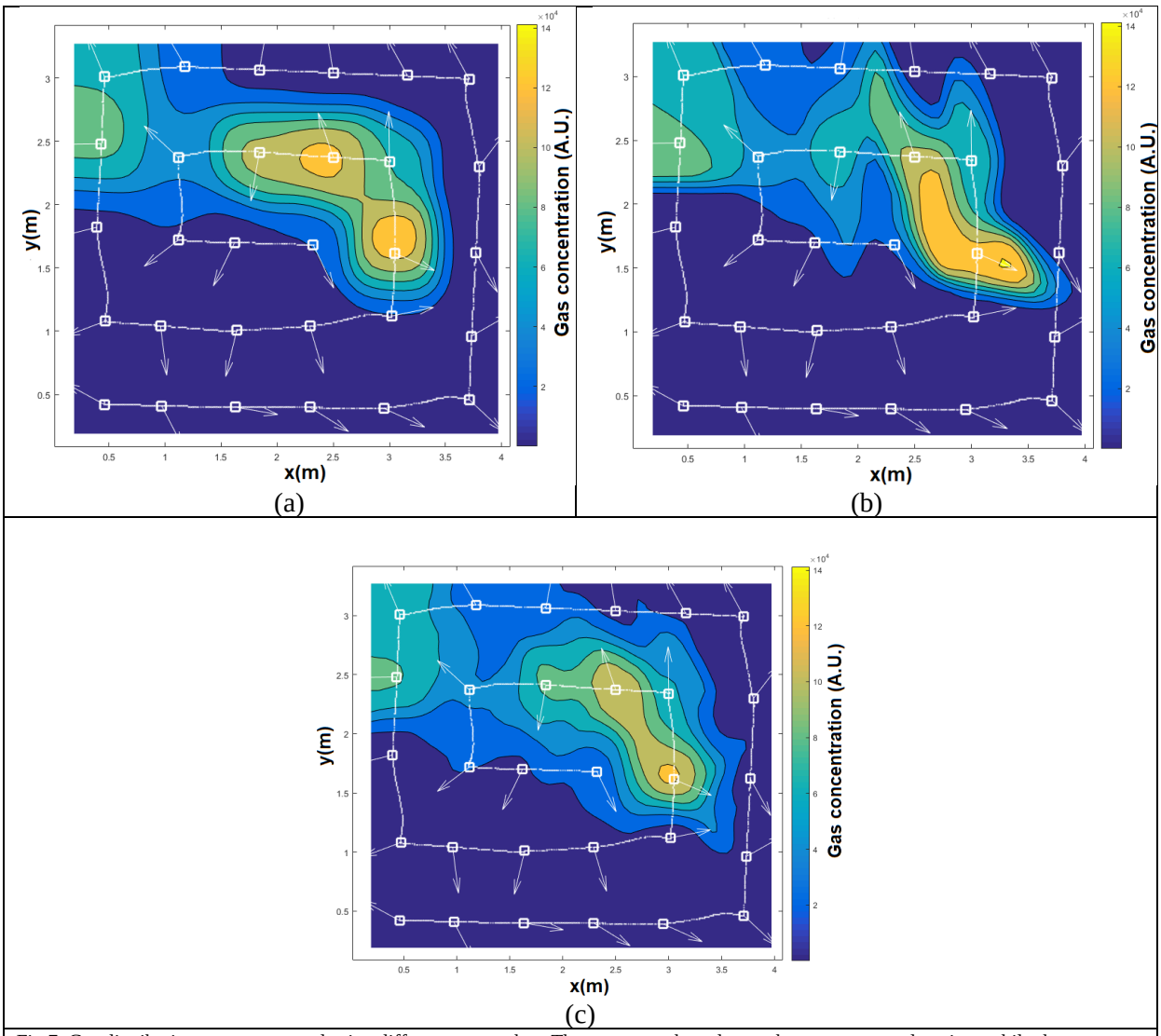

Fig 7: Gas distribution maps generated using different approaches. The square markers denote the measurement locations while the arrows denote the mean wind directions. (a) The Kernel DM+V [11] (no wind information). (b) The Kernel DM+V/W with mean wind information. (c) The Kernel DM+V/W, which uses airflow models instead of average wind vectors.

C.2 Gas Identity aware gas distribution mapping

An additional layer of information that has to be integrated in the task of gas distribution mapping is the identity of the detected gas concentrations. Multi compound gas distribution maps have been explored in the past in different publications [11, 12]. However, one of the limitations of these aforementioned algorithms is that they require a set of training data to perform gas discrimination and

therefore, measured concentrations of unknown gases can lead to erroneous results. This is due to the fact that neither the gas discrimination nor the gas distribution algorithms can process measurements from previously unknown gases. In Smokebot, we aim to develop a gas distribution mapping that can handle unknown interferents. As previously explained, unknown gases are expected to be present in complex scenarios.

We used the concept of an unsupervised algorithm, introduced in D2.6: Gas Detection and Identification, and we developed it further towards an algorithm that could be used in the task of online gas distribution mapping of multiple compounds. The latest iteration, which is called the Km-P algorithm not only allows to discriminate previously unknown gases, but also allows to compute class posteriors for each of the acquired measurements, which is a crucial feature for multi compound gas distribution mapping [12]. The Km-P algorithm was extensively evaluated in different scenarios, with different robotic platforms and with up to four unknown chemicals and the results show its feasibility. Recently, we submitted a journal article related to the Km-P algorithm to Sensors and Actuators B: Chemical, which is arguably one of the most reputed venue in the chemical and gas sensing community. The article titled "A Cluster Analysis Approach Based on Exploiting Density Peaks for Gas Discrimination with Electronic Noses in Open Environments" was submitted on May, 2017 and it is currently under review.

We tested the suitability of the Km-P algorithm for online gas discrimination in an uncontrolled environment. In a machine workshop, we set two artificial gas sources (ethanol and acetone) evaporating from plastic containers. The robot was first placed outside of the workshop, where it learned the baseline (i.e. clean air) response. After the baseline response was learned, the robot was manually driven inside the workshop, where it explored the environment for several minutes. Fig. 8 shows the setup of the experiments.

Fig 8: Experimental setup where the Km-P algorithm was evaluated. Two gas sorces were placed in the environment and the robot successfully learned the presence of the two airborne chemicals (ethanol and propanol).

Figure 9 show the outputs from the Km-P Algorithm. The plots show an estimation of the gas concentration (Fig 9.a) and the class posteriors (Fig 9.b). The Km-P algorithm re-learns a new

discrimination model first when a given number of unknown measurements are detected and thereafter in a periodical way.

Notice that early in the experiment, when the sensors are at baseline response, the Km-P algorithm estimates a high posterior for clean air (yellow line). Later, the robot moves towards the first gas source and, after a few consecutive measurements near the gas source, acquires enough data to learn a new type of gas. The output from the Km-P algorithm (Fig 9.b) shows a high confidence with respect to analyte 1 (blue line), which was learned as the robot was collecting measurements. When the robot moves away from the gas source, and the sensors start to recover their baseline response level, the Km-P algorithm predicts a higher posterior for clean air. As the robot moves forward towards the second gas source, a new model is learned and as can be seen in Fig 9.b, a new class is learned and a high posterior is estimated by the algorithm as data is collected (brown line) as the robot moves away from the gas source, the likelihood of air rises.

