



A Digital Twin Framework for Environmental Sensing with sUAS

Derek Hollenbeck¹ · YangQuan Chen¹

Received: 10 October 2020 / Accepted: 23 November 2021 / Published online: 20 April 2022
© The Author(s) 2022

Abstract

This paper proposes a digital twin (DT) framework for point source applications in environmental sensing (ES). The DT concept has become quite popular among process and manufacturing industries for improving performance and estimating remaining useful life (RUL). However, environmental behavior, such as in gas dispersion, is ever changing and hard to model in real-time. The DT framework is applied to the point source environmental monitoring problem, through the use of hybrid modeling and optimization techniques. A controlled release case study is overviewed to illustrate our proposed DT framework and several spatial interpolation techniques are explored for source estimation. Future research efforts and directions are discussed.

Keywords Remote sensing · Small unmanned aircraft systems (sUAS) · Remotely piloted aircraft system (RPAS) · Digital twin · Spatial-temporal modeling

1 Introduction

Environmental sensing and monitoring includes a variety of applications such as measuring local air quality, soil remediation, oil spills, ground water quality-tracing, etc. These applications can include both fixed and mobile sensors. For example, sensing ecosystem flux can be done with fixed sensors in the soil, above ground sensors, with mobile sensors (e.g. ground vehicles or sUAS [36]) or remote sensing techniques (such as with satellite imagery or manned aircraft [47]). Typically fixed sensor approaches work well to provide precise ecosystem flux measurements (e.g. chambers, autochambers, and eddy covariance towers

[13]). In terms of site level emissions, fixed open path lasers have been used to evaluate and compare with accepted Environmental Protection Agency (EPA) methods (e.g. dairy farms) [10]. Within the past few years, several works in the area of environmental sensing with sUAS have been done (see the comprehensive review by [18]). Typically the sensors consist of low cost chemical sensors (e.g. amperometric, metal-oxide, non-dispersive infrared, or photo-ionization detectors) or higher cost tunable diode laser absorption spectroscopy (traditional open/closed cell - TDLAS or backscatter based - sTDLAS). Other techniques like optical gas imaging, which uses thermal camera based sensing have been used with sUAS. For instance, the sTDLAS (e.g. Pergam Methane mini-G (SA3C50A)) has been used for determining spatial methane in permafrost [38]. In the oil and gas industry, the TDLAS (e.g. open path laser spectrometer - OPLS [22], or path integrated) has been used for detection [50, 51] and for quantification [27], such as in quantifying the Alberta Methane Field Challenge (AMFC) site emissions [52]. The use of fixed-wing sUAS have been used as well in both detection [15] and quantification [8].

Due to the large variety of approaches and applications of environmental sensing, the focus of this paper will be on ground-to-air emissions (as an extended versions of the conference paper [30]) with an application to gas plumes (e.g. fugitive methane). The motivation of studying fugitive methane (CH_4) gas, stems from the fact that methane has a large green house gas (GHG) warming

DH was supported in part by the National Science Foundation Research Traineeship Grant DGE - 1633722. DH and YQC were supported in part by State of California's Strategic Growth Council grant "Mobile Biochar Production For Methane Emission Reduction And Soil Amendment" (SGC#2019CCR20014, 2019–2023).

✉ YangQuan Chen
ychen53@ucmerced.edu

Derek Hollenbeck
dhollenbeck@ucmerced.edu

¹ Department of Mechanical Engineering, University of California, Merced, 5200 N. Lake Rd., Merced, CA, 95343, USA

potential relative to CO_2 and occurs from both natural and anthropogenic sources. Modeling gas plumes in general can be done using a variety of tools (e.g. deterministic and stochastic approaches) which provide information about how the plume may develop given a set of initial and boundary conditions. However, these visualization require a certain degree of model fidelity to capture realistic behavior (i.e. plume meandering and trapping). Computational fluid dynamics approaches can be applied but suffer from high computational cost. Stochastic methods have shown better computational efficiency but for a given initial condition the output is an ensemble (due to the inherent randomness) and presents difficulty in implementing continuous time. On the other hand, hybrid modeling (which encompasses a combination of stochastic and deterministic methods) can provide computational efficiency while maintaining highest possible fidelity. In practice, computational efficiency of these models would need to be suitable for deployment on a laptop. Thus, there exists a trade-off between model fidelity and run-time. This is important for applying concepts such as DT to further improve on ES work.

What is a DT? Some of the earliest work of using physical twins (or hardware twins) can be seen as early as National Aeronautics and Space Administration's (NASA) Apollo program [16]. In this case there existed at least two space vehicles or rovers at once. This allowed for the continuous testing of code, maneuvers, and conditions before sending them to the deployed version. Software can also be introduced to the hardware twin, as in the 'iron bird' which combined a physical interface of a plane cockpit with an aerodynamics simulation for training purposes. The simulation was used for a visual aide as well as to compute the meteorological conditions and resulting forces on the plane [16]. Some of the early reports of the current concept of digital twin (as digital equivalence of a physical product) showed up in a product lifecycle management course at the University of Michigan in 2003 taught by Michael Grieves. Since then a white paper was written in reference to virtual factory replication [28]. It is explained that the digital twin (DT) in this scenario contains three parts: (1) physical products in real space (2) virtual products in virtual space and (3) and connections that tie together the real and virtual products. The use cases for DT are: (1) conceptualization - visualization of the physical and virtual products, (2) comparison - compare physical product behavior with the virtual product to improve results and (3) collaboration - solutions found can be applied immediately to other factories (or systems). For ES, the environment can be full of complexity and having a DT can provide a deeper understanding between the sensors and physical parameters such as source rate, location, and atmospheric conditions. Additionally, DT's have the ability to update the

system parameters, over time, as well as give insight on performance metrics.

This paper proposes a DT framework for ES from a point source emission as an extended version of the conference paper in [30]. A controlled release case study is used to solve an optimization-based behavior matching problem of the DT framework, using the estimated concentration signal, $\hat{c}_m = f(\mathbf{x}, \mathbf{v}|\theta)$ (where \mathbf{x} is the position of the sUAS, \mathbf{v} represents the wind field and θ is the DT parameters), and the measured concentration signal, c_m , with a single sUAS measurement system. The mass balance method is applied to experimentally quantify the source emission rate. This paper extends the work in [30] by: defining DT levels and capabilities; adding additional details and figures of the DT model, approach, sensitivity of parameters, and how the observation problem is effected by sensor location (i.e. measurement operator); exploring additional spatial and spatial-temporal interpolation methods to quantify the source and compare data between the DT simulation and physical experiment; proposing the use of Mittag-Leffler function into the spatial-temporal gas distribution modeling framework. The paper is organized as follows: Section 2 describes the DT framework, Section 3 overviews the controlled release experiment, setup and behavior matching optimization, Section 4 we conclude on the simulation results, and Section 5 provides a discussion on future work to further this field. This paper proposes a DT framework for ES with a gas plume source (see Fig. 1).

2 The Proposed DT Framework

In order to discuss a framework there needs to be some agreed upon definition of DT. Over the years there have been several variants of the DT definition such as those shown in Table 1. The Mechatronics Embedded Systems and Automation Lab has since added a more general definition which is as follows, "A Digital Twin is the combination of multiple, individual, and detailed simulation models (continuous, discrete, hybrid), where its interconnection represents the dynamics of a complex system, which is updated periodically (windowed or real time) with the system information in order to reflect the system current status as well as predict its future behavior and possible faults". For ES, the model needs to be selected such that there is a trade off between fidelity and computational speed. Considering the quote from statistician George Box, "all models are wrong, but some are useful", we need the desirables of the DT model to be capable of: (D1) exhibiting relevant digital artifacts with suitable level of fidelity, (D2) short or long time-scale

Table 1 Definitions of Digital Twin [4]

Definition	Authors	Year
DT is a virtual, digital equivalent to a physical product	Grieves [28]	2003
Up-to-date representation of an actual physical asset in operation	Mathworks [5]	2019
Dynamic virtual representation of a physical object or system, usually across multiple stages of its lifecycle. It uses real-world data, simulation or machine learning models, combined with data analysis, to enable understanding, learning, and reasoning.	IBM [2]	2020
Software representations of assets and processes that are used to understand, predict, and optimize performance in order to achieve improved business outcomes.	GE Digital [6]	2019
DT is a perfect digital copy of the physical world: a digital twin. This twin would enable you to collaborate virtually, intake sensor data and simulate conditions quickly, understand what-if scenarios clearly, predict results more accurately, and output instructions to manipulate the physical world..	Deloitte [3]	2020
An integrated Multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin	NASA[26]	2012
A digital twin is a multi-faceted dynamic set of smart digital models of a system or a subsystem along with all its constituents, which accurately represent the design of a product, production process or the performance of a product or production system in operation.	Dufour et al. [23]	2018
DT are precise, virtual copies of machines or systems driven by data collected from sensors in real time, these sophisticated computer models mirror almost every facet of a product, process or service.	Tao et al. [46]	2019

characteristics evolves along with the real system given the same conditions, and (D3) can be used to solve relevant problems for the system or improving methodology [30].

The DT concept can also be broken into four different stages or levels based on their current implementation: (L1) Pre-Digital Twin, (L2) Digital Twin, (L3) Adaptive

Digital Twin, and (L4) Intelligent Digital Twin. Each stage has increasing levels of model sophistication. Aside from L1, all of the levels have a physical twin. As the DT levels go from L2 to L4 the data acquisition from the physical twin increases. Different system health and performance parameters can be extracted for predictive maintenance (from batch updates to real-time updates).

