Autonomous Sensing with Scientific Machine Learning for Monitoring Greenhouse Gas Emissions

Genevieve Flaspohler* [1,2] **Victoria Preston*** [1,2] **Nicholas Roy** [1] **John W. Fisher III** [1] **Adam Soule** [2]
Anna P.M. Michel [2]

1. Overview

Greenhouse gases are major contributors to global climate change; methane (CH_4), carbon dioxide (CO_2), and nitrous oxide (N_2O) are responsible for 80% of total atmospheric radiative forcing (Stocker et al., 2013). Natural sources of greenhouse gases, which include active volcanic sites, marshes, and estuaries, are the largest sources of greenhouse gas emissions; nearly half of all CO_2 emissions are from air-sea exchange alone (Solomon et al., 2007; Le Quéré et al., 2012). An increase in anthropogenic sources within the last 150 years, although volumetrically smaller than natural sources, has nevertheless overwhelmed natural sinks of these potent chemicals. We are therefore interested in characterizing natural "climate drivers" (e.g., air-sea flux) or natural processes that, as a result of increased anthropogenic emissions, are "climate driven" (e.g., melting permafrost).

To provide insight into the magnitude of gas flux from natural sources, spatio-temporal distributions of gases, and processes that drive local emissions, *in situ* instrumentation carried by a robotic platform is uniquely suited to study small- to mid-scale emission events for extended periods of time over large distances. However, collecting scientifically rich datasets of dynamic emission events requires adaptive, online planning under uncertainty and sophisticated learned representations of physical phenomena, posing challenges for even state-of-the-art machine learning and planning paradigms.

By **uniting classical scientific computing and machine learning techniques**, we aim to create novel models of environmental gas emissions that can be used in combination with **decision-making and navigation under uncertainty** to collect scientifically valuable, *in situ* data on natural greenhouse gas emissions.

*Equal contribution [1]CSAIL, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA [2]Deep Submergence Laboratory, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA. Correspondence to: Genevieve Flaspohler <gflaspo@csail.mit.edu>, Victoria Preston <vpreston@csail.mit.edu>.

Workshop on Tacking Climate Change with ML, Vancouver, BC, Canada, NeurIPS, 2019.

2. Motivating Field Studies

As motivation, we highlight two ongoing projects that examine gaseous emissions in natural environments: seasonal Arctic thawing and volcanic outgassing (field sites illustrated in Appendix A).

Seasonal Arctic Thawing In the high-Arctic, seasonal melt of landfast ice (e.g., ice over lakes, coasts), has been shown to trigger large, episodic outgassing of CO_2 and CH_4 (Zona et al., 2016; Lamarche-Gagnon et al., 2019; Karlsson et al., 2013). Quantifying emissions during the spring freshet period will help to refine emission budget models for the region. Arctic thawing is directly impacted by a warming climate, which triggers a positive feedback loop as longer thaw seasons in more of the Arctic environment lead to the release of sequestered gases (Ernakovich et al., 2014). Our work will use underwater, surface, and aerial vehicles to characterize outgassing at the land-ocean continuum.

Volcanic Outgassing Major volcanic eruptions produce gases, aerosols, and ash that influence climate processes such as cloud formation, solar reflectance, and upper ocean fertilization (Robock, 2000). Single volcanic eruptions can have a dramatic impact on global climate, as illustrated by the Tambora volcano 1816 eruption which caused a decrease in global temperatures up to $0.7°C$ and led to a wintry summer that resulted in major food shortages across the Northern hemisphere (Stothers, 1984). Even when not actively erupting, volcanic gases diffuse through the soil and escape through fissures and cracks in the crust, contributing to atmospheric concentrations. Our work will use aerial vehicles to monitor gas ratios in order to develop a predictive model of major geological events and passive emission rates.

3. Machine Learning and Autonomous Sensors for Emission Models

We propose the following unified framework for autonomous environmental monitoring: 1) using gas diffusion models from physics-based scientific modeling to plan maximally informative data collection paths, and 2) using the collected data to improve the science models via data-driven scientific machine learning. These two approaches should reinforce each other in a feedback loop; initial models can be used to select promising locations for collecting environmental data, and machine learning techniques can use these data to improve the emission models. This framework is illustrated in Fig. 1.

The combination of these approaches is particularly compelling: science models establish a rich structure for the data, but are brittle under noisy, uncertain measurements and difficult to update with streaming data; machine learning is incredibly flexible but suffers from low explainability and data efficiency, which is valuable in scientific applications.

