Estimating GHG Emissions by emulating atmospheric transport using machine learning

Nikhil Dadheech[1,*], Tai-Long He[1], and Alexander J. Turner[1]
[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, *nd349@uw.edu

1. Abstract

Previous work has shown that point sources for GHGs dominate their emission budget. Due to their localized nature, studying these point sources necessitates dense measurements. Fortunately, there has been a proliferation of dense observing systems for GHGs over the past decade. Estimating GHG emission fluxes through these observations require computing source-receptor relationship (also known as "footprint"). This relationship is constructed using full physics-based atmospheric transport models which are oftentimes computationally expensive as well as storage intensive as the number of measurements increases. Here, we present a deep-learning-based emulator of footprints at kilometer-scale. The model consists of encoder-decoder based U-net model which requires meteorological parameters as inputs, and it predicts footprint for an observation in near-real-time. We compared the accuracy of the emulator by one-on-one comparison with footprints generated by STILT model. The emulator footprints are further used in estimating GHG emission fluxes using inverse modelling. We replicated a previous case study from Turner et. al., 2020 using footprints from the emulator.

2. Motivation

$$\underbrace{y}_{\text{Measurements}} = \underbrace{H}_{\text{Footprints}}\underbrace{x}_{\text{Emissions}} + \underbrace{\in}_{\text{Error}}$$

Figure 1: Relating observations to emissions

$$\underbrace{\hat{x}}_{\text{Posterior}} = \underbrace{x_a}_{\text{Prior}} + (\underbrace{HB}_{\text{Prior covariance matrix}})^T(HBH^T + \underbrace{R}_{\text{Observational covariance matrix}})^{-1}(y - Hx_a)$$

Atmospheric transport relates observations to emissions using first equation. We can estimate emissions based on the second equation which uses prior and observations error covariance matrices.

3. Training data

Generated with STILT model

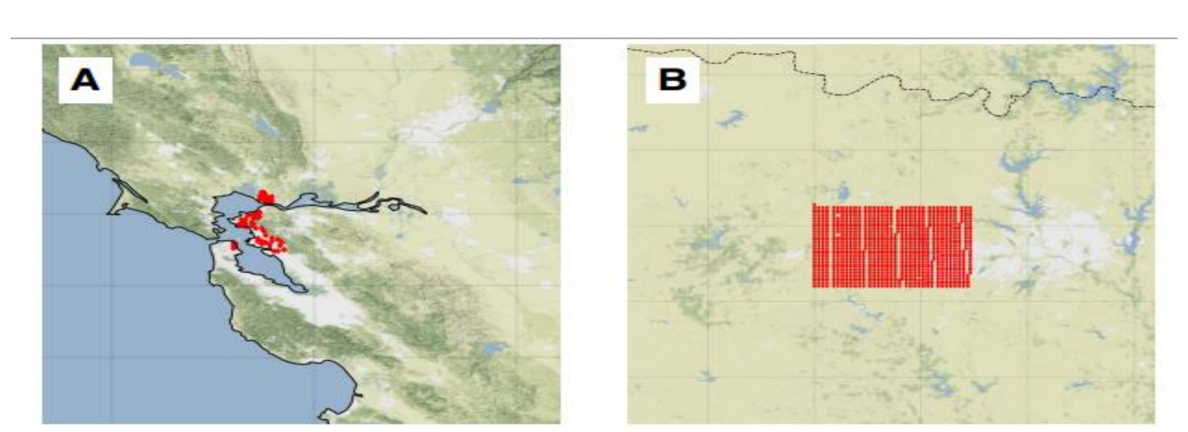

*Figure 2: Regions used in this study for development and evaluation of the emulator. **(A)** is the San Francisco Bay Area, CA and **(B)** is the Barnett Shale, TX region. The red dots represents locations of receptors for which footprints were constructed.*