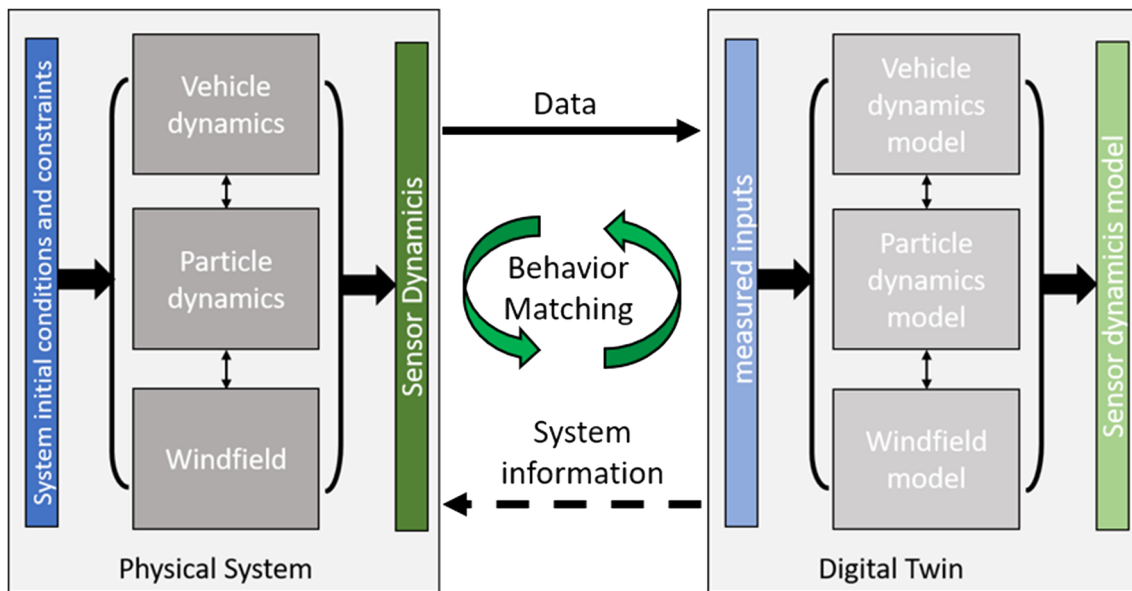


Fig. 1 Digital twin framework diagram

Additionally, these stages can start to incorporate machine learning (ML) in operator preferences as well as for the system/environment [34].

The proposed DT framework for a point source first starts on L1 with a proposed model for representing the complex system. In the case of gas concentration, the system is governed by a partial differential equation such as the advection diffusion equation,

$$\partial_t C = \nabla \cdot (D \nabla C) - \nabla \cdot (\mathbf{v}C) + \text{source}, \tag{1}$$

where the C represents the concentration of the gas, D is the diffusion coefficient, and $\mathbf{v} = [u, v]^T$ is the wind vector field. The source can be represented using Kronecker delta function $\delta_s = \delta(x - x_s)\delta(y - y_s)\delta(z - z_s)$, where $\mathbf{x}_s = [x_s, y_s, z_s]^T$ is the location of the point source. The wind field parameters can be solved generally by using the something like the incompressible Navier-Stokes equation,

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} - \nu \nabla^2 \mathbf{v} = -\nabla w + \mathbf{g}. \tag{2}$$

Here ν is the kinematic viscosity, w is the thermodynamic work, and \mathbf{g} is the gravitational acceleration. Solving these systems consecutively can be expensive and thus, some assumptions need to be applied to simplify the models. One such approach to solve this problem was proposed by Farrell et al. [25] and is summarized here for completeness. The first assumption is to break the concentration dynamics into discrete filaments (or packets of molecules) based on different length scales (i.e. large scale advection (a), intermediate scale turbulent mixing and stirring (m), and local small scale diffusion (d)). The transport of the k -th concentration molecule can then be described by, $\dot{\mathbf{x}}_k = \mathbf{v}(\mathbf{x}) = \mathbf{v}_a + \mathbf{v}_m + \mathbf{v}_d$. However, the small scale diffusion term can be absorbed into the shape of the filament, such that the filament's i -th location can be described as,

$$\dot{\mathbf{x}}_i = \mathbf{v}(\mathbf{x}) = \mathbf{v}_a + \mathbf{v}_{m_i}. \tag{3}$$

Now the only thing left to resolve is the wind vector field. One approach is to look at average wind vector such that $\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}'$, where the overbar ($\bar{\cdot}$) represents the average and the prime ($'$) denotes the deviation from the average. Substituting the decomposition of the wind vector into (2) and using the following assumptions the Navier-Stokes equation can be drastically simplified:

1. Coriolis forces, geostrophic winds, and molecular viscosity can be deemed small and neglected;
2. Measurements are conducted close to the ground where winds are relatively constant w.r.t. the altitude.

The wind field can then be represented by the following equations that resembles the viscous Burgers equation,

$$\frac{\partial \bar{u}}{\partial t} = -\bar{u} \frac{\partial \bar{u}}{\partial x} - \bar{v} \frac{\partial \bar{u}}{\partial y} + K_x \frac{\partial^2 \bar{u}}{\partial x^2} + K_y \frac{\partial^2 \bar{u}}{\partial y^2}$$

$$\frac{\partial \bar{v}}{\partial t} = -\bar{u} \frac{\partial \bar{v}}{\partial x} - \bar{v} \frac{\partial \bar{v}}{\partial y} + K_x \frac{\partial^2 \bar{v}}{\partial x^2} + K_y \frac{\partial^2 \bar{v}}{\partial y^2}. \tag{4}$$

This is realized by using the simplest K -closure method, which relates the diffusivity ($K_{x,y}$) to, $\overline{u'u'} = -K_x \frac{\partial \bar{u}}{\partial x}$ and $\overline{u'v'} = \overline{v'u'} = -\frac{1}{2}(K_x \frac{\partial \bar{v}}{\partial x} + K_y \frac{\partial \bar{u}}{\partial y})$ and similarly for the $\overline{v'v'}$, and $\overline{v'u'}$. For the simulation aspect, the wind field meandering can be produced by perturbing the flow with colored noise using a second order transfer function approach, $H(s) = Ga/(s^2 + bs + a)$ driven by white noise [25]. This colored noise is updated on the corners of the finite grid and adjacent boundary conditions are linearly interpolated (see Fig. 2). The meandering behavior (see Fig. 3) can be adjusted by scaling a, b, G and $K_{x,y}$. The interior nodes can be solved implicitly [37] or explicitly by finite differences [40].

The large scale advective term, \mathbf{v}_a , can be calculated by using a bilinear interpolation between the nearby wind vectors points or estimated using the nearest neighboring point relative to the filament position, \mathbf{x}_i . The grid size used to calculate the the wind field can remain course (large Δx and Δy) as it reflects the larger length scales of the model. Additionally the large spatial separation between points helps to increase simulation runtime. To maintain efficiency we used the nearest neighboring point for this work. To calculate the intermediate scale term, \mathbf{v}_m , we use a random process satisfying, $\dot{\omega} = A'\omega + B'\eta$ and $\mathbf{v}_m = C'\omega + D'\eta$, where $\omega \in R^n$ and η is a white noise random process with spectral density σ_η^2 . The matrices A', B', C' , and D' are appropriately sized. The resulting standard deviation grows as $G\sigma_\eta\sqrt{t}$. In the simplest 2D case $D' = \mathbf{0}$, $(B')^T = C' = [1, 1]$, and $A' = I$, where I is the identity matrix.

To calculate the small scale diffusion, \mathbf{v}_d the use of a growth term, $R(t)$, is needed. This growth term represents the standard deviation of the packet of molecules inside the discrete filament,

$$\frac{dR_i}{dt} = \frac{\gamma}{2R_i} \rightarrow R_i(t) = (R_{i0}^2 + \gamma t)^{1/2},$$

where $r_i(t) = \|\mathbf{x} - \mathbf{x}_i\|$ is the Euclidean distance from the sensor and the i -th filament, $R_{i0} = R_i(0)$ and γ

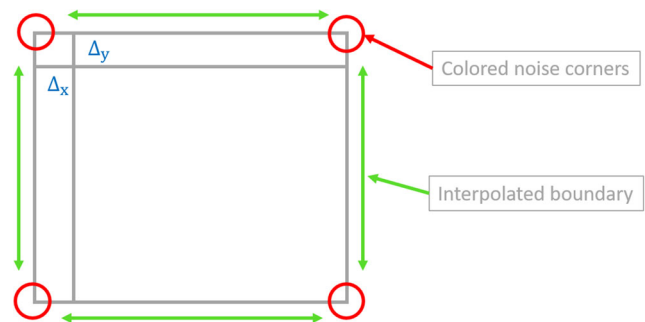


Fig. 2 Scheme for updating the colored noise in the corners interpolating the adjacent boundaries

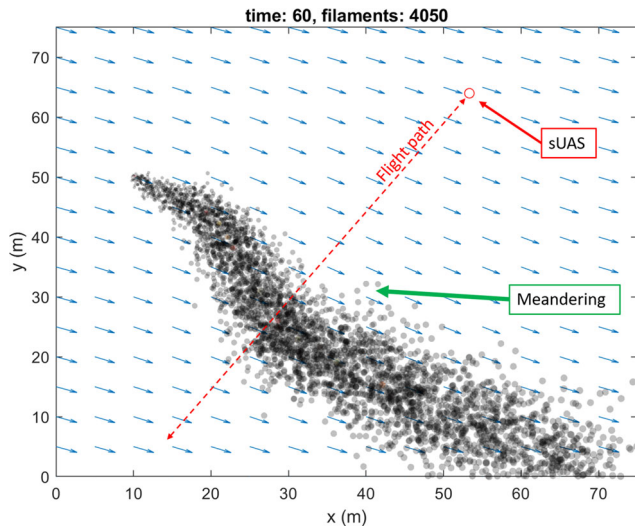


Fig. 3 MATLAB example of the digital twin with meandering

is a positive scalar. This allows for the concentration of individual molecules in a filament to be calculated by,

$$C_i(\mathbf{x}, t) = \frac{Q}{\sqrt{8\pi^3 R_i^3(t)}} \exp\left(-\frac{r_i^2(t)}{R_i^2(t)}\right), \quad (5)$$

where Q is the filament source rate in molecules per filament. The total concentration measured, $C(\mathbf{x}, t)$, can be calculated given an effective sensor area, A_e , where N represents the number of filaments inside $A_e = \pi r_e^2$.

