Source Identification and Field Reconstruction of Advection-Diffusion Process from Sparse Sensor Measurements

Arka Daw*
Virginia Tech
Blacksburg, United States.
darka@vt.edu

Kyongmin Yeo
IBM T.J. Watson Research Center
Yorktown Heights, United States.
kyeo@us.ibm.com

Anuj Karpatne
Virginia Tech
Blacksburg, United States.
karpatne@vt.edu

Levente Klein
IBM T.J. Watson Research Center
Yorktown Heights, United States.
kleinl@us.ibm.com

Abstract

Inferring the source information of greenhouse gases, such as methane, from spatially sparse sensor observations is an essential element in mitigating climate change. While it is well understood that the complex behavior of the atmospheric dispersion of such pollutants is governed by the Advection-Diffusion equation, it is difficult to directly apply the governing equations to identify the source information because of the spatially sparse observations, i.e., the pollution concentration is known only at the sensor locations. Here, we develop a multi-task learning framework that can provide high-fidelity reconstruction of the concentration field and identify emission characteristics of the pollution sources such as their location, emission strength, etc. from sparse sensor observations. We demonstrate that our proposed framework is able to achieve accurate reconstruction of the methane concentrations from sparse sensor measurements as well as precisely pin-point the location and emission strength of these pollution sources.

1 Introduction

Methane is one of the potent greenhouse gasses [4, 3, 10] that is emitted into the atmosphere through leakages in natural gas systems, raising livestocks, or via natural sources such as wetlands. These methane emissions caused by human activities have further been identified as a major contributor to climate change [1, 9, 8]. Thus, inferring the location and emission strength of pollution sources such as methane leaks are both essential in monitoring the air-quality as well as mitigating the climate change. However, one of the major challenges in localization of these emission sources is the unavailability of high-resolution methane concentration maps, and these individual pollution sources are to be estimated from the limited concentration measurements from a sparse sensor network. Atmospheric inverse models that aim to either reconstruct the concentration field [7] or identify pollution source information [5] of airborne pollutants have been extensively studied in statistics and applied mathematics communities. However, due to the ill-posed nature of the inverse problem, it still remains as a challenging topic. Recently, potential of deep learning approaches for the inverse problem has been demonstrated [2].

*Work done during internship at IBM T.J.Watson Research Center.

Machine Learning and the Physical Sciences workshop, NeurIPS 2022.

We develop a multi-task learning framework that can provide a high-fidelity reconstruction of the spatio-temporal concentration field and identify the emission characteristics of the pollution sources such as their location, and emission strengths from a time series of sensor measurements. We propose a novel 3D diffusive-masked convolution to gradually propagate the information from the sensor locations to the unobserved field using a diffusion process. We demonstrate that our multi-task model is able to achieve accurate reconstruction of the pollution concentrations from sparse sensor measurements as well as precisely pin-point the location and emission strength of these pollution sources.

2 Background and Problem Setup

Let there be N_s number of potential emission sources and N_o number of sensors. We assume that the atmospheric dispersion is determined by the following Advection-Diffusion Equation:

$$\frac{\partial \phi}{\partial t} + \mathbf{u}\nabla\phi - K\nabla^2\phi = \sum_{i=1}^{N_s} q_i(\mathbf{x}, t); \;\; \mathbf{x} \in \mathbb{R}^2, t \in \mathbb{R}^+ \tag{1}$$

where ϕ is the concentration, $\mathbf{u}$ is the wind velocity field, K is the turbulent diffusivity, and $q_i(x,t)$ is the emission strength of the $i-th$ source.