4. Development of deep learning emulator

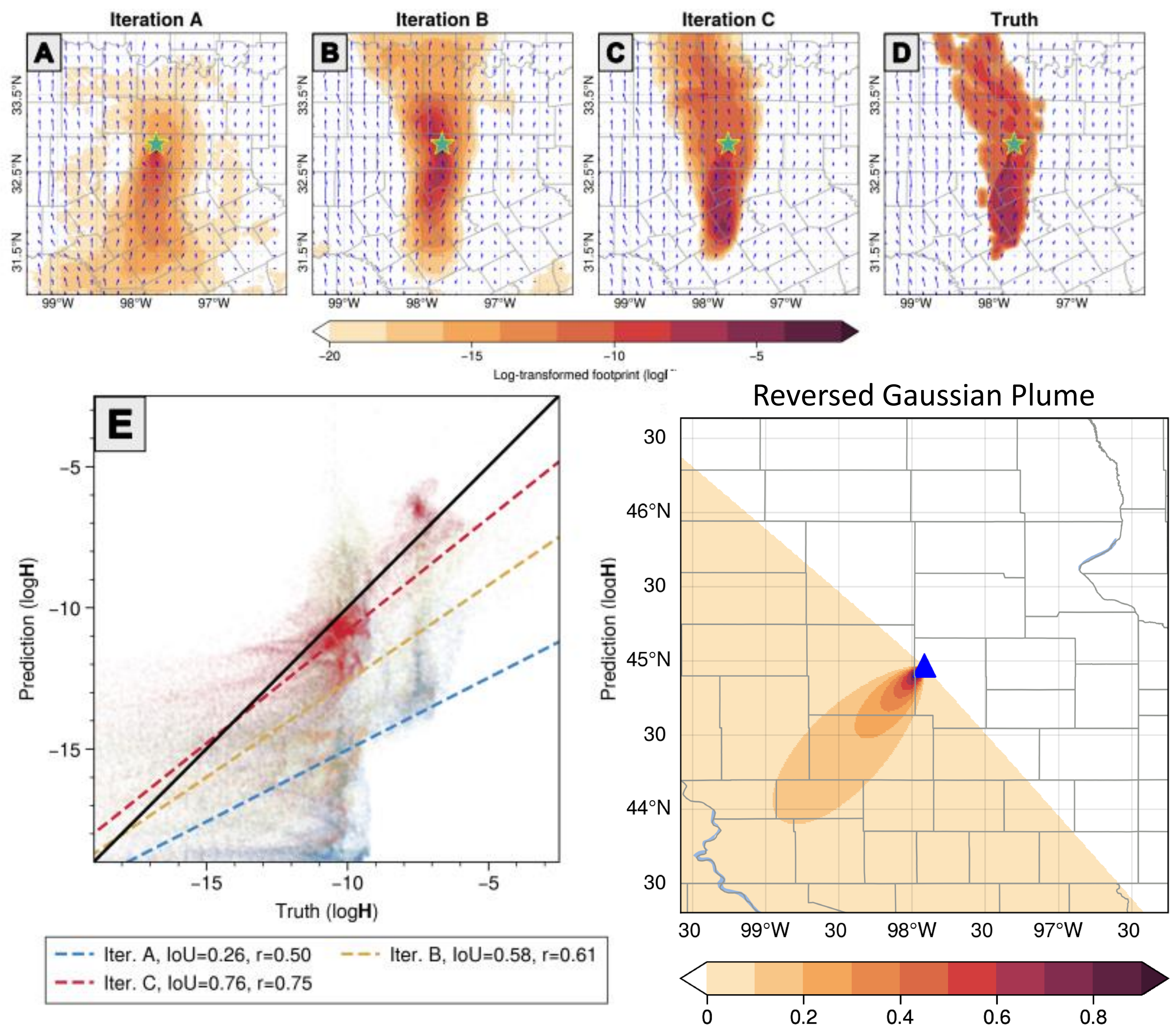

Figure 3: Evolution of emulator for three different instances A, B, and C during training. Figure D shows the "truth" (STILT footprint). Figure E shows the comparison between predictions and truth for A, B, and C.