$$C(\mathbf{x}, t) = \sum_{i=1}^N C_i(\mathbf{x}, t) \quad (6)$$

The effective area can be approximated as a circle with radius r_e . The sensor can be modeled as a low pass filtered system with thresholding, $\dot{c} = f_{BW}(C(\mathbf{x}, t) - c)$. The sensor bandwidth is f_{BW} and threshold is λ . The measured concentration, c_m would then be,

$$c_m(t) = \begin{cases} c(t), & \text{if } c(t) > \lambda \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

This current model can represent the gas plume as 2D flow. However, in most scenarios the gas needs to be represented as a 3D flow usually with building resolving capabilities. To start we will examine just rural case with no buildings. Furthermore solving the wind field for 3D case results in an increase in computation because the number of discrete points grows as $1/\Delta^3$ if $\Delta = \Delta x = \Delta y = \Delta z$.

In [30], the model was extended to 3D by utilizing a detection probability based on experimental observations in [31, 45]. The experiments showed that the detection probability varies as a function of altitude and wind speed. An example of this can be seen in Fig. 4, where the sigmoid function can be fit to the probability of detection given different wind conditions or atmospheric stability classes.

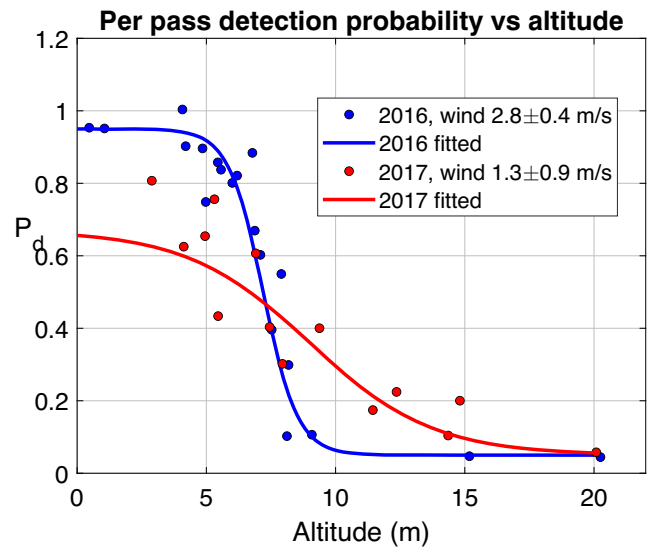


Fig. 4 Contrast in measurements of two different wind conditions stable (2016) vs unstable (2017) under controlled methane release of 5 standard cubic feet per hour (SCFH) at 30m downwind [31]

By using this knowledge one can essentially extend the 2D model to 3D. Given that sensors do not always detect the target gas, we can introduce a conditional detection probability P_d . Letting the P_d be a function of downwind distance d , altitude z , and wind velocity $|\mathbf{v}_a|$. Making the assumption that the lateral or cross wind position does not affect the probability with respect to the simulation, the relation becomes, $P_d(\delta_d|z) = f(d, z, |\mathbf{v}_a|)$. Given the nature of the mass balance method, the downwind distance d does not readily change, and therefore can be approximated as

$$P_d(\delta_d|z) \approx \frac{P_d(\delta)}{1 + e^{M(z-z_{bias})}}. \quad (8)$$

Here M is the slope at the half probability point, and z_{bias} is the altitude at which half probability occurs. From the sigmoid curves in [31] (see Fig. 4), $P_d(\delta) = [0.62, 0.9]$, $M = [0.42, 1.5]$, and $z_{bias} = [7.2, 9]$, where the average can be chosen to mirror the conditions of the experiment. However, if the gas is not in contact of the ground the distribution does not capture the vertical behavior. By using a Gaussian Normal distribution,

$$P_d(\delta_d|z) \approx P_d(\delta) \exp\left(\frac{-(z - z_{bias})^2}{2\sigma_z^2}\right), \quad (9)$$

we can alleviate this limitation. Although this only works if the source location and downwind distance are known ahead of time. Additionally, there is a relationship between the value of M , z_{bias} , and the atmospheric stability classes that needs to be further investigated. The atmospheric stability classes (A-F), first introduced by Pasquill in 1961 and later reformulated by Gifford, are typically referred to as the Pasquill-Gifford (PG) curves [32]. The PG curves

relate wind speed and solar irradiance to the dispersion coefficients that govern the spread of the plume. This is represented in the Gaussian plume model, or sometimes referred to as the infamous ‘bell curve’,

$$C(d, d_l, z, H) = \frac{Q_p}{\pi \sigma_{d_l} \sigma_z U} \exp\left(\frac{-d_l^2}{2\sigma_{d_l}^2}\right) \times \left[\exp\left(\frac{-(z-H)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+H)^2}{2\sigma_z^2}\right) \right], \tag{10}$$

where d_l is the lateral (or crosswind) distance from the plume centerline, $H = h + \Delta H$ is the effective height of the plume, h is the stack height, Q_p is the source rate in g/s and ΔH is the plume rise [7]. The plume rise term can be calculated using models such as Holland’s formula which is dependant on temperature and wind speed. However, Holland’s formula was designed for exhaust stacks and also takes into consideration the pressure, stack diameter, and temperature of the ambient air and stack. As a result if the stack velocity is small, such as with sub-surface leaks that diffuse through the ground, the plume rise term tends toward zero. Therefore other factors should be considered, such as lapse-rate, which takes into consideration how the temperature changes as a function of altitude. Furthermore, for small leaks, even surface temperature may come into effect by introducing vertical momentum through natural convective forces. Holland found that there is a small correlation between the plume rise and the temperature gradient near the ground. Briggs mentioned that the plume rise is only dependent on the temperature gradient of the air to which the plume is rising. It is also mentioned that the gradient near the ground is not a good representation of the gradient at higher altitudes [17]. Thus, this part of the DT framework will be reserved for future research.

The simulation source rate, \hat{Q}_* , given in molecules per second, and is determined by the number of filaments that are discretely released from the point source location. The particular way the filaments are released depends on two parameters, the number of filaments per puff n_f and the number of puffs per filament n_p . The filament source rate can then be defined by $Q = \hat{Q}_*/(n_f n_p)$, which is used in (5). The different combinations of these parameters can result in qualitative changes in the measured signal from the sensor. This qualitative behavior can be matched to the measured signal of the physical system and sensor (i.e. behavior matching). The behavior matching can be undertaken given an observable data set that captures these qualitative changes within the model. The key point here is observable data and thus the question becomes, how to get it? The answer depends generally on two things: sensing in the right place and sensing at the right time. Given a set of observations \mathbf{s} we want to infer the states or plume field, \mathbf{x}_i . This observation problem can be formulated as

a state estimation problem using orthogonal decomposition methods based on singular value decomposition(SVD) [24]. Essentially the problem can be represented as a \mathcal{L}_2 minimization,

$$\mathbf{s} = \mathbf{H}\mathbf{x} \approx \mathbf{H}\Phi\mathbf{v}, \tag{11}$$

$$\mathbf{v} \in \arg \min_{\mathbf{v}} \|\mathbf{s} - \mathbf{H}\Phi\tilde{\mathbf{v}}\|_2^2, \tag{12}$$

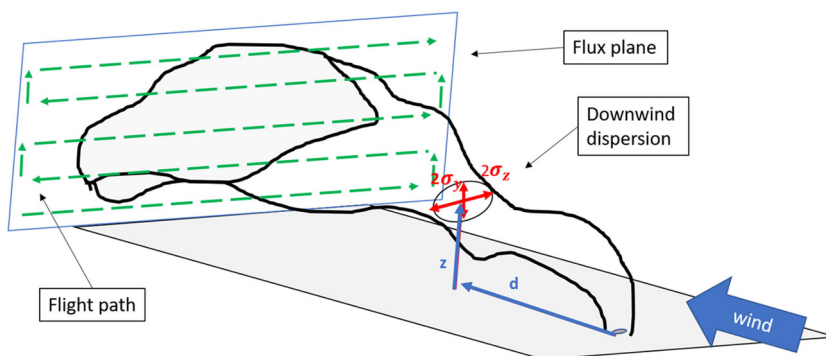
where Φ represents the orthogonal modes and \mathbf{v} the associated coefficients. This approach can yield solutions that are often ill-posed and not necessarily unique. However, it can be shown that the estimation error depends on the approximation basis Φ (orthogonal modes) as well as the measurement operator, \mathbf{H} ,

$$\|\mathbf{x}_i - \hat{\mathbf{x}}_i\| = \|(\mathbf{I} - \Phi(\mathbf{H}\Phi)^+\mathbf{H})\mathbf{x}_i\|. \tag{13}$$

The symbol $+$ represents the Moore-Penrose pseudo-inverse. The measurement operator depends entirely on the sensor location and not the orthogonal modes. Therefore, the choice of sensor location is an important task. For the single sUAS measurement system, this comes in the form of path planning and is subject to disturbances from changes in the stability of the wind. When applied to the mass balance method, path planning involves scanning back and forth at different altitudes inside a region of interest. The size of this region can incorporate twice the expected plume dispersion (deviation from the plume centerline) σ_{d_l} and σ_z , such as in [31] to ensure a likely encapsulation of the plume (see Fig. 5). This approach aims to capture rich signal data for conducting the mass balance method. However, this process takes time to complete with a single sensor system, and thus can present spatial-temporal issues when reconstructing the mass flux.