Also, let the time-series of measurements from the N_o sensors be $\Phi = \{\Phi_1, \Phi_2, ..., \Phi_{N_o}\}$, where Φ_i denotes the (time varying) sensor measurement of the i-th sensor, i.e., $\Phi_i = [\Phi_i^1, \Phi_i^2, ..., \Phi_i^{N_t}]$, with N_t denoting the number of time steps. The location of the i-th sensors are denoted by $\mathbf{x}_i^\Phi = (x_i^\Phi, y_i^\Phi)$. Also, note that the wind velocity field $\mathbf{u}$ is available on a $N_x \times N_y$ uniform-grid, such that N_x and N_y are the number of discretizations of the input domain in x and y directions respectively. We further assume that our pollution sources are point sources, and can be realized as $q_i(x,t) = c_i\delta(x - x_i^s, y - y_i^s)$, where c_i is the emission strength of the i-th emission source, (x_i^s, y_i^s) is the location of the source, and δ is the Dirac delta function.

Problem Statement: Given the time series measurements of the sensor networks $\Phi \in \mathbb{R}^{N_o \times N_t}$, their locations $\mathbf{x}^\Phi = \{\mathbf{x}_1^\Phi, \mathbf{x}_2^\Phi, ..., \mathbf{x}_{N_o}^\Phi\}$, and the wind velocity field $\mathbf{u}$, the spatio-temporal field reconstruction problem can be formulated as estimating the methane concentration ϕ on the regular grid $N_t \times N_x \times N_y$. Additionally, we are also interested in estimating the emission characteristics of the N_s potential emission sources, such as their constant emission strengths $\mathbf{c} = [c_1, c_2, ..., c_{N_s}]$, and their locations $\mathbf{x^s} = \{\mathbf{x_1^s}, \mathbf{x_2^s}, ..., \mathbf{x_{N_s}^s}\}$.

3 Proposed Method

3.1 Multi-task learning framework

Multi-task learning [12, 11] is a learning paradigm, where the knowledge from one task can be utilized to improve the performance of the model on other similar tasks. Here, we aim to leverage the strong correlation between the two tasks of identifying the source information and reconstructing the concentration field. Note that, through the advection-diffusion equation, the concentration field is strongly coupled with the source characteristics given $\mathbf{u}$. We propose to learn a shared encoder $g(\boldsymbol{\theta}) : [\Phi, \mathbf{x}^\Phi, \mathbf{u}] \rightarrow \mathbf{z}$ which aims to learn a latent representation $\mathbf{z}$. Then, this latent representation $\mathbf{z}$, is fed into task-specific decoders $f_1(\boldsymbol{\omega}) : \mathbf{z} \rightarrow \phi$ and $f_2(\boldsymbol{\psi}) : \mathbf{z} \rightarrow [\mathbf{c}, \mathbf{x}^s]$ to find the solution of the inverse problem and obtain the reconstructed field.

The multi-task learning framework can then be optimized using the objective function $\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\omega}, \boldsymbol{\psi}) = \mathcal{L}^{\text{recon}}(\boldsymbol{\theta}, \boldsymbol{\omega}) + \mathcal{L}^{\text{inverse}}(\boldsymbol{\theta}, \boldsymbol{\psi})$, where $\mathcal{L}^{\text{recon}}(\boldsymbol{\theta}, \boldsymbol{\omega})$ is the loss function for reconstructing the spatio-temporal concentration field, and $\mathcal{L}^{\text{inverse}}(\boldsymbol{\theta}, \boldsymbol{\psi})$ is the loss function for the inverse problem.

3.2 Field Reconstruction using Diffusive Masked Convolution

We employ a series of masked convolutions for the reconstruction of the spatio-temporal concentration field from the sparse observations. We first define a "masked" convolution operation by restricting the convolution operation on the masked region centered around the sensor locations where we actually

have the observations. The masked convolution operation is defined as,

$$x' = \begin{cases} W^T(X \odot M)\frac{\text{sum}(\mathbb{1})}{\text{sum}(M)} + b, & \text{if sum}(M) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where W and b, respectively, denote the weights and bias of a convolutional filter, X is the current feature (pixel) values for the convolutional sliding window, M is a corresponding mask, and $\odot$ represents element-wise multiplication operation. Note that the masked convolution is defined similar to the Partial Convolutions [6] introduced for an inpainting problem.

