Otto
Lamminpää
Jet Propulsion Laboratory, California Institute of Technology
Sean Crowell, LumenUs Scientific LLC
Will Keely, LumenUs Scientific LLC, Jet Propulsion Laboratory, California Institute of Technology
Greg McGarragh, Colorado State University Cooperative Institute for Research in the Atmosphere
Chris O'Dell, Colorado State University Cooperative Institute for Research in the Atmosphere
Steffen Mauceri, Jet Propulsion Laboratory, California Institute of Technology
Poster
Remote sensing of atmospheric carbon dioxide (CO2) carried out by NASAs Orbiting Carbon Observatory-2 (OCO-2) satellite mission and the related Uncertainty Quantification (UQ) effort involves repeated evaluations of a state-of-the-art atmospheric physics model. The retrieval, or solving an inverse problem, requires substantial computational resources. Machine Learning (ML) methods show tremendous potential for this remote sensing retrieval problem due to (1) the potential for learning new relationships in model parameters, state vector elements, and measured quantities, without implementing explicit physics parameterizations as hypotheses, which is time-consuming and limited by our current understanding, and (2) the incredible speed gains realized by a fully trained and validated ML model [Mishra and Molinaro, 2021, David et al., 2021]. The training process can be computationally costly, but the trained model predicts state variables and uncertainties as a model evaluation in seconds. By comparison, the operational OCO-2 retrieval takes 3-5 minutes per sounding, most of which is spent on the complex radiative transfer (RT) calculations. We apply modern ML techniques to the trace gas retrieval problem by investigating two paths. The first pathway involves using a fast Gaussian Process (GP) emulator of the operational RT to create a surrogate retrieval for the OCO-2 Level 2 Full Physics (L2FP) processing pipeline. The second pathway will develop "direct" retrieval approaches that link atmospheric and surface parameters with radiances. This process starts from the work of Breon et al. [2022], but moves beyond the classical neural network (NN) approach to incorporate more modern techniques. This will be explored by implementing methods like Deep Ensembles and Mixture Density Networks.

In this work, we propose and implement a statistical emulator to speed up the computations in the OCO-2 L2FP physics model. Our approach is based on Gaussian Process (GP) Regression, leveraging recent research on Kernel Flows and Cross-Validation to efficiently learn the kernel function in the GP. The resulting retrieval algorithm, which we call L2GP, takes advantage of the fast and accurate GP RT and analytic Jacobians in an optimal estimation (OE) framework to L2FP. L2GP provides uncertainty estimates and averaging kernels, which allow for assimilation of L2GP predicted XCO2 into inversion models [Crowell et al., 2019, Peiro et al., 2022]. We demonstrate our method by replicating the behavior of OCO-2 forward model within measurement error precision, and further show that in control cases, our method reproduces the CO2 retrieval performance of OCO-2 setup with orders of magnitude faster computational time. Our proposed approach is not only fast but also highly accurate (its relative error is less than 1%).
Meeting homepage
IWGGMS-20 Workshop
Download to PDF