To advance the pre-Digital Twin model (L1) to a Digital Twin (L2) we need to have the physical twin exist and have the virtual system model of the physical twin. This step requires a system identification like optimization of the pre-Digital Twin model, given a dataset with observable properties and tunable parameters to minimize on. Once this minimization is complete the Digital Twin will allow for testing different path planning strategies to improve mass balance method results. The tunable parameters (see Fig. 6) in this DT model consist of: the effective radius r_e , the number of filaments per puff n_f , the number of puffs per second n_p , the intermediate scale mixing and stirring dispersion σ_η , and a correction factor c_f where the corrected simulation source rate is $(\hat{Q}_*)_c = c_f \hat{Q}_*$. The value of \hat{Q}_* is chosen to be the same as the physical source rate, such that $Q = (\hat{Q}_*)_c/(n_f n_p)$, and c_f is determined through behavior matching optimization.

Fig. 5 An example curtain flight path for mass balance calculations



3 Case Study

The experimental data used to convert the L1 Digital Twin to L2 was based off a controlled release single source field experiment in the Merced Vernal Pools and Grassland Reserve just north of the University of California, Merced campus. The source setup included a pure methane bottle connected to a pressure regulator, followed by several meters of PFTE (Polytetrafluoroethylene) tubing, where a rotameter flowmeter gauge was placed for controlling the release rate. Additional PFTE tubing was used to connect the output of the flowmeter to a 5 Gallon bucket filled with coarse rocks (≈ 1 cm in diameter). The rocks served to slow the stack velocity down and let the wind carry the methane, simulating a surface point source. The source rate was initially set to 10 SCFH, whereas after laboratory

testing the source rate was calibrated to be $Q_* = 12.9$ SCFH. The landscape of the site was relatively flat with little topology and no surface obstructions (see Fig. 7a). A sUAS, namely the DJI M210 (see Fig. 7b), was equipped with the OPLS sensor, a highly sensitive spectrometer (detection rates of 10ppb s^{-1}) and acquisition rates of 5Hz [22]. The atmospheric stability during the experiment was determined to be C and B (slightly to moderately unstable) using the PG curves [32]. The wind speed and direction was measured with a RMYoung 3D ultrasonic anemometer. The experiment consisted of 16 downwind flight paths of horizontal transects ascending from ground level till a clean air fetch was reached, referred to as a curtain (see Fig. 5).

Once the experimental data is collected, the most data information rich flights are to be selected to perform behavior matching with the L1 digital twin. In this

Fig. 6 The timeseries behavior and sensitivity with tuning parameters θ given one pass through the plume

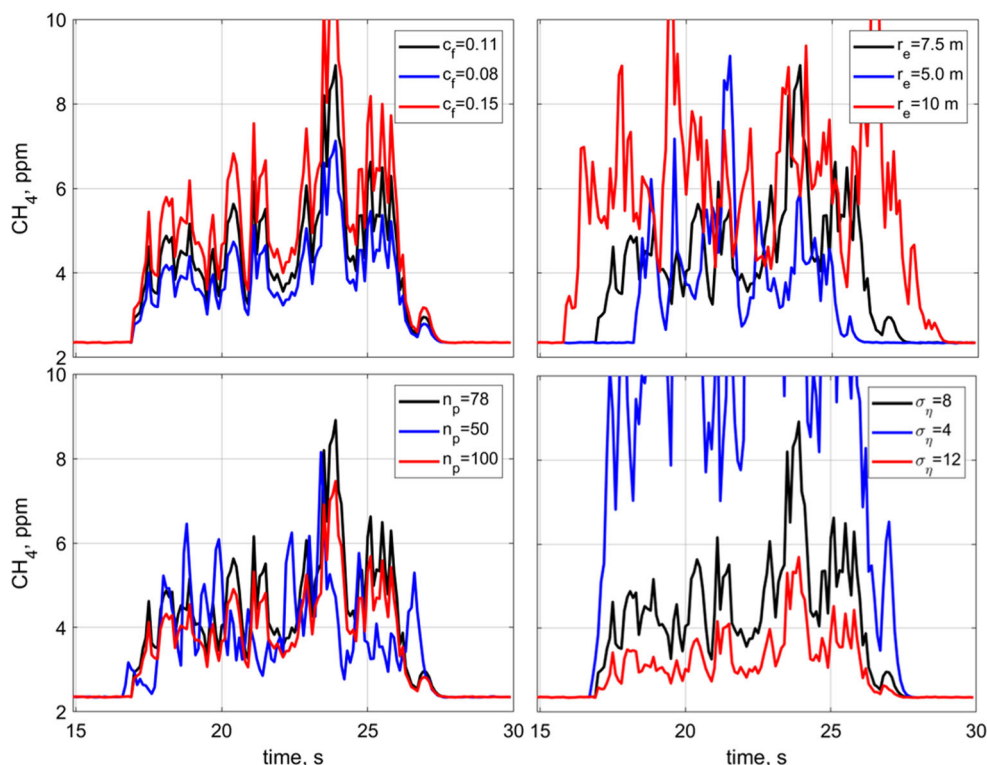
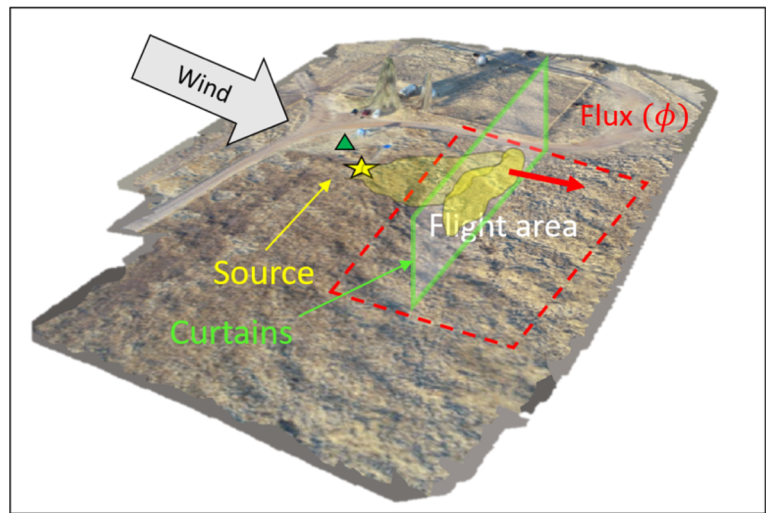


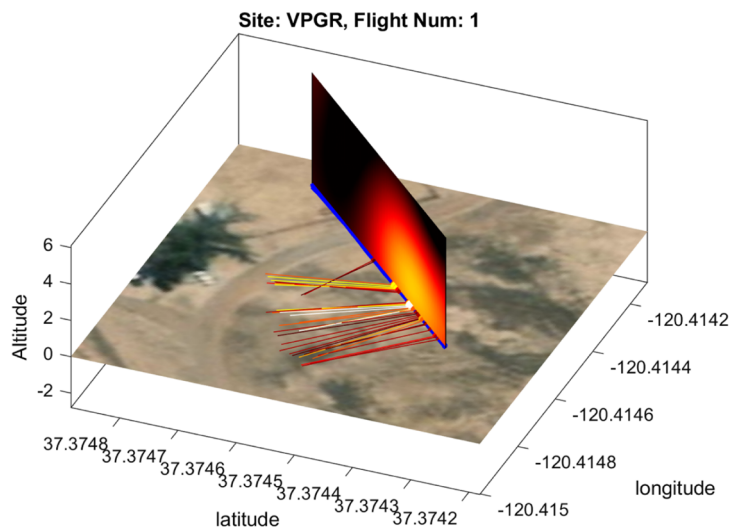
Fig. 7 (a) Experiment location at Merced Vernal Pools and Grassland Reserve. The yellow star indicates the source, the green triangle is the RMYoung anemometer, and the flight area is shown red dotted line. (b) DJI Matrice 210 (M210) equipped with a sensitive methane sniffer developed by [22] and attached to the DJI Skyport. (c) Example mass balance measurement for one curtain flight



(a)



(b)



(c)

experiment, some of the curtain flights (such as flights 13 and 14) resulted in no detected emissions and cannot be used to behavior match the model (see Table 2). After analysing the 16 flights using the mass balance method (see Fig. 7c), flight 1 showed stable wind conditions, consistent methane detection throughout the ascending transects, and good estimation of the true source rate. Thus, flight 1 was chosen for behavior matching the sensor signal of the L1 digital twin to the experimental sensor signal (see Fig. 8). The simulation source rate \hat{Q}_* was set to ($\hat{Q}_* = 2.4 \times 10^{24}$ molecules/s). Using the PG curves given the current meteorological conditions (i.e. atmospheric stability) the initial condition of σ_η was chosen to be 8 m. The wind speed and direction measured from the experiment was used as a direct input into the L1 digital twin, using the mean wind field approximation. Using the behavioral characteristics of the hyperparameters shown in Section 2 and in an attempt to reduce complexity, let $n_f = 1$, such that $n_f n_p$ has units as filaments per second. Heuristically tuning the hyperparameters, to give a best guess of $\theta = [c_f, r_e, n_p, \sigma_\eta]$, the following initial conditions were chosen: $c_f = 0.2$, $r_e = 5$ m, $n_p = 40$ puffs/sec, and $\sigma_\eta = 8$ m. It can be also seen that there is a lag present in the system (see Fig. 8) which can be attributed to model inconsistencies as well as initial conditions of the plume. Since there is currently no implementation of plume rise (future research), the average altitude where concentration detection occurred was used to determine the z_{bias} term and the vertical dispersion was set to $\sigma_z = 5$ m.