Then, the mask, M, is updated using a diffusion process after each masked convolution layer. The diffusion process is modeled by a Spatial Gaussian Convolutional Kernel of size $k \times k$ denoted by $\tilde{W}_k \in \mathbb{R}_+^{k \times k}$ as follows:

$$\tilde{W}_k(i,j) = exp(-\frac{1}{2\sigma_k^2}\Big[(i - \frac{k-1}{2})^2 + (j - \frac{k-1}{2})^2\Big]); \;\; i,j \in \{0,1,..,k-1\}, \tag{3}$$

where i, j are the i-th row and j-th column of the $k \times k$ kernel, σ_k is the only learnable parameter of the kernel. In other words, if σ_k is large then the spatial-convolution kernel $\tilde{W}_k$ would diffuse the mask values over a larger spatial region. Thus, the learnable-parameter σ_k controls the rate of diffusion between the different layers.

The spatial kernel $\tilde{W}_k$ can be repeated across time once (since the size of the kernel along time is one to ensure time-invariance) and the input C_{in} and output C_{out} channels to form the time-invariant Diffusion Kernel W_k^{diff}, and the mask is updated using the diffusion as $M' = (W_k^{\text{diff}})^T M$.

After the mask update we additionally clip the mask values greater than one. We first initialize the mask at the input layer M_0 by the sparse sensor-network locations, as shown in Equation 4.

$$M_0(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \mathbf{x}^{\mathbf{\Phi}} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Then, we stack multiple Diffusive Masked Convolution layers so that the mask M_0 gradually grows after each convolution layer and ultimately we can reconstruct the full field from the sparse sensor measurements. Note that this formulation using the Diffusive Masked Convolution can encode arbitrarily placed sensors as the input mask. Thus, by training such a model with different sparse sensor measurements, it is possible to predict the concentration fields on unseen sensor-network configurations.

3.2.1 Gaussian Negative Log-likelihood Formulation for Uncertainty Quantification

Here, we employ a Gaussian model, where $\phi \sim \mathcal{N}(\phi_\mu, \sigma^2)$ to quantify the uncertainty in the predicted concentration field. The negative log-likelihood loss function is given as,

$$\mathcal{L}^{\text{recon}} = \frac{1}{2}\mathbb{E}_{x \sim p_{data}}[\log \sigma^2 + \frac{(\phi_{true} - \phi_\mu)^2}{\sigma^2}] \tag{5}$$

where ϕ_μ and σ^2 and denote the predicted mean and variance of the concentration field respectively. Later, we demonstrate that the Gaussian negative log-likelihood loss formulation prevents over-fitting on the ground truth concentration field and offers a smooth estimation of the predicted mean concentration ϕ_μ.

3.3 Emission Characteristics Estimation

Next, we consider the task of solving the inverse problem of estimating the emission characteristics. One of the challenges is that the number of potential sources can vary, thus, the decoder f_2 should be able to handle this varying output size. We propose to divide the 2D spatial domain ($N_x \times N_y$) into a $S \times S$ grid. For each cell in the $S \times S$ grid we predict the following source characteristic vector $[p_i, x_i, y_i, c_i]$, where $i \in \{1, 2, ..., S^2\}$, p_i is the probability of containing a source in the grid cell i, x_i and y_i represent the relative location of the source with respect to the top left corner of the cell, and c_i represents the emission strength of the source.

The objective function for estimating the emission characteristics can be evaluated as follows:

$$\mathcal{L}^{\text{inverse}} = \lambda_{src} \sum_{i=1}^{S^2} \mathbb{1}_{ij}^{src} [(x_i^s - \hat{x}_i)^2 + (y_i^s - \hat{y}_i)^2] + \lambda_{src} \sum_{i=1}^{S^2} \mathbb{1}_{ij}^{src} [(\sigma(\hat{p}_i) - 1)^2] + \lambda_{nosrc} \sum_{i=1}^{S^2} \mathbb{1}_{ij}^{nosrc} [(\sigma(\hat{p}_i))^2] + \lambda_{src} \sum_{i=1}^{S^2} \mathbb{1}_{ij}^{src} [(c_i - \text{Softplus}(\hat{c}_i))^2] \quad (6)$$

where, $\mathbb{1}_i^{src}$ denotes if the source occurs in the grid cell i, and $\mathbb{1}_{ij}^{src}$ denotes if the j-th source characteristic vector in grid cell i is "responsible" for that prediction.