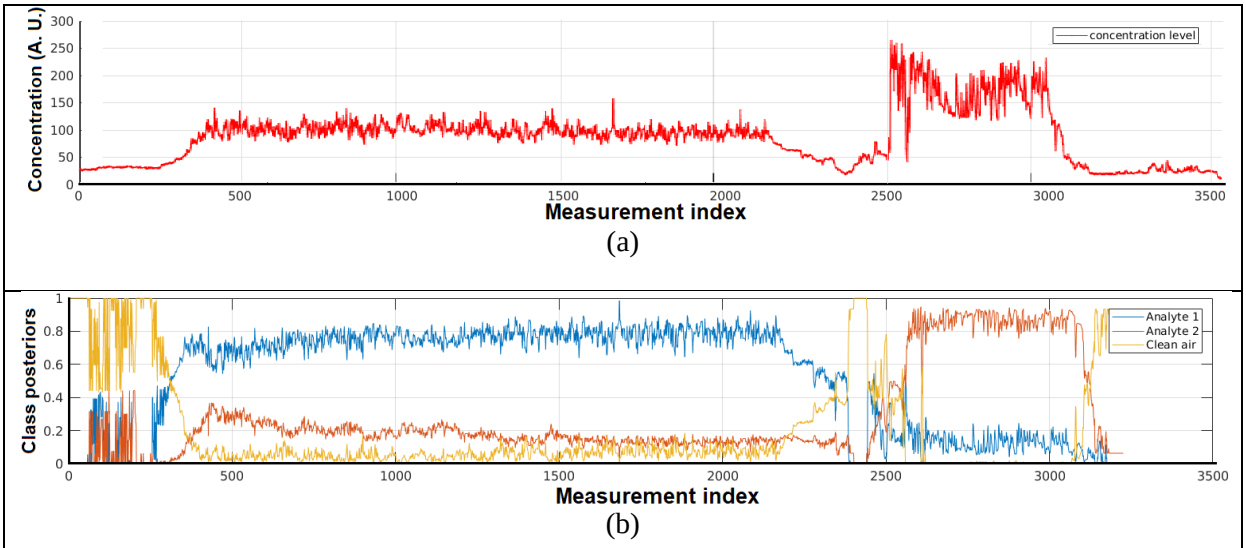

Fig 9: Exploration run in the presence of two unknown chemicals. (a) Gas concentration indicator. (b) Gas identity class posteriors estimated using the Km-P algorithm.

It is worth notice again that the algorithm started with zero knowledge or training data before moving into the exploration scenario. Such challenge is expected in an emergency situation, where gas leaks might have occurred. While firefighters might have some a-priori knowledge about the type of gases that might be present, it is highly likely that unknown interferents may be present, for example, as by-products of fire or explosions.

D Summary and Outlook

During the second year into Smokebot's project life cycle, a novel airflow distribution mapping was developed. As previously stated, airflow modeling has seldom been addressed at micro-scales, and while there exist a small number of solutions, these algorithms are either too simplistic or

computationally expensive to be carried out on-board a mobile robot. Thus, we proposed a data driven approach to conduct. The distinctive characteristic of the solution developed in Smokebot is the extrapolation method, used to estimate airflow PDFs at non visited location. This particular extrapolation algorithm models airflow as a combination of laminar and turbulence characteristics. We proved that such method can outperform conventional extrapolation techniques.

Regarding gas distribution mapping, we explored how to incorporate the output of different low-level tasks. First, we showed that airflow models can be integrated in the gas distribution maps. For this purpose, we developed the Kernel DM+V/Γ algorithm, which builds upon the airflow dependent algorithm originally proposed by Reggente et al. in [8]. We showed how the use of an airflow model, instead of mean airflow vectors can capture the complexities of turbulent flow in a better way, since gas concentrations are extrapolated through a set of multimodal kernels. These multimodal kernels reflect conditions likely to be observed in uncontrolled scenarios, where more than one dominant airflow or can be present.

Another aspect that we explored in Smokebot was the creation of multi compound gas distribution maps under the presence of multiple unknown airborne chemicals. In order to do so, we developed the Km-P gas discrimination algorithm, which can be used in uncontrolled environments and that can estimate class posteriors for the gas identities, which is a crucial requirement when creating multi-compound maps. We tested the capabilities of the Km-P algorithm in a complex environment where two gas sources where placed. With no a-priori knowledge about the identity or the number of gas sources, the Km-P algorithm successfully predicted the presence of two chemicals in the environment in an online way.

As previously mentioned, the above contributions resulted in two publications. The first one, related to airflow modeling was a dual journal/conference publication (see annex) while the Km-P unsupervised gas discrimination algorithm is currently under review. In the upcoming months, the following tasks will be further developed as follows:

- A new publication, related to the Kernel DM+V/ Γ will be submitted. This publication will include the developments presented in this report and also will include a comprehensive numerical evaluation (currently being carried out) that will show how the inclusion of airflow models can improve the quality of the gas distribution maps. This task can be completed by M36 of the project's development cycle.
- The Km-P will be further improved. In its current implementation, the algorithm updates its gas discrimination model periodically and, when the number of measurements substantially grows, the updating of the model requires higher computational resources. An improved version of the algorithm will update the model only when required and the implementation will be optimized in order to lower the computational demands. In addition, the Km-P algorithm will be combined with a multi compound gas distribution mapping algorithm, such as the one presented in [12]. This will require improvements on the original algorithm in order to handle measurements with unknown identities and the re-computation of the gas

distribution map, when a new gas identity has been learned. These tasks will be conducted during Year 3 of the project´s life cycle.

- We will integrate gas distribution mapping with the high bandwidth sensor nodes developed in T2.1. This effort has already started and experiments with Smokebot's novel gas sensing units have been carried out in cooperation between UWAR and ORU. A conference publication, that will showcase the capabilities of the novel sensor box against conventional sensors, is likely to be submitted before the end of the project.
- Integration of gas distribution and airflow mapping with tasks related to WP1. In particular, we will use the information from T1.2 and T1.3 to compute obstacle-dependent gas and airflow maps. While this task has already started, further development and validation are required. These tasks will be conducted during Year 3 of the project´s life cycle.
- Integration of gas distribution mapping with tasks from WP4: Situation Analysis and WP5: Autonomous behavior. For example, the produced gas distribution maps will be integrated in the General Disaster Information Model (T4.1) so it can be used by related tasks such as detection and prediction of hazards (T4.2) and self-preservation (T5.2).

References

[1] Hernandez Bennetts V, Lilienthal AJ, Neumann PP, Trincavelli M. Mobile Robots for Localizing Gas Emission Sources on Landfill Sites: Is Bio-Inspiration the Way to Go? *Frontiers in Neuroengineering*. 2011;4:20. doi:10.3389/fneng.2011.00020.

[2] Beaucage, P., Brower, M.C., Tensen, J. Evaluation of four numerical wind flow models for wind resource mapping (2014) Wind Energy, 17 (2), pp. 197-208.

[3] A. Manor and S. Berkovic, "Bayesian inference aided analog downscaling for near-surface winds in complex terrain," Atmospheric Research, vol. 164165, pp. 27 – 36, 2015.

[4] S. Kato and S. Murakami, "New ventilation efficiency scales based on spatial distribution of contaminant concentration aided by numerical simulation," American Society of Heating, Refrigerating and Airconditioning Engineers (ASHRAE) Transactions, pp. 09–330, 1988.

[5] G. Cao, H. Awbi, R. Yao, Y. Fan, K. Sirn, R. Kosonen, and J. J. Zhang, "A review of the performance of different ventilation and airflow distribution systems in buildings," Building and Environment, vol. 73, pp. 171 – 186, 2014.