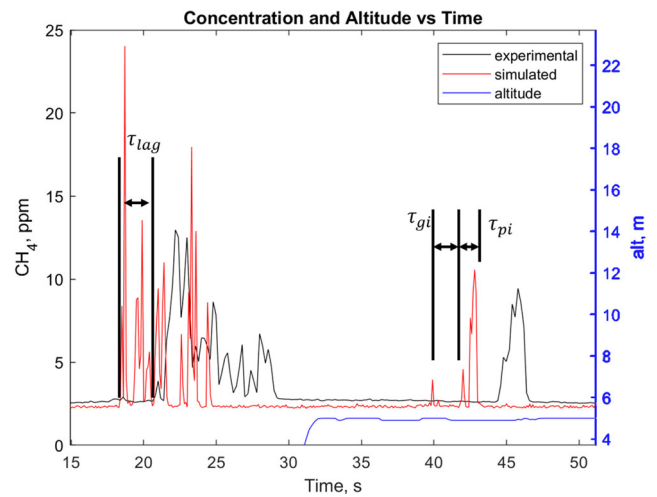


Fig. 8 The concentration time series in ppm during flight curtain 1, where $n_p = 40$ puffs/sec and the symbols τ_{gi} , τ_{pi} , and τ_{lag} represent the i -th gap length, i -th pulse length, and the time lag between the simulation and experimental data [30]

If we represent the L1 digital twin given hyperparameters θ as a function at time t ,

$$\{\hat{c}_m, \mathbf{x}_i\}_t = \{f(\mathbf{x}, \mathbf{v}|\theta)\}_t, \tag{14}$$

with the sUAS position (i.e. sensor position) \mathbf{x} and the wind speed and direction measurement using the ultrasonic anemometer (RMYoung) \mathbf{v} . The output consists of a concentration measurement $\hat{c}_m(t)$ and filament positions \mathbf{x}_i . In the physical experiment the locations of the

Table 2 Results from mass balance method for 16 curtain flights are shown below along with the average experimental wind speed, \bar{u}_e and the cos β_e between the normal vector of the flux plane

Flt	Q_*	$(\hat{Q}_*)_K$	$(\hat{Q}_*)_{IDW}^{p=2}$	$(\hat{Q}_*)_{IDW}^{p=1}$	$(\hat{Q}_*)_{KDM}$	$(\hat{Q}_*)_{KDM}^{\alpha=1}$	$(\hat{Q}_*)_{KDM}^{\alpha=0.1}$	\bar{u}_e	cos β_e
1	14.2±3.9	13.0±2.8	12.1±2.6	12.4±2.6	12.0±3.2	12.2±3.2	12.2±3.2	1.82±0.43	0.83
2	11.3±3.7	7.0±2.3	6.5±2.2	6.9±2.3	7.0±2.7	7.0±2.6	7.0±2.6	1.75±0.57	0.96
3	2.4±1.5	1.0±0.7	1.1±0.8	1.1±0.8	1.5±1.3	1.5±1.4	1.5±1.4	1.18±0.50	0.67
4	9.6±4.3	6.6±3.0	7.8±3.5	7.3±3.3	8.2±4.2	8.1±4.1	8.1±4.1	1.30±0.51	0.87
5	6.8±3.4	6.2±3.1	6.9±3.4	6.7±3.3	7.6±4.2	7.6±4.1	7.6±4.1	1.25±0.58	0.91
6	7.5±1.8	12.8±3.1	13.3±3.2	13.2±3.2	13.8±4.5	13.8±4.2	13.8±4.3	2.27±0.54	0.97
7	4.6±2.1	6.7±3.3	7.5±3.7	7.5±3.6	8.0±4.4	7.9±4.3	7.9±4.3	1.37±0.57	0.67
8	3.2±1.0	9.2±3.0	8.7±2.8	9.1±2.9	9.0±3.2	9.1±3.2	9.0±3.2	1.93±0.57	0.93
9	4.6±2.5	6.3±3.5	7.1±3.9	6.9±3.8	7.2±4.3	7.3±4.3	7.3±4.3	1.49±0.63	0.83
10	3.5±2.2	4.1±2.7	3.9±2.5	3.9±2.6	4.2±3.3	4.4±3.3	4.3±3.3	1.45±0.81	0.86
11	2.6±1.0	8.4±2.2	9.5±2.5	9.3±2.4	9.8±3.3	9.8±3.2	9.8±3.2	1.25±0.54	0.52
12	2.3±1.0	7.5±2.0	8.5±2.3	8.4±2.3	9.0±2.8	8.9±2.7	8.9±2.7	0.90±0.47	0.43
13*	0.9±0.3	15.6±5.4	16.8±5.8	15.7±5.4	17.7±6.9	17.3±6.8	17.5±6.8	2.47±0.87	0.98
14*	1.5±0.4	22.5±6.0	23.3±6.2	22.9±6.1	24.4±7.8	24.6±7.6	24.6±7.6	2.52±0.66	0.94
15	9.0±2.0	13.2±2.9	12.1±2.7	13.2±2.9	11.4±3.4	11.7±3.2	11.8±3.3	2.07±0.46	0.96
16	7.6±2.6	5.4±1.9	5.1±1.7	5.3±1.8	4.7±3.1	4.4±3.2	4.5±3.2	0.52±0.27	0.27

The flow rates (experimental Q_* and simulation \hat{Q}_*) are in SCFH, the subscripts denote the spatial interpolation method, and the superscripts are the associated parameters, where K is Kriging, IDW is inverse distance weighting, and KDM is Kernel DM+V

plume filaments are not observable given that methane gas is not visible to the human eye. The plume can potentially be observed using a large number of high quality sensors, of which, is expensive, difficult to set up, and only be inferred using spatial interpolation techniques. Thus, the only output available to optimize over is the concentration measurement, $c_m(t)$, in the least squares sense. In order to compensate for the stochastic behavior of the plume dynamics, the cost function was designed to average $n_T = 20$ trials. Unfortunately, this leads to multiple possible solutions to the minimization. Therefore, additional constraints on the hyperparameter space, Ω , and regularization can be added to constrain the magnitude of the concentration measurement such that,

$$\theta = \arg \min_{\theta \in \Omega} J(\theta), \quad \Omega = \{\theta | \theta_{min} \leq \theta \leq \theta_{max}\}, \quad (15)$$

$$J(\theta) = \frac{1}{n_T} \sum_{k=1}^{n_T} \left[0.2 \left(\sum_{i=1}^{n_t} \{c_m\}_i - \sum_{j=1}^{n_t} \{\hat{c}_m\}_j \right)^2 + 0.8(c_{max} - \hat{c}_{max})^2 \right], \quad (16)$$

where $c_{max} = \max(\mathbf{c}_m)$. The cost function weights were chosen to balance the effects of each term and were determined from trial and error.

The solution to (15) can be found using a variety of optimization methods and techniques, such as Gradient Descent or Stochastic Gradient Descent. Faster convergence can be achieved using Hessian based approaches but due to the cost of the forward model calculation and considering complexity of implementation these methods were not considered. Other gradient free approaches are more desirable for this problem such as Genetic Algorithm, Extremum Seeking Control, or Pattern (Direct) Search can be applied. With simplicity in mind, the Pattern Search Optimization was used here to behavior match the digital twin to the experimental data. After 65 iterations and over 390 cost function evaluations (7880 forward model runs) the gradient of the cost function $dJ/dt \approx 0$ (see Fig. 9). The model parameters were found to be: $c_f = 0.11$, $r_e = 7.5$ m, $n_p = 78$ puffs/s, and $\sigma_\eta = 8$ m.

To calculate and perform the mass balance method, sometimes called the box method, requires a way to measure the flux entering or leaving a control volume. Typically, upwind and downwind curtain flights are used to measure the flux. However, in the case of a controlled release the downwind curtain flight is sufficient to calculate the flux, granted the plume is encapsulated in the flight path (see Fig. 5). Using the timeseries vector $\mathbf{c}_m = [c_m(\Delta t), c_m(2\Delta t), \dots, c_m(n_t \Delta t)]^T$ with the corresponding measurement locations, $\{\mathbf{x}\}_1^{n_t}$, a concentration matrix, \mathbf{C}_m can be estimated on an n by m spatial grid through spatial interpolation methods. The flux calculation can be then

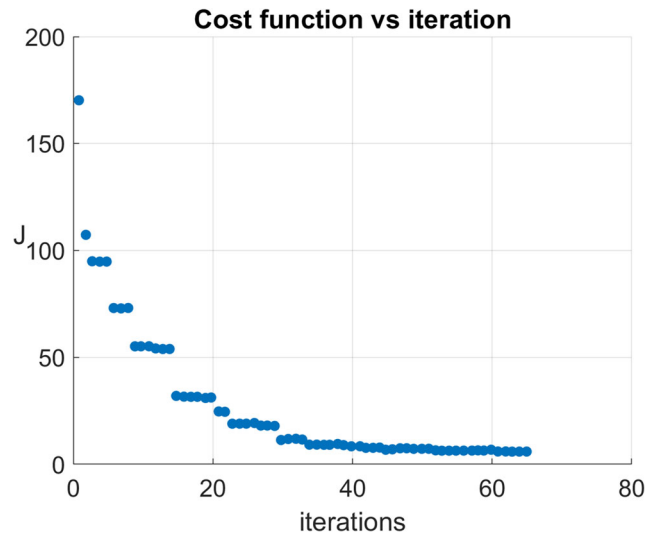


Fig. 9 Pattern search method convergence plot for behavior matching of flight 1 given the cost function, $J(\theta)$. The algorithm converged as mesh size tolerance was reached to ≈ 5.813

given as the integration over curtain area with the average wind vector multiplied by the cosine angle of the normal vector to the curtain with the enhanced concentration,

$$\phi = \int_A \bar{\mathbf{v}} \cdot \hat{\mathbf{n}}[(\mathbf{C}_m - C_b)]dA. \quad (17)$$