Table 1: Comparison of computational & storage expenses for constructing footprints for 1000 measurements using STILT vs emulator

Simulation Case	#CPU cores	#Servers	Hours
Sequentially (STILT)	1	1	1000
Parallel on 1 node (STILT)	32	1	31.25
Parallel on 10 nodes (STILT)	32	10	3.125
Emulator using GPU	1-4	1	0.08

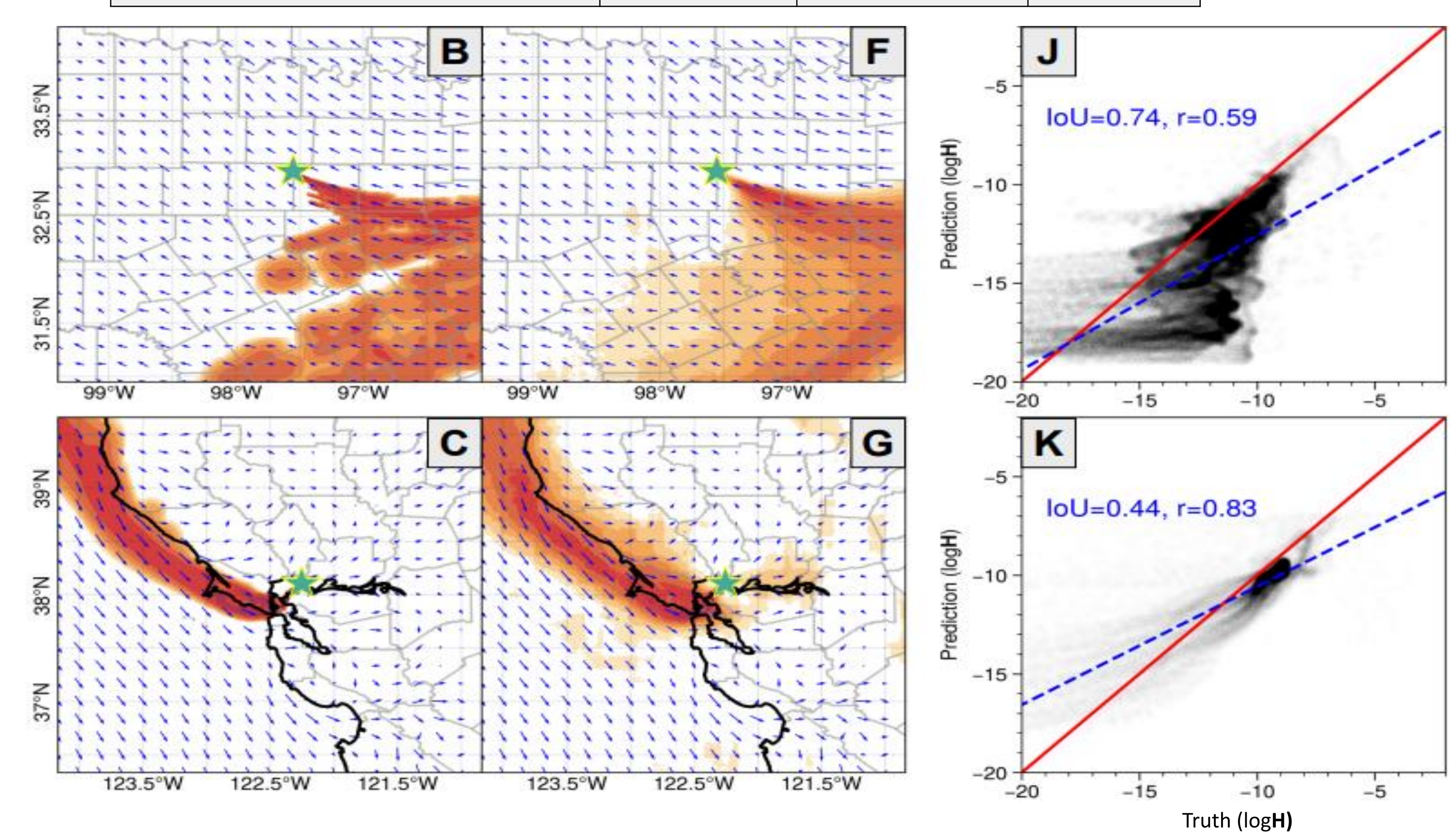

Figure 4: Comparison of footprints generated with emulator against STILT footprints for validation data (independent from the model training dataset).