[6] V. Hernandez Bennetts, A. J. Lilienthal, A. A. Khaliq, V. Sese and M. Trincavelli, Towards Real-World Gas Distribution Mapping and Leak Localization Using a Mobile Robot with 3D and Remote Gas Sensing Capabilities. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2013, 2327-2332.

[7] E. Schaffernicht, V. Hernandez and A. J. Lilienthal. Mobile Robots for Learning Spatio-temporal Interpolation Models in Sensor Networks – The Echo State Map Approach . 2017 IEEE International Conference on Robotics and Automation (ICRA). To appear

[8] M. Reggente and A. J. Lilienthal, Using Local Wind Information for Gas Distribution Mapping in Outdoor Environments with a Mobile Robot. Proceedings of IEEE Sensors, 2009, 1715-1720.

[9] M. Reggente and A. J. Lilienthal, The 3D-Kernel DM+V/W Algorithm: Using Wind Information in Three Dimensional Gas Distribution Modelling with a Mobile Robot. Proceedings of IEEE Sensors, 2010.

[10] A. J. Lilienthal, M. Reggente, M. Trincavelli, J. Blanco, and J. Gonzalez, A Statistical Approach to Gas Distribution Modelling with Mobile Robots – The Kernel DM+V Algorithm. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2009, pp. 570-576.

[11] A. Loutfi, S. Coradeschi, A. J. Lilienthal and J. Gonzalez, Gas Distribution Mapping of Multiple Odour Sources using a Mobile Robot. Robotica, 27:2, 2009, pp. 311-319

[12] V. Hernandez Bennetts, E., Victor Pomareda, A. J. Lilienthal, S. Marco and M. Trincavelli, Combining Non Selective Gas Sensors on a Mobile Robot for Identification and Mapping of Multiple Chemical Compounds . Sensors MDPI 2014, 14: 9, 2014,17331 - 17352 .

Annex

Probabilistic Air Flow Modelling Using Turbulent and Laminar Characteristics for Ground and Aerial Robots

Victor Hernandez Bennetts, Tomasz Piotr Kucner, Erik Schaffernicht, Patrick P. Neumann, Han Fan, and Achim J. Lilienthal

***Abstract*—For mobile robots that operate in complex, uncontrolled environments, estimating air flow models can be of great importance. Aerial robots use air flow models to plan optimal navigation paths and to avoid turbulence-ridden areas. Search and rescue platforms use air flow models to infer the location of gas leaks. Environmental monitoring robots enrich pollution distribution maps by integrating the information conveyed by an air flow model. In this paper, we present an air flow modelling algorithm that uses wind data collected at a sparse number of locations to estimate joint probability distributions over wind speed and direction at given query locations. The algorithm uses a novel extrapolation approach that models the air flow as a linear combination of laminar and turbulent components. We evaluated the prediction capabilities of our algorithm with data collected with an aerial robot during several exploration runs. The results show that our algorithm has a high degree of stability with respect to parameter selection while outperforming conventional extrapolation approaches. In addition, we applied our proposed approach in an industrial application, where the characterization of a ventilation system is supported by a ground mobile robot. We compared multiple air flow maps recorded over several months by estimating stability maps using the Kullback–Leibler divergence between the distributions. The results show that, despite local differences, similar air flow patterns prevail over time. Moreover, we corroborated the validity of our results with knowledge from human experts.**

***Index Terms*—Aerial systems: perception and autonomy, environment monitoring and management, field robots, mapping.**

I. Introduction

AIR Flow Modelling (AFM) is the task of creating spatio-temporal models of variables such as wind speed and direction. AFM has been extensively studied in meteorology and atmospheric research [1]. AFM is a challenging task since instantaneous patterns in the wind flow are hard to predict. Especially at surface level, laminar (e.g. uniform) patterns are rarely observed and instead, air flow shows high spatio-temporal variations. Temporal variations are observed at different scales ranging from long term variations (years) to turbulent variations (minutes to seconds). Spatial variations are also observed at different scales namely, macro scales (distances up to 5000 km), meso scales (distances up to 2000 km) and micro scales (distances smaller than 2 km). The vast majority of the AFM approaches are aimed at macro and meso scales and only a limited number of methods are available for micro scale, turbulence-dominated regimes.

Nonetheless, micro scale air flow models have a wide range of applications in different fields of study, among others, mobile robotics. For example, air flow models can provide valuable information to aerial robots. An example of the applicability of AFM is [2], presented by Al-Sabban *et al.* The authors propose a method based on Gaussian models and Markov Decision Processes to plan a path for a fixed-wing UAV that considers time-varying wind fields in order to minimize the onboard energy usage by planning for efficient exploration trajectories. Thus, knowledge about the distribution of the air flow is critical.

Another line of research that benefits from AFM in micro scales is Mobile Robot Olfaction (MRO), where gas sensitive robots perform tasks such as gas leak localization and gas distribution mapping. In MRO, wind data can be used to determine the exploration trajectory of the robot [3] and to improve modelling of the spatial distribution of gas concentrations [4]. Recently, in [5], it is shown that active air flow generation can help to localize gas leaks, emphasizing in this way, the limitations of the current use of natural air flow information. In most of the previous application examples, average values from wind speed and direction were used by the models. However, averaged wind information neither captures the complexities of turbulence nor fully considers the spatial variability of the air flow [6].

The use of mobile robots in previously unexplored applications, such as occupational health and pollution surveillance, are further fields where robot-aided AFM can be useful. An example of such application is proposed in [7], where a hybrid pollution monitoring system was presented. The hybrid system comprises both, a stationary sensor network and a mobile

Manuscript received September 10, 2016; accepted January 16, 2017. Date of publication January 31, 2017; date of current version February 16, 2017. This paper was recommended for publication by Associate Editor F. Kendoul and Editor J. Roberts upon evaluation of the reviewers' comments. This work was supported in part by the Swedish Knowledge Foundation project RAISE (20130196) and the H2020-ICT project SmokeBot (645101).

V. Hernandez Bennetts, T. P. Kucner, E. Schaffernicht, H. Fan, and A. J. Lilienthal are with the AASS Research Centre, School of Science and Technology, Örebro University, Örebro 701 82, Sweden (e-mail: victor.hernandez@oru.se; tomasz.kucner@oru.se; erik.schaffernicht@oru.se; han.fan@oru.se; achim.lilienthal@oru.se).

P. P. Neumann is with the Bundesanstalt für Materialforschung und -prüfung, Berlin 12205, Germany (e-mail: Patrick.Neumann@bam.de).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LRA.2017.2661803

2377-3766 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

(a) (b)

Fig. 1. AFM application scenarios considered in this paper. (a) Aerial robotics. (b) Robot-aided ventilation characterization in a foundry hall.

robot. In the application scenario of [7], it is argued that the spatially-dense but temporally sparse models acquired with a robot can complement the temporally-dense, spatially sparse data acquired with an on-site sensor network. Similarly, air flow models, computed from robot data, can complement measurements acquired with air flow sensors placed at a limited number of locations. Both sources of information combined can then be used to characterize ventilation patterns in industrial and living spaces. Occupational health experts can use such models to understand how pollutants are spread in the work environment and apply corrective actions. Similarly, in the field of urban pollution monitoring, developers and authorities can benefit from robot-aided AFM and data from weather stations (which are usually sparse in number) to improve air pollution models. These models can in turn help to understand how different attributes of the urban form interact and affect pollution levels.