Here C_b represents the background concentration measurement. An example of a flux plane calculation can be shown in Fig. 7c. Several kinds of techniques can be applied to the timeseries data to generate \mathbf{C}_m , such as Ordinary Kriging, Inverse Distance Weighting (IDW) or statistical gas distribution modeling (e.g. Kernel DM+V/ W). For the case of Kriging, a linear unbiased estimator, can be solved using a

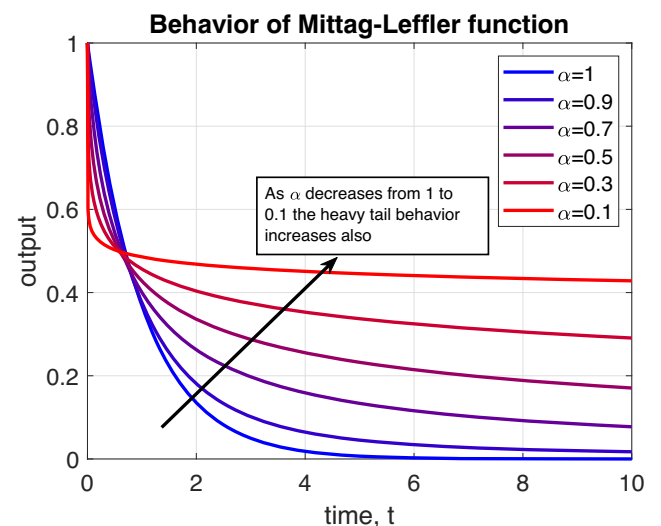


Fig. 10 Response of the single parameter Mittag-Leffler function for $\mathcal{E}_\alpha(-t^\alpha)$ as $\alpha \in [0, 1]$

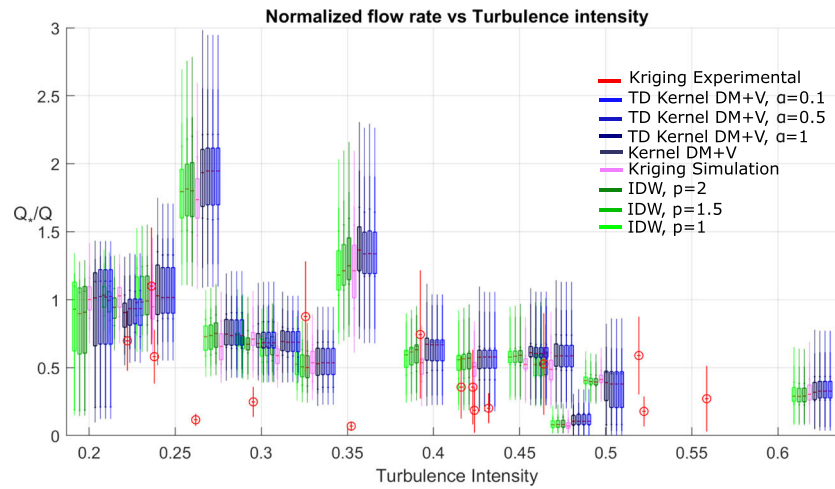


Fig. 11 Here the turbulence intensity, $T_i = \sigma_u/\bar{u}$, of the wind is plotted against the normalized flow rate, which trends toward $\hat{Q}_*/Q = 1$ as $T_i \rightarrow 0$. A boxplot of the source estimation for each flight (set of 10 simulations), given a spatial interpolation method, is shown for each T_i condition observed from the experimental data. A smaller

boxplot (shown within each boxplot) contains tails that represent the uncertainty. We can observe 1) several flights agree with experimental data 2) majority of flights underestimate the flow rate in short due to increased T_i and 3) the remaining flights that disagree are likely not capturing vertical plume behavior

variogram (or semi-variogram),

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (s(x) - s(x+h))^2. \quad (18)$$

Generally, the redundancy and closeness covariance matrix are calculated from fitting the experimental variogram to a function with respect to the spatial distance or lag, such as the exponential or Gaussian, and knowing the covariance of the measured points (referred to as the sill). This method requires second order stationarity in the spatial observations [49]. Unfortunately this is only approximately the case during stable atmospheric conditions. Methods such as IDW look at the inverse distance when determining

weights and does not consider other spatial statistics as in Kriging. There has been some efforts to improve this by looking at minimum error variance within IDW [12]. Wind information has been incorporated into Kernel DM+V/W approach in 3D [42]. Another adaptation to statistical gas distribution mapping approach was implemented in [29] by using the Gaussian plume kernel, which incorporates the wind information into the covariance function in determining the weights. The downside to this approach is that it assumes the source location. This is not ideal for determining the flux in this case. The spatial-temporal effects on modeling have been looked at as a separation between the spatial covariance and the exponential decay between temporal measurements [11, 53],

$$w_i^{(k)} = \mathcal{N}(|\{\mathbf{x}\}_i - \mathbf{x}^{(k)}|, \sigma) \varphi(t_*, t_i), \quad (19)$$

where \mathcal{N} is a Gaussian weight function, $(\cdot)^{(k)}$ represents the k -th grid point, $\{\mathbf{x}\}_i$ are the sensor measurements at time t_i and φ is given as,

$$\varphi(t_*, t_i) = e^{-a_\varphi(t_* - t_i)}. \quad (20)$$

However, when dealing with complex systems such as in turbulent environments there may be several factors that determine what value of a_φ is. If we look at the weight summation of exponentials, with elements increasing towards infinity, it can be described by fractional order dynamic behavior [35]. The Mittag-Leffler function [44] can be used to describe this behavior,

$$\mathcal{E}_{\alpha, \beta}(t) = \sum_{k=1}^{\infty} \frac{t^k}{\Gamma(\alpha k + \beta)}, \quad (21)$$

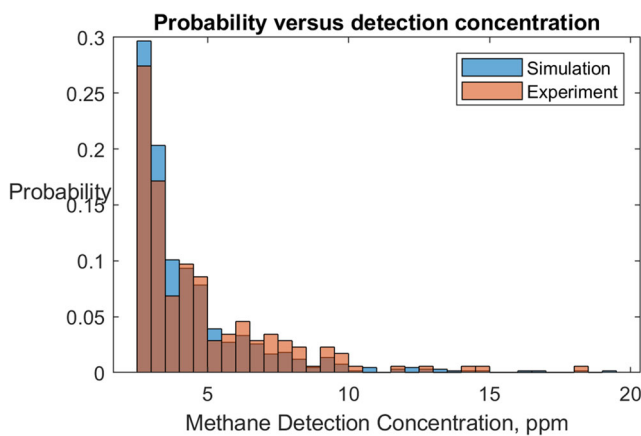


Fig. 12 The detection concentrations from the simulation and experimental sensor are compared in a probabilistic histogram. While the distributions show that the majority of indications are near background levels, they also agree with the less likely larger concentrations detected in the experiment

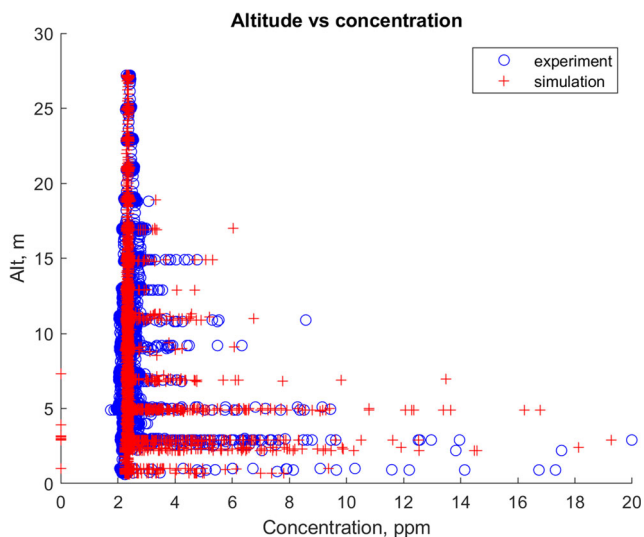


Fig. 13 The detection altitude versus measured concentrations from the simulation and experimental sensor are compared. The distributions match well with experiment and the probability of detection declines rapidly as the altitude increases from 10-15 m

such that when $\alpha = \beta = 1$, $\mathcal{E}_{1,1}(t) = e^t$. Given that $\beta = 1$ and letting α decrease we can see in Fig. 10 that the tail behavior decays much more slowly than the exponential function. By combining the approach in (19) with (21) we can apply a new weighting scheme with tuning parameter, α ,

$$w_i^{(k)} = \mathcal{N}(|x_i - x^{(k)}|, \sigma) \mathcal{E}_{\alpha,1}(-a_\varphi |t_* - t_i|^\alpha). \tag{22}$$

The value of t_* can be chosen for each grid point based on the time t_i of the closest measured point that minimizes $|\{\mathbf{x}\}_i - \mathbf{x}^{(k)}|$. For the single sensor mapping problem we test the change in sensitivity for $\alpha = 1, 0.5$, and 0.1 .