The neural network architecture and implementation details of the proposed framework using diffusive-masked convolution are provided in the Appendix A.

4 Results

Experiment Setup: We performed our experiments on 4000 simulations of the forward problem of the advection-diffusion equation, with varying source locations and their emission strengths under stochastically generated wind conditions. In each simulation, the number of sources is randomly selected between one to four. These set of simulations were there divided into training and test sets to train our proposed multi-task learning framework.

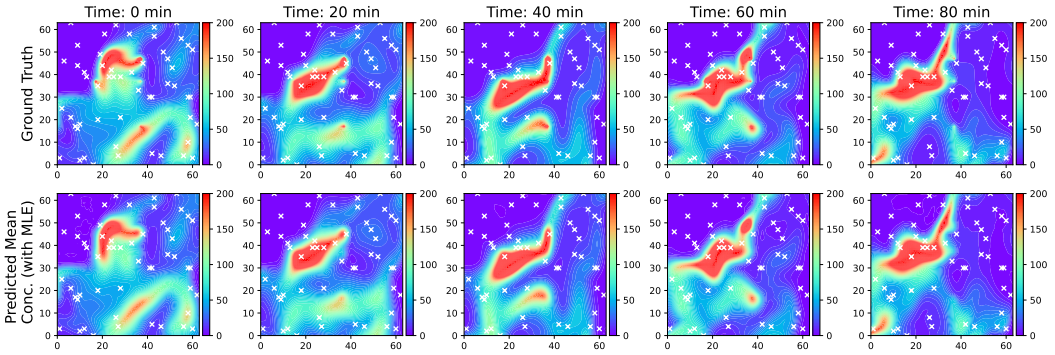

Figure 1: Demonstrating the reconstruction of the global field from sparse sensor measurements on a representative test example. The white crosses denote the position of the sensors. Top row: shows the ground truth concentration fields at various time stamps, Bottom row: the predicted mean of the concentration fields

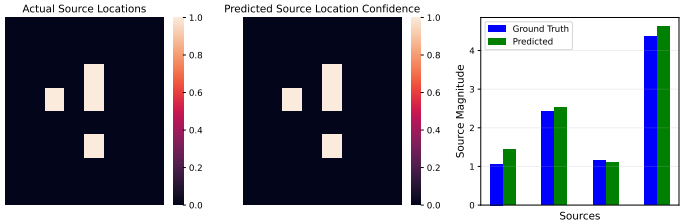

Figure 2: Localization of the pollution sources (left and middle) and their emission strength estimation (right) on a representative test example.

Evaluating Proposed Framework on Simulations: The proposed multi-task learning framework was trained on 80% of the simulations and tested on the rest. To evaluate the performance on the global field reconstruction task, we compute the Mean Squared Error (MSE) between the predicted concentration field and the ground truth concentration field on the test set. We are able to obtain high fidelity reconstructions of the global concentration field using our proposed diffusive masked convolution with a MSE of 0.051. An example of the reconstructed concentration field on a test case is shown in Figure 1. To evaluate the performance on the inverse problem, i.e., localization and emission