The contributions of this paper with respect to micro scale AFM in general and mobile robotics in particular are:

1) We introduce a novel AFM algorithm for micro scale environments that considers laminar and turbulent characteristics of the air flow. The algorithm uses data collected in a sparse number of measurement locations and computes a model that predicts joint Probability Density Functions (PDFs) of wind speed and direction at, e.g., non visited locations. Furthermore, an implementation of our proposed algorithms is available in a public repository for the interested reader.[1]
2) We explore the applicability of our algorithm in aerial robotics using data collected with a quadcopter (Fig. 1(a)). The produced air flow models show the air flow distributions and their variations under different environmental conditions.
3) We demonstrate that the proposed method has a better performance, compared to standard extrapolation methods. Moreover, we show the stability of our algorithm to the selection of its functional parameters.
4) We address a practical application, namely the characterization of the ventilation system of an industrial scenario (the foundry hall in Fig. 1(b)), using the proposed framework. The obtained results and input from human experts suggest that useful air flow models can be estimated in turbulence dominated environments.

[1]A Matlab implementation of the proposed algorithm is available for download at https://github.com/OrebroUniversity/LT-Algorithm

II. Related Work

As previously stated, most of the AFM-related work has been centred around macro/meso scales with low temporal resolution (e.g. months, days), using weather stations [1]. In micro-scales, simple AFM approaches include the use of trace gases, as presented in [8], or the use of sophisticated Computational Fluid Dynamics (CFD) simulations [9]. However, these methods are not directly applicable to mobile robotics. For example, CFD-based techniques require precise a-priori knowledge on boundary conditions and moreover, these techniques require high computational resources, which may not be available on-board mobile platforms.

Data-driven AFM techniques are a viable option for mobile robotics platforms. To perform AFM, mobile robots should be able to acquire wind information while exploring the environment. In the case of ground robots, this is accomplished by equipping the platforms with dedicated wind sensors (i.e. anemometers). In the case of flying robots such as quadcopters, the use of anemometers is restricted due to the limited payloads and the disruptions introduced by the rotors of the platform. Thus, wind information is estimated from inertial measurement units that consist of accelerometers, gyroscopes, magnetometers, as shown in [10]–[12].

In robotic applications, air flow is normally considered as a disturbance to the control problem and has seldom been explicitly modelled. To the authors best knowledge, there are only a few works that explicitly address the problem of AFM in the context of mobile robotics. In [13], Kowadlo *et al.* proposed a naïve physics air flow model for gas source localization. This algorithm is rule-based and the confidence in a limited number of hypotheses is updated according to acquired wind measurements. More recently, the work presented in [14] introduced a probabilistic method for mapping dynamics in the context of robotics. The authors model velocity fields as a set of Semi-Wrapped Gaussian Mixture Models to model the flow of discrete objects (e.g. people, vehicles) mainly. In addition, the paper explored the applicability of their approach in the context of AFM by estimating a wind field from a set of sparse robot measurements, but they did not include domain-specific knowledge such as turbulence. Rodriguez and co-authors presented in [15] a wind field estimation algorithm for flying robots. The authors first estimated the wind vectors at different positions by combining different sensors, such as GPS, and then used the estimated wind vectors to fit Weibull distributions to model only the wind speed at different altitudes. In comparison, the AFM algorithm presented in this work models both, wind speed and direction at different locations in an explored area.

III. Algorithm

The data-driven Air Flow Modelling (AFM) approach presented in this paper uses localized wind speed and direction measurements, acquired at a sparse number of locations, as inputs. The algorithm's output is a 2-D model (i.e. a map) in which a discrete representation of the air flow distribution (i.e. a histogram) is modelled for each query location. The AFM process is divided into two stages, namely the estimation of the

joint Probability Density Function (PDF) at the measurement locations, and an extrapolation process to infer joint PDFs at non-visited locations.

A. PDF Estimation at Measurement Locations

In the first stage of the computation process, joint PDFs of the wind speed (ν) and wind direction (θ) are iteratively computed, at distinct measurement locations m, as data is being acquired. The joint PDFs are modelled as a discrete, two dimensional space Γ^m that has a total of N_s states. Each state J in Γ^m is a (θ, ν) tuple of wind direction and speed. Given measurements $z = (\theta, \nu)$, the joint pdf Γ^m and hence, the probability of state J given the set of measurements up to time t (1a) is given as:

$$\Gamma^m(J|\mathbf{z}_{1:t}) = \frac{\Gamma^m(J|\mathbf{z}_{1:t-1}) + p_{J,t}}{\sum_{j=1}^{N_s}\Gamma^m(j|\mathbf{z}_{1:t})} \tag{1a}$$

$$p_{J,t} = \frac{\bar{p}_{J,t}\mathcal{N}_s(z_t|J)}{\sum_{i=1}^{N_s}\bar{p}_{J_i,t}\mathcal{N}_s(z_t|J_i)} \tag{1b}$$

$$\bar{p}_{J,t} = \sum_{i=1}^{N_s} p_{J_i,t-1}\mathcal{C}(J|J_i, \mathbf{z}_{1:t}) \tag{1c}$$

Equation (1a) updates the probability for each state J in Γ^m by $p_{J,t}$. N_s denotes the number of states in Γ^m. $p_{J,t}$ is estimated using a Bayes filter [16], as defined by (1b) and (1c). The Bayes filter incorporates knowledge about the wind sensing device using the sensor model $\mathcal{N}_s$ as follows:

$$\mathcal{N}_s(z|J) = \frac{1}{2\pi\sigma_\theta\sigma_\nu}\mathbf{exp}\left(-\frac{\Delta(z_\theta, J_\theta)^2}{2\sigma_\theta^2} - \frac{|z_\nu - J_\nu|^2}{2\sigma_\nu^2}\right) \tag{2}$$

In the above equation, Δ computes the orthodromic (i.e. great circle) distance [17] between measurement z_θ and the state center J_θ accounting in this way, for the circularity of the wind direction variable. Parameters σ_θ and σ_ν correspond to the confidence intervals with respect to wind direction and speed. The value of parameters σ_θ and σ_ν are often provided by the sensor's manufacturer. Wind variability is accounted for in the Bayes filter by defining a transition model $\mathcal{C}$ (1c) as follows:

$$\mathcal{C}(J|J_i, \mathbf{z}_{1:t}) = \mathcal{C}(J|J_i, z_{1:t-1}) + \mathcal{N}_s(z_t|J) \tag{3}$$

The transition model $\mathcal{C}$ between current state J and a previous state J_i assigns probabilities based on the frequency of each transition pair (J, J_i). This means that higher posteriors will be assigned to states that do not exhibit intermittent transitions but rather stable/frequent conditions over time. This allows to discriminate between dominant wind states and turbulence. $\mathcal{C}$ does not make a priori assumptions and it is learned as data is recorded, according to (3).

Fig. 2 shows an example run to compute Γ^m from 1500 randomly generated measurements. The measurements simulate high turbulence conditions and a stable, dominant air flow direction centred at (250°, 0.15 m/s) and with standard deviations of 25° in θ and 0.08 m/s in ν. A total of 70% of the data points correspond to randomly generated noise. Notice that, compared to a simple histogram, the proposed method clearly identifies the dominant air flow direction by assigning higher probabilities to the corresponding states J.

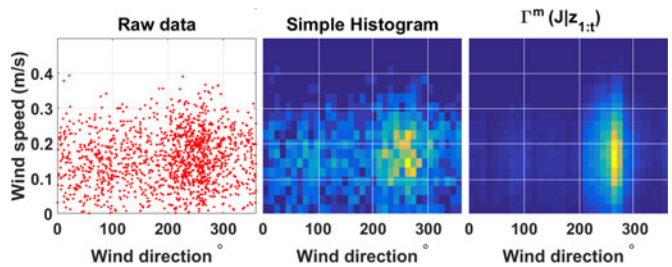

Fig. 2. Comparison of PDFs generated with a simple histogram and the proposed method (Γ^m).