After optimization of the L2 digital twin, using the experimental position and wind measurements, each flight was simulated 10 times. The spatial interpolation was carried out using Kriging, IDW, and Kernel DM/ V to see a comparison in quantitative results across the different approaches. It can be observed that all the spatial interpolation methods tend to perform similarly. Also, for the spatial-temporal method using (22) does not show noticeable performance improvement. Future work can be done to address the optimal configurations of these methods and their hyperparameters. It can be observed that the source rate estimations fluctuate in value depending on current weather conditions and turbulence (see Table 2). Given small deviations in the wind, the sUAS curtain flight path can appropriately capture the spatial data through the mass balance plane. When the wind deviates slightly, the data we capture can sometimes be spread out over the flux plane (i.e. like a blurred image). As a result, this can lead to higher source rate estimations or to poor detection and characterization of the plume. Since the

wind direction fluctuations correlate to different stability classes it makes sense that there should be some correlation with the turbulence intensity as well (see Fig. 11). Thus, mindfulness of the current atmospheric conditions are important when planning and conducting these flights. Furthermore, the wind behavior, to some extent, is always changing in magnitude and shifting in direction. It seems this behavior is very rarely calm and stable, which may put physical limitations on the sensitivity of accurate spatial sampling with single mobile sensor systems (at least in the case for air). Comparing the experimental data to that of the digital twin one can observe qualitative similarities in the distribution of sensor detections, as seen in Fig. 12. Additionally, larger less frequent spikes at lower altitudes were observed that are consistent with what was experimentally observed (see Fig. 13).

4 Conclusion

In this work we develop a L2 DT model for environmental sensing application of a point source emission on the ground using sUAS. We give background details on spatial interpolation and modeling of sparsely measured concentration points, and also propose the use of the Mittag-Leffler function for capturing fractional order dynamics of the spatial-temporal gas distribution modeling. We compared the results from [30] with different spatial interpolation and modeling techniques. The results (shown in Fig. 11) indicate that the choice of spatial interpolation and modeling technique is less sensitive to method choice and more dependent to the DT model itself. The average mass balance calculation (excluding flights 13 and 14) of the source rate for the experiment using Kriging was determined to be 6.4 ± 2.4 SCFH, and the average mass balance calculation of the source rate for the simulation results are 7.7 ± 2.6 SCFH for Kriging, 8.1 ± 3.4 SCFH for Kernel DM+V, 7.8 ± 2.7 SCFH for IDW($p = 2$), 7.8 ± 2.7 SCFH for IDW($p = 1.5$), 7.9 ± 2.7 SCFH for IDW($p = 1$), 8.1 ± 3.3 SCFH for TD Kernel DM+V($\alpha = 1$), 8.1 ± 3.4 SCFH for TD Kernel DM+V($\alpha = 0.5$), and 8.1 ± 3.4 SCFH for TD Kernel DM+V($\alpha = 0.1$). Although the DT simulation results do not reflect the true source rate, they do reflect the experimental emission estimates, which is the goal of the DT. The qualitative behavior of the DT was also observed to be comparable to the mass balance calculations of the experiment as a function of turbulence intensity, with more accurate estimates when atmospheric conditions are stable.

This work also provided an extension to the conference paper in [30] by: (1) defining the DT and the associated levels and capabilities; (2) extending the details on the L2 DT model used in this study including the figures

showcasing the flight path, probability of detection per transect pass, and sensitivity of model parameters in the timeseries behavior matching process; (3) giving an explanation of the behavior matching problem in the context of optimization and the measurement operator H ; and (4) extending the details on the case study setup and choice of flights for behavior matching.

5 Future Efforts

Future research into making the digital twin (DT) smart is desirable as well as getting the digital twin to L3 (allows for real-time updates between the physical system and DT as well as utilize machine learning operators) and L4 (DT model utilizes adaptive techniques such as reinforcement learning between the physical system and DT for understanding system performance and features) [34]. For the DT to be smart it needs to have 5 characteristics such as Cognizant (aware of capabilities and limitations), Taskable (handle high-level, often vague, instructions), Reflective (learn from experience to improve performance), Ethical (adhere to a system of societal and legal norms), Knowledge Rich (reason over a diverse body of knowledge) [1]. Efforts in cloud computing can be leveraged with the L3 and L4 DT's to make the system smarter with tools such as Microsoft Azure and Amazon Web Services.

Just like the controlled release scenario in an open environment [30, 45] or in a real world scenario with active well sites and obstructions [50–52] there is a need to understand benchmarks for a given environment. For example, a full scale experiment in Oklahoma City such as Joint Urban 2003 test [9], where a sulfur hexafluoride (SF_6) gas was released from a known source with several sensors located throughout the city. The resulting experiment was coupled with high fidelity modeling, such as FEM3MP [19], to explore source estimation techniques [21]. However, the data sets are not readily available for others to benchmark. Other test scenarios such as the Mock Urban Setting Test (MUST) have been used in similar explorations [33]. In both of these examples the sensors are fixed and there is little to no public availability of the data sets (i.e. only shared among participants). This presents a need for a benchmark data set or framework (such as the DT) to provide a scenario based environment to test source term estimation techniques etc. In addition, open source data archiving of the data set enables reproducibility within the research field.

Single mobile sensor systems (e.g. sUAS with OPLS) have inherent spatial-temporal issues that limit gas distribution mapping usage to certain (stable) atmospheric stability classes, otherwise the accuracy can degrade and uncertainty grow. Utilizing a grid of sensors will ultimately

improve the field reconstruction efforts, as it can better separate the spatial-temporal effects and provide opportunities for sensor location optimization [48]. A group of mobile sensors can be implemented and controlled using techniques such as Central Voronoi Tessellations [20]. Furthermore, the desired sampling locations can potentially take advantage of adaptive sparse sampling methods [39] based on information metrics such as information entropy. Data driven approaches, given rich sampling locations can then be potentially applied (e.g. Dynamic Mode Decomposition [43], Compressive Sensing [14]). Even machine learning techniques may then be available such as shallow neural networks with sparse sampling [24] or physics informed neural networks (PINN) [41]. PINN's have been used in similar problems such as semantic inpainting [54]. Utilizing the power of machine learning with physical constraints combined with adaptive spatial sampling techniques could lead to more accurate reconstruction of flux planes, perhaps providing better source estimation and localization results overall.

Acknowledgements The authors would like to thank Jairo Viola for his helpful discussions on digital twins. The authors also thank the reviewers for helpful comments in improving this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Smart and Autonomous Systems (2018). <https://www.nsf.gov/pubs/2018/nsf18557/nsf18557.htm>
2. Digital twin: Helping machines tell their story (2020). <https://www.ibm.com/topics/what-is-a-digital-twin>
3. Digital twins (2020). <https://www2.deloitte.com/us/en/insights/focus/tech-trends/2020/digital-twin-applications-bridging-the-physical-and-digital.html#endnote-10>
4. Smart control engineering via digital twins (SCE via DT) (2020). <http://mechatronics.ucmerced.edu/digital-twin>
5. What is a digital twin? (2020). <https://www.mathworks.com/discovery/digital-twin.html>
6. What is a digital twin? (2020). <https://www.ge.com/digital/blog/what-digital-twin>
7. Abdel-Rahman, A.A.: On the atmospheric dispersion and Gaussian plume model. In: WWAI'08: Proceedings of the 2nd International Conference on Waste Management, Water Pollution, Air Pollution, Indoor Climate, pp. 31–39 (2008)