characteristics estimation of the pollution sources, we compute the precision, recall and relative location MSE, which provides us an estimate of the detection efficiency of the pollution sources. We also compute the MSE of the source magnitude strengths to quantify the error in approximating the emission strengths. We are able to demonstrate a fairly high recall of 94.5%, and a precision of 71.0%. This suggests that the model is able to recover 94.5% of the pollution sources that were present in the simulations. However, the low precision suggests that the model also generates about 30% false positives in its predictions. Further, we observe that once a grid cell detects a pollution source, it can exactly pin-point the location of the source inside the grid with a relative MSE of $1.35e-07$. We also demonstrate an accurate estimation of the emission strength for the pollution sources having an error of 0.19. Figure 2 shows the ability of our proposed framework to accurately identify the pollution sources as well as approximate their emission strengths on an example case. We also demonstrate in Appendix Figures 5 and 4 that the Gaussian-Negative Loglikelihood formulation for estimating uncertainties allows the estimate of the mean concentration field to be a smooth function, as the large fluctuations near the pollution sources are taken care by the predicted variance, thereby preventing oscillatory solutions due to the Delta-function-like behaviors around the sources.

5 Conclusion

We present a multi-task learning framework for identifying potential pollution sources and obtaining reconstruction of spatio-temporal concentration fields from sparse sensor measurements. We also propose a novel diffusive-masked convolution operations that employs the diffusion process in performing masked-convolutions. The diffusive-masked convolution is realized by a spatial-Gaussian convolution kernel, followed by diffusing the mask to nearby regions. Thus, by stacking multiple such layers we are able to iteratively diffuse the information from a sparse sensor measurements (represented using the initial mask) in a principled manner until the learned representations spread over the entire spatio-temporal domain. We also demonstrate precise reconstruction of the concentration field along with accurate localization and emission strength estimation of the pollution sources on the test simulations using our proposed framework.

6 Broader Impact

The problem of reconstructing the spatio-temporal fields of a complex-time evolving dynamical system from sparse sensor observations is common in the field of science and engineering. Examples of such cases include reconstructing the oceanic currents from sparse measurements from a bouy network, or reconstructing the seismic waves from measurements taken at sparse station networks. In this paper, we present a novel and principled approach of reconstructing the global fields from sparse sensor measurements motivated by the process of diffusion. Our proposed framework is obtain high fidelity reconstruction of the target spatio-temporal field from any arbitrarily placed sensor network (with variable number of sensors), without the need to retrain the model. In addition, we also provide a generic framework to solve inverse problems that require precise localization of sources and their characterizations. Again, a wide range of problems in the field of science and engineering fall under this category such as locating the epicenter of seismic activities from sparse station observations.

References

[1] Mingkui Cao, Keith Gregson, and Stewart Marshall. Global methane emission from wetlands and its sensitivity to climate change. *Atmospheric environment*, 32(19):3293–3299, 1998.

[2] Kai Fukami, Romit Maulik, Nesar Ramachandra, Koji Fukagata, and Kunihiko Taira. Global field reconstruction from sparse sensors with voronoi tessellation-assisted deep learning. *Nature Machine Intelligence*, 3(11):945–951, 2021.

[3] Robert W Howarth. A bridge to nowhere: methane emissions and the greenhouse gas footprint of natural gas. *Energy Science & Engineering*, 2(2):47–60, 2014.

[4] Robert W Howarth, Renee Santoro, and Anthony Ingraffea. Methane and the greenhouse-gas footprint of natural gas from shale formations. *Climatic change*, 106(4):679–690, 2011.

[5] Youngdeok Hwang, Hang J Kim, Won Chang, Kyongmin Yeo, and Yongku Kim. Bayesian pollution source identification via an inverse physics model. *Computational Statistics & Data Analysis*, 134:76–92, 2019.

[6] Guilin Liu, Fitsum A Reda, Kevin J Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European conference on computer vision (ECCV)*, pages 85–100, 2018.

[7] Xiao Liu, Kyongmin Yeo, and Siyuan Lu. Statistical modeling for spatio-temporal data from stochastic convection-diffusion processes. *Journal of the American Statistical Association*, 117 (539):1482–1499, 2022.

[8] Fiona M O'Connor, O Boucher, N Gedney, CD Jones, GA Folberth, R Coppell, P Friedlingstein, WJ Collins, J Chappellaz, J Ridley, et al. Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review. *Reviews of Geophysics*, 48(4), 2010.