B. Extrapolation Stage

The extrapolation stage uses $\Gamma^m_{1:N_m}$, to predict PDFs Γ^q at non-visited query locations q. In this paper we propose the $\mathcal{LT}$ algorithm, which computes Γ^q as the linear combination of a laminar flow component ($\mathcal{L}$) and a turbulent component ($\mathcal{T}$). The laminar component $\mathcal{L}$, is computed using a set of N_a artificial samples $\hat{z}_i = (\hat{\theta}_i, \hat{\nu}_i)$ from $\Gamma^m_{1:N_m}$ as follows:

$$\mathcal{L}_q(J|\Gamma^m_{1:N_m}) = \frac{\sum_{i=1}^{N_a}\mathcal{N}_s(\hat{z}_i|J)}{\sum_{j=1}^{N_s}\mathcal{L}_q(j|\Gamma^m_{1:N_m})} \tag{4a}$$

$$\hat{\nu}_i = \frac{\sum_{r=1}^{N_m}\omega_r^q\nu_r}{\sum_{r=1}^{N_m}\omega_r^q} \tag{4b}$$

$$\hat{\theta}_i = \text{atan2}\left\{\sum_{r=1}^{N_m}\omega_r^q\cos(\theta_r), \sum_{r=1}^{N_m}\omega_r^q\sin(\theta_r)\right\} \tag{4c}$$

$$\omega_r^q = \text{RBF}(|\mathbf{q}_i - \mathbf{m}_r|, \sigma) \tag{4d}$$

Equation (4b) implements the Nadaraya-Watson [18] estimator and (4c) computes the weighted circular mean [17]. Both computations use a Radial Basis Function (RBF) kernel with a bandwidth σ (4d). However, mean values $\hat{\theta}_i$ (4c) convey little to no information regarding multimodal (e.g. turbulent) wind flow distributions and moreover, $\hat{\theta}_i$ is undetermined if wind vectors are opposed to each other. A good indicator of multimodality is the circular variance (CV), which is the variation in the angles around a mean direction. CV is bounded between 0 and 1, reaching its maximum when the wind vectors are opposed to each other [19].

Thus, the $\mathcal{LT}$ algorithm estimates Γ^q by weighting between a laminar component ($\mathcal{L}$) and a turbulent component ($\mathcal{T}$) using a turbulence indicator ρ, that depends on the circular variance of $\mathcal{T}^q$, as follows:

$$\Gamma^q(J|\Gamma^m_{1:N_m}) = (1-\rho)\mathcal{L}^q(J|\Gamma^m_{1:N_m}) + \rho\mathcal{T}_q(J|\Gamma^m_{1:N_m}) \tag{5a}$$

$$\mathcal{T}_q(J|\Gamma^m_{1:N_m}) = \frac{\sum_{i=1}^{N_m}\omega_r^q\Gamma^m_{1:N_m}(J|\mathbf{z}_{1:t})}{\sum_{j=1}^{N_s}\mathcal{T}(j|\Gamma^m_{1:N_m})} \tag{5b}$$

$$\rho = \frac{1}{1 + e^{-\kappa CV(\mathcal{T}_q)}} \tag{5c}$$

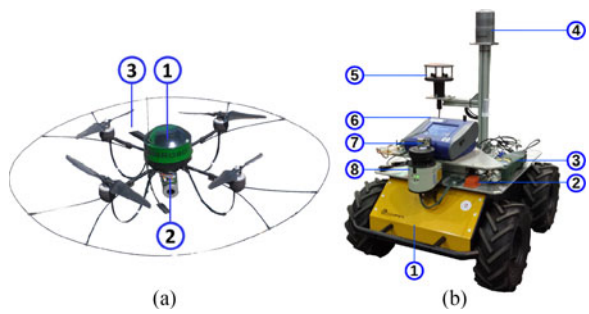

Fig. 3. (a) Aerial and (b) ground robots. (a): 1- Airrobot AR100-B quadrocopter. 2- Gas sensing module. 3- Transportation tube. (b): 1- Husky A-200. 2- Inertial Measurement Unit (IMU). 3- Photoionisation gas detector. 4- Velodyne HDL-32E scanner. 5- Windsonic anemometer. 6- Dusttrak particle sensor. 7- air quality measurement module. 8- LMS121 scanner.

Notice that, while the $\mathcal{L}$ component (4a) is histogram computed from N_a weighted averages of artificial samples drawn from the PDFs Γ^m, the $\mathcal{T}$ component approximates the turbulent distributions by performing a linear combination of the PDFs Γ^m weighted by the RBF kernel.

The extrapolation stage depends on three free parameters, namely σ (4d), κ (5c) and N_a (4a). σ controls the smoothness of the extrapolation in $\mathcal{L}$ and the mixing of the PDFs Γ^m in $\mathcal{T}$. κ corresponds to the steepness of the logistic function that computes the turbulence indicator ρ. N_a controls the number of artificial samples used to compute $\mathcal{L}$. In Section V, we show the stability of the $\mathcal{LT}$ extrapolation method with respect the selection its parameters.

IV. Experimental Setup

The proposed AFM framework was evaluated with real world data collected in two different challenging scenarios (indoor and outdoor) with two different robot platforms. In the outdoor scenario (Fig. 1(a)), a micro UAV, namely an AirRobot AR100-B (AirRobot GmbH & Co. KG, Germany), was deployed to collect wind measurements (Fig. 3(a)). The wind vector is estimated by fusing the micro UAV's on-board sensors (GPS and IMU) to compute the parameters of the wind triangle as presented in [12], [20]. The wind triangle is commonly used in aerial navigation and describes the relationships between the flight vector $\vec{v}$, the ground vector $\vec{w}$, and the wind vector $\vec{u}$. We carried out 11 measurement tours, each one with an approximate duration of 14 to 19 minutes. In each tour, the micro UAV was programmed to explore the experimental areas following a sweeping trajectory, where it stopped at approximately 30 waypoints to take wind samples for about 20 s, with a sampling rate of 8 Hz.

For the indoor scenario, we consider an industrial foundry that casts bronze bearings (Fig. 1(b)). Several activities are carried out in this scenario: For example, bronze scrap is melted and then transported to locations where bearings are being cast. During data collection, personnel was instructed to conduct their assigned tasks as usual and thus, no disruption or artificial conditions were introduced [7]. Moreover, the ventilation system of the facility was operating according to the schedule defined by the foundry's operators.

In the case of the indoor scenario, wind measurements were collected with a ground robot (Husky A-200[2]) equipped with an HDL-32E 3D range sensor for localization and 3D mapping (Fig. 3(a)). Wind sensing was performed with a Windsonic anemometer. We performed 12 measurement tours through the foundry each one of a duration between 1.5 to 2 hours. In each measurement tour, the robot stopped in different way-points and collected wind measurements for 120 s, with a 4 Hz sampling rate. The first six tours were recorded during a measurement week in October 2015, while the other five tours were conducted in February 2016. As shown in Fig. 3, the payload of both robots include additional sensors that are not used in this work.

To assure robust robot localization, which is critical to associate measurements with spatial locations, we used GPS for the aerial robot and in the case of the ground robot, we employed the Normal Distribution Transform Monte Carlo Localization (NDT-MCL) in large, 3-D environments [21].