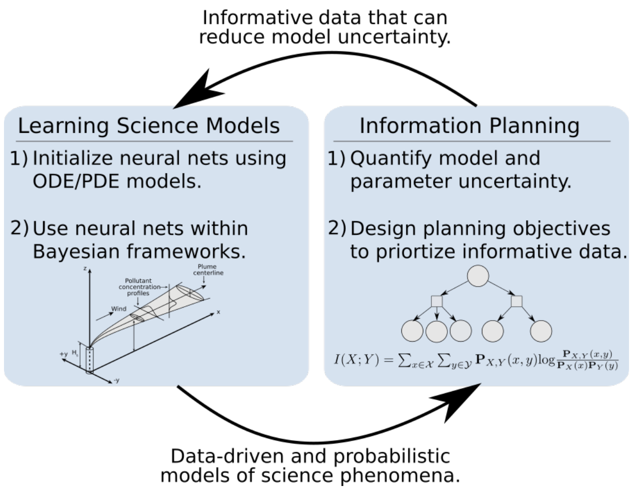

Figure 1. The proposed sensing and learning framework. We propose potential techniques and technical issues for the problems of learning science-driven models and planning informative data-collection missions.

Learning Gas-Emission Models Our approach will leverage recent work for improving scientific models using machine learning techniques (Han et al., 2018; Raissi & Karniadakis, 2018). Our scientific model for emission modeling will be based on a simple and widely-used physics-based model of gas diffusion - the Gaussian plume model. Gaussian plumes are an approximate solution to the advection-diffusion differential equations (Stockie, 2011):

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} + \frac{\partial}{\partial y}\left\{K_y\frac{\partial C}{\partial y}\right\} + \frac{\partial}{\partial z}\left\{K_z\frac{\partial C}{\partial z}\right\}, \quad (1)$$

where C is the gas concentration, $\{x, y, z\}$ describe distance from the source in $\mathbb{R}^3$, and $\theta = \{u, K_y, K_z\}$ are parameters.

In Fig. 1, we suggest two potential techniques for combining machine learning and scientific models for gas emissions. The first is directly initializing a neural network with solutions to the advection-diffusion differential equations (Eq. 1), thereby enforcing initial structure in the network. The advantage of this method could be to reduce the amount of training data necessary for training the neural network (Raissi & Karniadakis, 2018). The second technique would leverage the ability for neural nets to act as nonlinear measurement models (Johnson et al., 2016). We would subsequently use these nets within a Bayesian framework to estimate the parameters of the science model online with streaming data and uncertainty quantification.

Adaptive Data Collection Given the objective of learning better scientific models of gas emissions, a robotic agent should collect data in uncertain or inaccurately modeled regions in order to improve the existing model. Previous work in information planning for environmental applications has shown that selecting sensing locations according to **information-theoretic objectives** allows for accurate and data-efficient environmental models (Flaspohler et al., 2019; Hitz et al., 2017; Singh et al., 2010). Estimating information theoretic objectives requires accurate estimates of model uncertainty. This constraint requires a learned gas-emission model that allows for accurate uncertainty quantification.

Selecting optimal sampling locations to maximize an information-theoretic objective function is a difficult planning problem. The robot must trade-off between taking exploratory actions to learn the model, and exploitative actions to take advantage of it's knowledge. This can be approached as a reinforcement learning (Stadie et al., 2015), POMDP (Arora et al., 2017), or bandit problem (Krause et al., 2008). However, each of these approaches have limitations in large, continuous, dynamic and partially observable environments. We plan to first develop a simulation framework for evaluating proposed planners and then to consider planning adaptive and efficient explore-exploit behaviors.

4. Impact

Natural greenhouse gas emissions from environmental processes play a significant role in global climate, but are challenging to monitor due to their scale and variability. As anthropogenic forcing begins to change the climate system, these natural emission sources may enter into positive or negative feedback loops that affect the climate system in ways that are important to model and predict. We propose a sensing and learning framework that leverages science models derived from first principles and machine learning techniques to collect informative observations that adapt the models to complex, real phenomena. Several key technical challenges include integrating few-shot and online learning into scientific machine learning models, learning uncertainty metrics, and performing near real-time decision-making in large, continuous, high-dimensional environments. Within the climate science community, our work will produce valuable datasets of *in situ* measurements and models of natural greenhouse gas emissions.

References

Arora, A., Furlong, P. M., Fitch, R., Sukkarieh, S., and Fong, T. Multi-modal active perception for information gathering in science missions. In *Proc. Int. Symp. Auton. Robots*, pp. 1–27, 2017.