5. Estimating emission fluxes using emulator

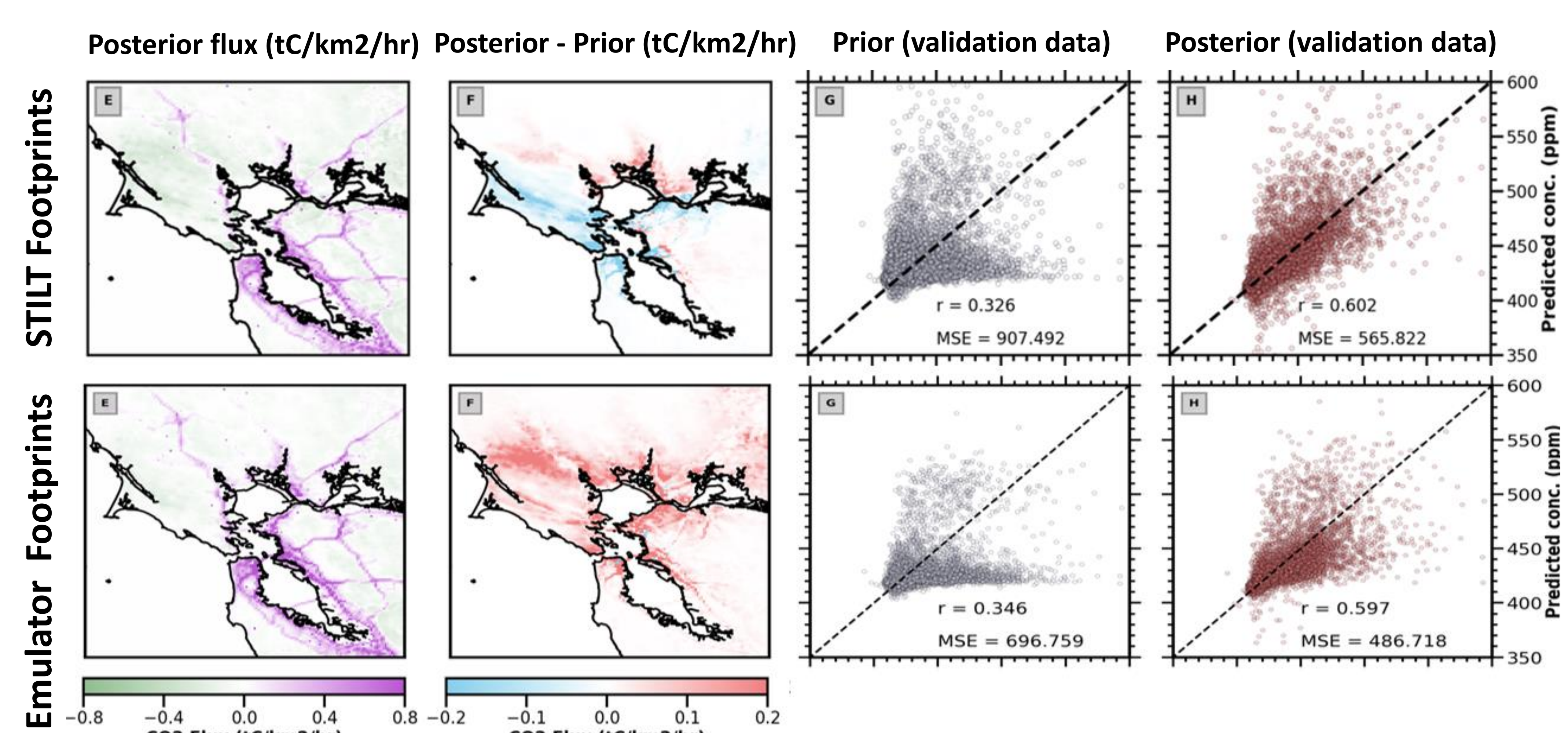

Figure 5: CO2 flux inversion in San Francisco Bay Area (replicating Turner et. al., 2020). This figure shows posterior emission fluxes, their difference with prior from Feb – April 2020. The next two columns are comparisons between predicted and observed concentration from the data independent from flux inversion.

6. Conclusion & future work

- We have developed a deep learning emulator which is trained on footprints simulated using full-physics based STILT model.
- Independent validation data and emission flux comparison shows that the magnitude and spatial patterns of footprints are well-predicted
- Further, we are working on increasing the generalizability of the emulator over the CONUS
- Increasing the scope of emulator to the footprints for the satellite obs.
- Use the emulator to compute footprints on the fly for dense observations and near-real-time emission monitoring.

7. Acknowledgements

This work is funded by NASA FINESST Award #80NSSC22K1557, NASA ECF Grant #80NSSC21K1808, Environmental Defense Fund, and the Integral Environmental Big Data Fund (UW)