V. Results

The wind vector estimation approach used in this work was evaluated in two different modes, namely hovering and flight. For the validation during hovering mode, the micro UAV was manually positioned close to an anemometer (2 m to 5 m), in a way that the generated air flow of the micro UAV's rotors did not influence the anemometer. Then, the micro UAV was controlled automatically for about 20 min using the on-board GPS to collect wind data. For the in-flight (moving over ground) validation, four waypoints were placed around an anemometer in a square of size 30×30 m^2. Three trials were performed, each taking approximately 6 min, in which the micro UAV autonomously visited each waypoint twice. For the evaluation of these results, the assumption has to be made that the wind vector measured at the anemometer and the micro UAV are comparable, i. e., that the wind field does not change drastically over the distance between the reference measurement with the anemometer and the position of the micro UAV. The calculated RMSE (between in-flight measurements and the anemometers) for the wind speed and direction is ± 0.60 m/s and $\pm 14^\circ$ with the micro UAV hovering close to the anemometer and ± 0.36 m/s and $\pm 15^\circ$ with the micro UAV flying around. The interested reader can consult [12] for further details on the evaluation process.

A quantitative evaluation of the $\mathcal{LT}$ algorithm was performed using micro UAV data coming form 11 trials, over two different days with varying environmental conditions. At each trial, the micro UAV was commanded to stop at N_m waypoints to collect wind data. The sensor model $\mathcal{N}_s$ was defined according to the experimental validation of the wind triangle approach described above. Thus, parameters σ_θ and σ_ν were set to 15° and to 0.36 m/s respectively.

The generated air flow maps were evaluated using two different leave-one-out cross-validation (LOOCV) processes. In both cases, the data from a single measurement location was

[2]http://www.clearpathrobotics.com/husky/

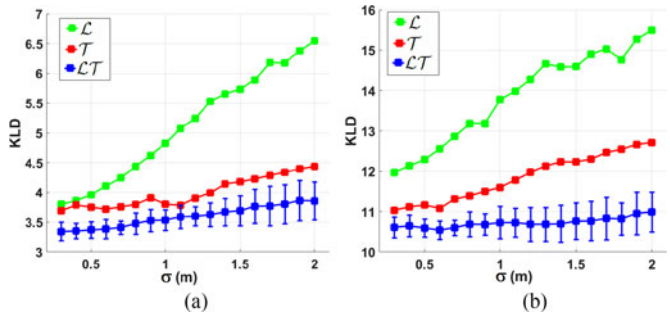

Fig. 4. Performance curves of the proposed algorithm (blue errorbar curves) against laminar-only (green curves) and turbulence-only (red curves) extrapolation approaches. Lower KLD values denote a higher similarity between distributions. (a) Γ^m vs Γ^q. (b) Γ^{raw} vs Γ^q.

kept for evaluation and the data from the $N_m - 1$ remaining locations was used to train air flow models using three different approaches namely, a laminar-only approach ($\mathcal{L}$, (4a)), a turbulence-only approach ($\mathcal{T}$, (5b)) and the the $\mathcal{LT}$ algorithm proposed in this work.

In the first LOOCV, a PDF Γ^m was computed from the left-out data and the predictive capability of the generated air flow maps was evaluated using the Kullback Leibler Divergence (KLD) between the prediction Γ^q and Γ^m, computed from the left out measurements. In the second LOOCV, the left-out data was used to compute a histogram Γ^{raw}, which was compared against the prediction Γ^q using the KLD. Fig. 4(a) and (b) show the performance of the $\mathcal{L}$, $\mathcal{T}$ and $\mathcal{LT}$ maps computed with $\sigma = [0.1, 0.2, ..., 1.0]$ and $\kappa = [1, 5, 10, 20, 30]$. Considering all parameter combinations, the 11 runs and the number of waypoints per run, around 15000 models were trained and tested for each LOOCV. In both evaluations, the $\mathcal{LT}$ algorithm has a better performance compared to $\mathcal{L}$-only and $\mathcal{T}$-only models. In addition, the $\mathcal{LT}$ algorithm shows a high robustness in the selection of its parameters (κ, σ), as denoted by the blue performance curve. It can be seen in Fig. 4(a) and (b) that while the selection of the kernel bandwidth σ has a substantial effect on the performance of the $\mathcal{L}$-only and $\mathcal{T}$-only models (green and red lines respectively), the performance of the $\mathcal{LT}$ algorithm remains relatively stable, outperforming both, $\mathcal{L}$-only and $\mathcal{T}$-only models.

We evaluated the sensitivity of the $\mathcal{LT}$ algorithm to the number of artificial samples (N_a, (4a)). With data from a single micro UAV run, we computed the KLD between a model computed with $N_a = 200$ against models computed with increasing $N_a = [10, 20, \cdots, 190]$. The small KLD values shown in the error bar plot in Fig. 5 denote a high degree of similarity between models, proving in this way, the stability of the $\mathcal{LT}$ algorithm to the selection of N_a.

Fig. 6(a) and (b), show examples of predictions at left-out locations. The left plot in both figures shows the left-out data, the center plot shows the histogram computed from the left-out data (Γ^{raw}) and the right plot corresponds to the predicted Γ^q. In both examples, Γ^q fairly matches Γ^{raw}, as denoted by the computed KLD. In addition, $\rho = 0$ is predicted for both examples. This means that the predicted wind flow distribution

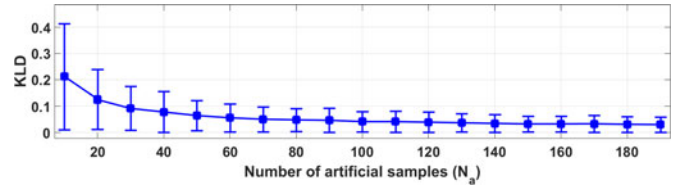

Fig. 5. Performance curve for the selection of N_a. A model computed with $N_a = 200$ is compared against models computed with an increasing number of artificial samples N_a.

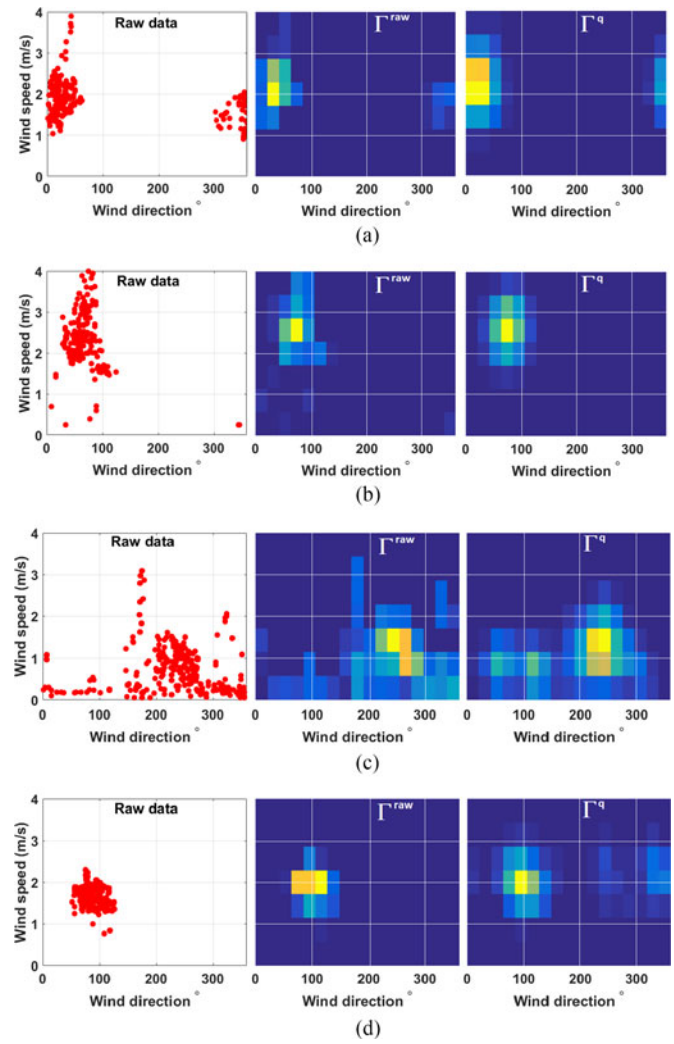

Fig. 6. Comparison between left out data, its corresponding histogram Γ^{raw} and the predicted Γ^q at different locations. (a) $\rho = 0.0$, KLD = 0.69 (b) $\rho = 0.0$, KLD = 0.32 (c) $\rho = 0.90$, KLD = 1.83 (d) $\rho = 0.79$, KLD = 9.02.