Ernakovich, J. G., Hopping, K. A., Berdanier, A. B., Simpson, R. T., Kachergis, E. J., Steltzer, H., and Wallenstein, M. D. Predicted responses of arctic and alpine ecosystems to altered seasonality under climate change. *Global Change Biology*, 20(10):3256–3269, 2014.

Flaspohler, G., Preston, V., Michel, A. P., Girdhar, Y., and Roy, N. Information-guided robotic maximum seek-and-sample in partially observable continuous environments. *IEEE Robotics and Automation Letters*, 4(4):3782–3789, 2019.

Han, J., Jentzen, A., and Weinan, E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.

Hitz, G., Galceran, E., Garneau, M.-È., Pomerleau, F., and Siegwart, R. Adaptive continuous-space informative path planning for online environmental monitoring. *J. Field Robot.*, 34(8):1427–1449, 2017. ISSN 15564959.

Johnson, M., Duvenaud, D. K., Wiltschko, A., Adams, R. P., and Datta, S. R. Composing graphical models with neural networks for structured representations and fast inference. In *Advances in neural information processing systems*, pp. 2946–2954, 2016.

Karlsson, J., Giesler, R., Persson, J., and Lundin, E. High emission of carbon dioxide and methane during ice thaw in high latitude lakes. *Geophys. Res. Lett.*, 40(6):1123–1127, 2013.

Krause, A., Singh, A., and Guestrin, C. Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. *J. Mach. Learn. Res.*, 9:235–284, 2008.

Lamarche-Gagnon, G., Wadham, J. L., Lollar, B. S., Arndt, S., Fietzek, P., Beaton, A. D., Tedstone, A. J., Telling, J., Bagshaw, E. A., Hawkings, J. R., et al. Greenland melt drives continuous export of methane from the ice-sheet bed. *Nature*, 565(7737):73, 2019.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., Van der Werf, G., Ahlström, A., et al. The global carbon budget 1959–2011. *Earth Syst. Sci. Data*, 5(2):1107–1157, 2012.

Raissi, M. and Karniadakis, G. E. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357:125–141, 2018.

Robock, A. Volcanic eruptions and climate. *Reviews of geophysics*, 38(2):191–219, 2000.

Singh, A., Ramos, F., Whyte, H. D., and Kaiser, W. J. Modeling and decision making in spatio-temporal processes for environmental surveillance. In *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 5490–5497, 2010.

Solomon, S., Qin, D., Manning, M., Averyt, K., and Marquis, M. *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, volume 4. Cambridge University Press, 2007.

Stadie, B. C., Levine, S., and Abbeel, P. Incentivizing exploration in reinforcement learning with deep predictive models. *arXiv preprint arXiv:1507.00814*, 2015.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M., et al. Climate change 2013: The physical science basis, 2013.

Stockie, J. M. The mathematics of atmospheric dispersion modeling. *Siam Review*, 53(2):349–372, 2011.

Stothers, R. B. The great tambora eruption in 1815 and its aftermath. *Science*, 224(4654):1191–1198, 1984.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., et al. Cold season emissions dominate the Arctic tundra methane budget. *Proc. Natl. Acad. Sci. U.S.A.*, 113(1):40–45, 2016.

Figure 2. **Application Environments:** Shown are the two environments in which we intend to deploy robotic vehicles to gather observations of greenhouse gas emissions. (Left) The image from Cambridge Bay shows the surface vessel used during preliminary field work in summer 2018. (Right) Venting sites at Yellowstone National Park, where a small gas sensor was incorporated into a drone vehicle and tested. (Center) Map Imagery © 2019 TerraMetrics

A. Field work

Preliminary field work has been conducted for both motivating applications (illustrated in Fig. 2). In summer 2018, an unmanned surface vehicle with onboard gas extraction equipment was deployed in Cambridge Bay, Nunavut, Canada to examine the content of methane and carbon dioxide in a local river system immediately after the seasonal ice cover began to retreat. Preliminary results indicated significantly elevated concentrations of both gases in the water column, which ventilated rapidly into the atmosphere. Future trials at this location in summers 2020 and 2021 will focus more on identifying the source of these gases and quantifying the rate of ventilation in the atmosphere.

In spring and summer 2019, a small gas sensor was tested at Yellowstone National Park. This work was generally to test the feasibility of detecting multiple gas species in the atmosphere around active volcanic vents. Further trials with the sensor package on a drone, and integrating the sensor measurements into an autonomy framework, are planned starting in early 2020.