8. Allen, G., Hollingsworth, P., Kabbabe, K., Pitt, J.R., Mead, M.I., Illingworth, S., Roberts, G., Bourn, M., Shallcross, D.E., Percival, C.J.: The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots. *Waste Management* (2018)
9. Allwine, K.J., Flaherty, J.E.: Joint Urban 2003: Study overview and instrument locations. Tech. rep., Pacific Northwest National Lab.(PNNL), Richland, WA (United States) (2006)
10. Arndt, C., Leytem, A., Hristov, A.N., Zavala-Araiza, D., Cativiela, J., Conley, S., Daube, C., Faloona, I., Herndon, S.: Short-term methane emissions from 2 dairy farms in California estimated by different measurement techniques and US Environmental Protection Agency inventory methodology: A case study. *J. Dairy Sci.* **101**(12), 11461–11479 (2018)
11. Asadi, S., Fan, H., Bennetts, V.H., Lilienthal, A.J.: Time-dependent gas distribution modelling. *Robot. Auton. Syst.* **96**, 157–170 (2017)
12. Babak, O., Deutsch, C.V.: Statistical approach to inverse distance interpolation. *Stoch. Env. Res. Risk A.* **23**(5), 543–553 (2009)
13. Baldocchi, D., Detto, M., Sonnentag, O., Verfaillie, J., Teh, Y.A., Silver, W., Kelly, N.M.: The challenges of measuring methane fluxes and concentrations over a peatland pasture. *Agr. Forest. Meteorol.* **153**, 177–187 (2012)
14. Baraniuk, R.G.: Compressive sensing [lecture notes]. *IEEE Signal Process. Mag.* **24**(4), 118–121 (2007)
15. Barchyn, T.E., Hugenholtz, C.H., Myshak, S., Bauer, J.: A UAV-based system for detecting natural gas leaks. *Journal of Unmanned Vehicle Systems* **6**(1), 18–30 (2017)
16. Boschert, S., Rosen, R.: Digital twin—the simulation aspect. In: *Mechatronic Futures*, pp. 59–74. Springer (2016)
17. Briggs, G.A.: Plume rise: A critical survey. Tech. rep., Air Resources Atmospheric Turbulence and Diffusion Lab., Oak Ridge Tenn (1969)
18. Burgués, J., Marco, S.: Environmental chemical sensing using small drones: A review. *Science of The Total Environment*, pp 141172 (2020)
19. Chan, S.T., Leach, M.J.: A validation of FEM3MP with joint urban 2003 data. *J. Appl. Meteorol. Climatol.* **46**(12), 2127–2146 (2007)
20. Chen, Y.Q., Wang, Z., Liang, J.: Automatic dynamic flocking in mobile actuator sensor networks by central voronoi tessellations. In: *IEEE International Conference Mechatronics and Automation*, 2005, Vol. 3, pp. 1630–1635. IEEE (2005)
21. Chow, F.K., Kosović, B., Chan, S.: Source inversion for contaminant plume dispersion in urban environments using building-resolving simulations. *Journal of Applied Meteorology and Climatology* **47**(6), 1553–1572 (2008)
22. Christensen, L.E.: Miniature tunable laser spectrometer for detection of a trace gas. US Patent 9,671,332 (2017)
23. Dufour, C., Soghomonian, Z., Li, W.: Hardware-in-the-loop testing of modern on-board power systems using digital twins. In: *Proceedings of the 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, pp. 118–123. IEEE (2018)
24. Erichson, N.B., Mathelin, L., Yao, Z., Brunton, S.L., Mahoney, M.W., Kutz, J.N.: Shallow neural networks for fluid flow reconstruction with limited sensors. *Proceedings of the Royal Society A* **476**(2238), 20200097 (2020)
25. Farrell, J.A., Murlis, J., Long, X., Li, W., Cardé, R.T.: Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes. *Environmental Fluid Mechanics* **2**(1-2), 143–169 (2002)
26. Glaessgen, E., Stargel, D.: The digital twin paradigm for future NASA and US air force vehicles. In: *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, pp. 1818 (2012)
27. Golston, L., Aubut, N., Frish, M., Yang, S., Talbot, R., Gretencord, C., McSpirtitt, J., Zondlo, M.: Natural gas fugitive leak detection using an unmanned aerial vehicle: Localization and quantification of emission rate. *Atmosphere* **9**(9), 333 (2018)
28. Grieves, M.: Digital twin: manufacturing excellence through virtual factory replication. *White Paper* **1**, 1–7 (2014)
29. He, X., Steiner, J.A., Bourne, J.R., Leang, K.K.: Gaussian-based kernel for multi-agent aerial chemical-plume mapping. In: *Proceedings of the Dynamic Systems and Control Conference*, vol. 59162, pp. V003T21A004. American Society of Mechanical Engineers (2019)
30. Hollenbeck, D., Chen, Y.: Characterization of ground-to-air emissions with sUAS using a digital twin framework. In: *Proceedings of the 2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1162–1166. IEEE. <https://doi.org/10.1109/ICUAS48674.2020.9213959> (2020)
31. Hollenbeck, D., Dahabra, M., Christensen, L.E., Chen, Y.: Data quality aware flight mission design for fugitive methane sniffing using fixed wing sUAS. In: *Proceedings of the 2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 813–818. IEEE (2019)
32. Hunter, C.: A Recommended Pasquill-Gifford Stability Classification Method for Safety Basis Atmospheric Dispersion Modeling at SRS. Tech. rep., Savannah River Site (SRS) (2012)
33. Keats, A., Yee, E., Lien, F.S.: Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.* **41**(3), 465–479 (2007)
34. Madni, A.M., Madni, C.C., Lucero, S.D.: Leveraging digital twin technology in model-based systems engineering. *Systems* **7**(1), 7 (2019)
35. Magin, R.L.: Fractional calculus models of complex dynamics in biological tissues. *Comput. Math. Appl.* **59**(5), 1586–1593 (2010)
36. Manies, K., Yates, E., Christensen, L., Fladeland, M., Kolyer, R., Euskirchen, E., Waldrop, M.: Can a drone equipped with a miniature methane sensor determine methane fluxes from an Alaskan wetland? (2018)
37. Mohamed, N.: Fully implicit scheme for solving Burgers' equation based on finite difference method. *Egyptian J. Eng. Sci. Technol.* **26**, 38–44 (2018)
38. Oberle, F.K., Gibbs, A.E., Richmond, B.M., Erikson, L.H., Waldrop, M.P., Swarzenski, P.W.: Towards determining spatial methane distribution on Arctic permafrost bluffs with an unmanned aerial system. *SN Appl Sci* **1**(3), 1–9 (2019)
39. Pacheco, R.A., Rerolle, F., Lemoine, J., Hernandez, L., Meité, A., Bibaut, A., van der Laan, M., Arnold, B., Sturrock, H.J.: Finding hotspots: development of an adaptive spatial sampling approach *Scientific Reports* **10**(1). <https://doi.org/10.1038/s41598-020-67666-3> (2020)
40. Pletcher, R.H., Tannehill, J.C., Anderson, D.: *Computational Fluid Mechanics and Heat Transfer*. CRC press (2012)
41. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019)
42. Reggente, M., Lilienthal, A.J.: The 3D-Kernel DM+V/W algorithm: Using wind information in three dimensional gas distribution modelling with a mobile robot. In: *SENSORS*, 2010 IEEE, pp. 999–1004. IEEE (2010)
43. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28 (2010)

44. Shukla, A., Prajapati, J.: On a generalization of mittag-Leffler function and its properties. *J. Math. Anal. Appl.* **336**(2), 797–811 (2007)
45. Smith, B.J., John, G., Christensen, L.E., Chen, Y.: Fugitive methane leak detection using SUAS and miniature laser spectrometer payload: System, application and groundtruthing tests. In: 2017 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 369–374. IEEE (2017)
46. Tao, F., Qi, Q.: Make more digital twins. <https://doi.org/10.1038/d41586-019-02849-1> (2019)
47. Thorpe, A.K., Frankenberg, C., Thompson, D.R., Duren, R.M., Aubrey, A.D., Bue, B.D., Green, R.O., Gerilowski, K., Krings, T., Borchardt, J., et al.: Airborne DOAS retrievals of methane, carbon dioxide and water vapor concentrations at high spatial resolution: application to AVIRIS-NG (2017)
48. Ucinski, D.: Optimal measurement methods for distributed parameter system identification. CRC press (2004)
49. Wackernagel, H.: Ordinary Kriging. In: *Multivariate Geostatistics*, pp. 79–88. Springer (2003)
50. Whiticar, M., Christensen, L., Salas, C., Reece, P.: GHGMap: novel approach for aerial measurements of greenhouse gas emissions British Columbia. Geoscience BC Summary of Activities 2017: Energy, Geoscience BC, Report 2018-4 pp. 1–10 (2018)
51. Whiticar, M., Christensen, L., Salas, C., Reece, P.: Ghgmap: Detection of fugitive methane leaks from natural gas pipelines British Columbia and Alberta. Geoscience BC Summary of Activities 2018: Energy and Water, Geoscience BC, Report 2019-2 pp. 67–76 (2019)
52. Whiticar, M., Hollenbeck, D., Billwiller, B., Salas, C., Christensen, L.: Application of the bc GHGMapper™, platform for the Alberta Methane Field Challenge (AMFC). Geoscience BC Summary of Activities 2019: Energy and Water, Geoscience BC, Report 2020-02 pp. 87–102 (2020)
53. Xu, Y., Choi, J., Dass, S., Maiti, T.: Bayesian prediction and adaptive sampling algorithms for mobile sensor networks: Online environmental field reconstruction in space and time. Springer (2015)
54. Zheng, Q., Zeng, L., Karniadakis, G.E.: Physics-informed semantic inpainting: Application to geostatistical modeling. *J. Comput. Phys.* **419**, 109676 (2020)

Derek Hollenbeck is a Ph.D. student studying Mechanical Engineering at the University of California, Merced. Derek has research interests in fluid mechanics, controls, and dynamics with applications in small unmanned aerial systems (sUAS), search and localization of fugitive gases, fractional calculus, and cyber physical systems.

YangQuan Chen is a full professor at the School of Engineering, University of California, Merced. He was on the faculty of Electrical and Computer Engineering Dept. of Utah State University, Logan, Utah from 2000 to 2012 where he was a tenured associate professor and the Director for the Center for Self-Organizing and Intelligent Systems (CSOIS) from 2004–2012. Dr. Chen published many books, papers and patents. He served as the General Co-Chair for ICUAS 2014 and 2017, 2019 (International Conference on Unmanned Aircraft Systems), and on the editorial board for journals like *IEEE Trans. on Control Systems Technology*, *ISA Trans.*, *ASME Journal of Dynamics Systems, Measurement and Control*, *IFAC journals of Control Engineering Practice* as well as *IFAC Mechatronics, Fractional Calculus and Applied Analysis*, and *Nonlinear Dynamics*. He also serves as the Topical Editor-in-Chief (Field Robotics), *International Journal of Advanced Robotics Systems (IJARS)* and *Intelligent Service Robotics*. He serves as an editorial board member of *MDPI Applied Sciences, Sensors and Drones*. Most recently, he serves as the editorial board for the new Springer journal *IAI (Industrial Artificial Intelligence)*. His ISI Publons H-index is 68 with more than 20k citations, Google Scholar citation is 44.7k with H-index 92, H10-index 551. Dr. Chen was in the 2018, 2019, 2010 and 2021 Highly Cited Researchers List by Clarivate. He won Research of The Year awards from Utah State University (2012) and UC Merced (2020) respectively. In 2018, for his cumulative work in drones, Dr. Chen won “Senate Distinguished Scholarly Public Service award” that recognizes a faculty member who has energetically and creatively applied his or her professional expertise and scholarship to benefit the local, regional, national or international community.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.