is highly likely to have unimodal, well defined laminar wind flow distributions, which also concurs with the actual distribution of the left-out data.

Fig. 6(a) and (b) show examples where the $\mathcal{LT}$ algorithm predicts turbulent conditions (i.e. high ρ values). In Fig. 6(c), Γ^q successfully predicts a dominant wind speed and directions and low-speed turbulent flows with varying wind directions. Fig. 6(d) shows an example where the $\mathcal{LT}$ algorithm wrongly predicted the presence of a turbulent flow. This was due to the

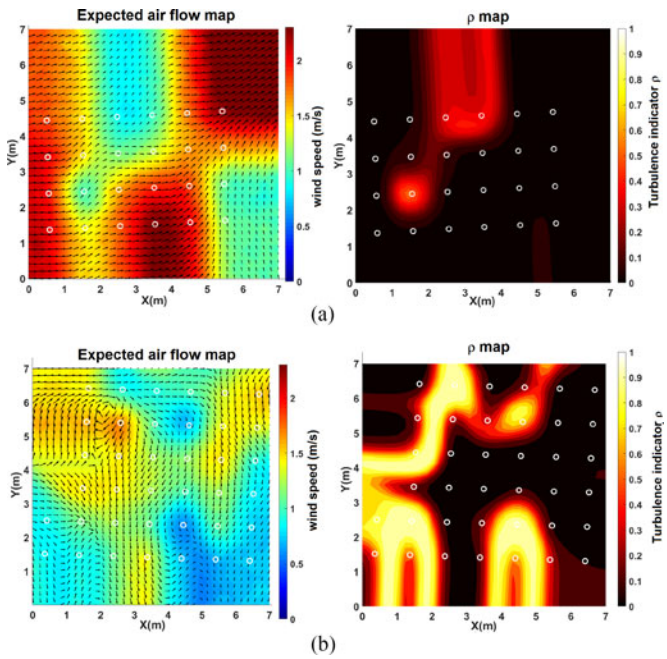

Fig. 7. Air flow maps generated for two different exploration runs with the micro UAV. The figure above displays the expected air flow maps and a distribution map of the turbulence indicator ρ. (a) 2011-07-07. (b) 2011-07-12.

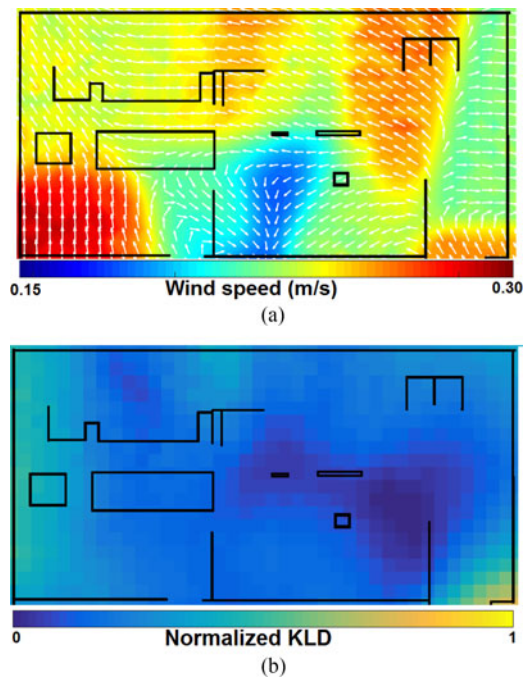

Fig. 8. Ventilation characterization results. (a) A resulting wind map computed from a single inspection run. (b) KLD based stability map. A lower (darker) value represents similar estimates of the wind distribution over time, while higher (brighter) values indicate different estimates.

spread of the wind direction measurements at the neighbouring locations. While Γ^q accurately predicts the dominant airflow, a multi modal distribution is wrongly predicted, compared with the left-out data.

Fig. 7(a) and (b) show air flow maps computed with data from two different quadcopter runs. For simplicity, the maps show the expected wind speed and direction at each query location. The maps were computed off-line using an unoptimized Matlab implementation of the $\mathcal{LT}$ algorithm on a standard PC (1.7 GHz Intel Core i7, 8GB RAM). With the previously mentioned computational resources and software implementation, updating both, the joint PDFs Γ^m and the transition models $\mathcal{C}$, takes an average of 12 ms. The time required to estimate a joint PDF Γ^q at a non visited location takes an average of 0.46 s using 50 samples artificially drawn from the joint PDFs Γ^m (see (4b) and (4c)).

The expected air flow map for the experiment performed on 2011-07-07 (Fig. 7(a)), shows stable flow conditions and a well defined wind direction while in the case of the experiments conducted on 2011-07-12 (Fig. 7(b)), low wind velocities, eddies and vortices are predicted at different locations in the explored environment. Distribution maps of the turbulence indicator ρ can convey useful information. In the experiment conducted on 2011-07-12, areas of high wind variability are predicted at different locations in the map. A quadcopter that aims to traverse a given area efficiently, as reported in [2], could use the ρ map to navigate only in areas with high wind stability. In the case of the experiment conducted on 2011-07-07, a more laminar, stable airflow distribution is predicted, as shown in the corresponding expected and ρ maps.

As stated before, one of the applications of AFM is the characterization of ventilation systems. We explore the applicability of our proposed AFM algorithm for ventilation characterization using a ground robot inspecting a 40 m $\times$ 70 m foundry hall. To account for the lack of long term data, since no stationary anemometers were placed in the foundry, we performed 11 measurement tours during two different seasons (winter and autumn), as explained in Section IV. The sensor model used to compute the wind maps was defined according to the anemometer's specifications provided by the manufacturer, namely $\sigma_\theta = 2^\circ$ and $\sigma_\nu = 0.2$ m/s.

Besides the wind maps, that can be considered as time limited snapshots of the air flow, we are interested in the stability of the results over time. Thereto, we compare the distributions in each corresponding grid cell using the Kullback Leibler Divergence (KLD). KLD is a well established measure for the distance between probability distributions. Here, we use the pairwise symmetric KLD, which has a low sensitivity to outliers due to the logarithm and hence will also emphasize areas that are constantly changing at a moderate level. As the comparison is done cell by cell, the result can be visualized as a grid stability map. All computations are done using the predictions Γ^q.

Fig. 8(a) shows one example of a computed wind map in the foundry hall while Fig. 8(b) shows the stability maps computed over the different measurement tours. Despite different value levels the maps show in general similar patterns. Most of the explored area can be considered stable while the bottom right corner shows quite diverse air flow patterns over time. This unstable patterns are due to the fact that the predictions at that location were made based on a few number of measurements. Overall, the stability map can serve as an overlay to the expected map (Fig. 8(a)) and serve as a confidence estimate. These results were evaluated together with a ventilation engineer at the foundry where the experiments were carried out. The conclusion

was that the produced air flow maps match the expected ventilation conditions at key locations in the foundry, for example, at the location of the furnaces.

VI. Discussion

In this paper, we presented a novel data-driven air flow modelling algorithm aimed at micro scale environments. A distinctive characteristic of our approach is the proposed $\mathcal{LT}$ extrapolation algorithm, which considers turbulent and laminar characteristics of air flow. We evaluated our algorithm using two different robotic platforms namely, a micro UAV and a ground based robot. The results show that our proposed method has a better performance, compared to standard extrapolation methods such as the Nadaraya-Watson estimator (which is used in our approach to compute the $\mathcal{L}$ component). In addition, we demonstrated that out proposed algorithm can convey useful information to mobile robots. For example, maps that show the wind variability can be used by e.g. path planning algorithms.

With what respects to ventilation characterization (i.e. the application example presented in this work), actual validation of the produced maps is rather difficult due to missing ground truth data, a problem very well known from mobile robot olfaction. Yet, on a qualitative level the results were discussed with a ventilation engineer from the foundry the data was collected in. The results correspond at key points to the assumptions about the airflow in the facility, while giving a much better idea what is happening on a finer level.

The results presented in this work are very promising, and they open up several directions for future work. First, the presented algorithms do not consider the geometry of the environment (e.g. walls, obstacles). A possible solution is to create a visibility matrix that considers the 3D model of the environment. In this way, when computing the wind flow conditions at a given cell, weights can be added to a given PDF considering whether or not there is an obstacle between the query cell and the sensing position.

Another aspect that is worth investigating is the time dependency in AFM. The presented algorithms compute a snapshot of the wind flow according to the measurements. While the histogram density modelling approach creates a transition model that aims to capture the wind variability over time, this transition model is not yet fully exploited to encode time dependency in the wind flow maps.

The models computed with the $\mathcal{LT}$ algorithm correspond to time limited snapshots of the wind distribution. Another interesting line for future work is to extend the $\mathcal{LT}$ algorithm to perform long term AFM. For practical reasons, this would require the use of data collected with stationary measurement stations that can complement the short term, spatially dense measurements that a mobile robot can acquire.

An additional direction to explore is a further investigation of the predictive capabilities of the proposed AFM algorithm. This extended numerical evaluation would consider measurements collected by stationary anemometers placed in a given target environment. Then, robot measurements can be used to create wind flow maps that in turn are used to predict the wind flow at locations where the corresponding ground-truth are provided by stationary anemometers.

References

[1] P. Beaucage, M. C. Brower, and J. Tensen, "Evaluation of four numerical wind flow models for wind resource mapping," *Wind Energy*, vol. 17, no. 2, pp. 197–208, 2014.

[2] W. H. Al-Sabban, L. F. Gonzalez, R. N. Smith, and G. F. Wyeth, "Wind-energy based path planning for electric unmanned aerial vehicles using markov decision processes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst*. Algarve, Portugal: IEEE, Feb. 2012, p. 13851250.

[3] P. P. Neumann, V. Hernandez Bennetts, A. J. Lilienthal, and M. Bartholmai, "From insects to micro air vehicles—A comparison of reactive plume tracking strategies," in *Proc. 13th Int. Conf. Intell. Auton. Syst.*, 2014, pp. 1–14.

[4] M. Reggente and A. J. Lilienthal, "The 3D-kernel dm+v/w algorithm: Using wind information in three dimensional gas distribution modelling with a mobile robot," in *Proc. IEEE Sensors*, 2010, pp. 999–1004.

[5] A. Murai, K. Yoshimoto, R. Takemura, H. Matsukura, and H. Ishida, "Robotic gas source localization assisted by active airflow generation," in *Proc. IEEE Sensors*, Nov. 2015, pp. 1–4.

[6] V. Hernandez Bennetts, A. J. Lilienthal, P. Neumann, and M. Trincavelli, "Mobile robots for localizing gas emission sources on landfill sites: Is bio-inspiration the way to go?" *Front. Neuroeng.*, vol. 4, no. 20, p. 4: 20, 2012.

[7] V. Hernandez *et al.*, "Towards occupational health improvement in foundries through dense dust and pollution monitoring using a complementary approach with mobile and stationary sensing nodes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst*. Daejeon, Korea: IEEE, 2016, pp. 131–136.

[8] S. Kato and S. Murakami, "New ventilation efficiency scales based on spatial distribution of contaminant concentration aided by numerical simulation," *ASHRAE Trans.*, vol. 94, no. 2, pp. 309–330, 1988.

[9] G. Cao *et al.*, "A review of the performance of different ventilation and airflow distribution systems in buildings," *Building Environ.*, vol. 73, pp. 171–186, 2014.

[10] T. Tomić, K. Schmidt, A. Mathers, and S. Haddadin, "The flying anemometer: Unified estimation of wind velocity from aerodynamic power and wrenches," in *Proc. IEEE/RSJ Intl. Conf. Intell. Robots Syst*. Daejeon, Korea: IEEE, 2016, pp. 1637–1644.

[11] L. Sikkel, G. de Croon, C. D. Wagter, and Q. Chu, "A novel online model-based wind estimation approach for quadrotor micro air vehicles using low cost MEMS IMUs," in *Proc. IEEE/RSJ Intl. Conf. Intell. Robots Syst.* Daejeon, Korea: IEEE, 2016, pp. 2141–2146.

[12] P. P. Neumann and M. Bartholmai, "Real-time wind estimation on a micro unmanned aerial vehicle using its inertial measurement unit," *Sensors Actuators A: Phys.*, vol. 235, pp. 300–310, 2015.

[13] G. Kowadlo and R. A. Russell, "Improving the robustness of nave physics airflow mapping, using Bayesian reasoning on a multiple hypothesis tree," *Robot. Auton. Syst.*, vol. 57, no. 67, pp. 723–737, 2009.

[14] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts, and A. J. Lilienthal, "Tell me about dynamics! Mapping velocity fields from sparse samples with Semi-Wrapped Gaussian Mixture Models," in *Proc. RSS 2016 Workshop: Geometry Beyond—Representations, Phys. Scene Understanding Robot.*, 2016.

[15] L. Rodriguez, J. A. Cobano, and A. Ollero, "Wind field estimation and identification having shear wind and discrete gusts features with a small UAS," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst*. Daejeon, Korea: IEEE, 2016, pp. 5638–5644.

[16] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. Cambridge, MA, USA: MIT Press, 2005.

[17] P. Berens, "Circstat: A MATLAB toolbox for circular statistics," *J. Stat. Softw.*, vol. 31, no. 10, 2009.

[18] E. Nadaraya, "On estimating regression," *Theory Probability Appl.*, vol. 9, no. 1, pp. 141–142, 1964.

[19] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Hoboken, NJ, USA: Wiley, 2000.

[20] P. Neumann, S. Asadi, A. J. Lilienthal, M. Bartholmai, and J. Schiller, "Micro-drone for wind vector estimation and gas distribution mapping," *IEEE Robot. Autom. Mag.*, vol. 19, no. 1, pp. 50–61, Mar. 2012.

[21] J. Saarinen, H. Andreasson, T. Stoyanov, J. Ala-Luhtala, and A. Lilienthal, "Normal distributions transform occupancy maps: Application to large-scale online 3d mapping," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 2233–2238.