[9] Dave Reay and Pete Smith. *Methane and climate change*. Routledge, 2010.

[10] ED Schulze, S Luyssaert, P Ciais, A Freibauer, IA Janssens, Jean-François Soussana, P Smith, John Grace, I Levin, B Thiruchittampalam, et al. Importance of methane and nitrous oxide for europe's terrestrial greenhouse-gas balance. *Nature geoscience*, 2(12):842–850, 2009.

[11] Trevor Standley, Amir Zamir, Dawn Chen, Leonidas Guibas, Jitendra Malik, and Silvio Savarese. Which tasks should be learned together in multi-task learning? In *International Conference on Machine Learning*, pages 9120–9132. PMLR, 2020.

[12] Yu Zhang and Qiang Yang. A survey on multi-task learning. *IEEE Transactions on Knowledge and Data Engineering*, 2021.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [No]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] Is applicable to a wide set of problems in science and engineering.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [No]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No]

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No]

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No] Appendix was strongly discouraged. Did not have space to include in main paper.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [No]

(b) Did you mention the license of the assets? [N/A]

(c) Did you include any new assets either in the supplemental material or as a URL? [N/A]

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Network Architecture for Proposed Multi-task Learning Framework

Our encoder-decoder architecture presented in the 3 is realized using a U-Net model with skip connections, where each of the conventional convolutional layers (Conv3d in our case) are replaced by our proposed diffusive masked-convolutional layers. In our experiments, the encoder and decoder are comprised of 4 layers each.

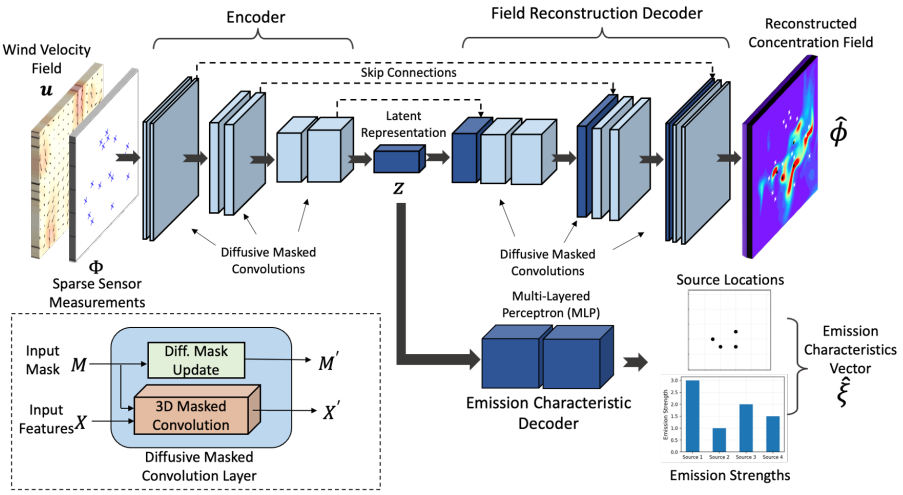

Figure 3: Neural Network Architecture of our proposed multi-task learning framework.

B Additional Visualization of the Results

In this section, we provide some additional visualizations to qualitatively analyze the behavior of our proposed framework. First, for a fixed value of the y-coordinate ($y = 30$ and $y = 37$) we plot the variations in the concentration fields w.r.t. the x-axis at various times in Figure 4. We observe that the predictive intervals estimated by our model always engulf the ground truth concentration. Additionally, we show that the Gaussian-Negative Loglikelihood (NLL) formulation for estimating uncertainties allows the estimate of the mean concentration field to be a smooth function, as the large fluctuations near the pollution sources are taken care by the predicted variance, thereby preventing an oscillations solution. This behavior can also be seen in Figure 5, where we see that the absolute error without the NLL shows rectangular oscillatory patterns, which is nor observed for the model trained with NLL.

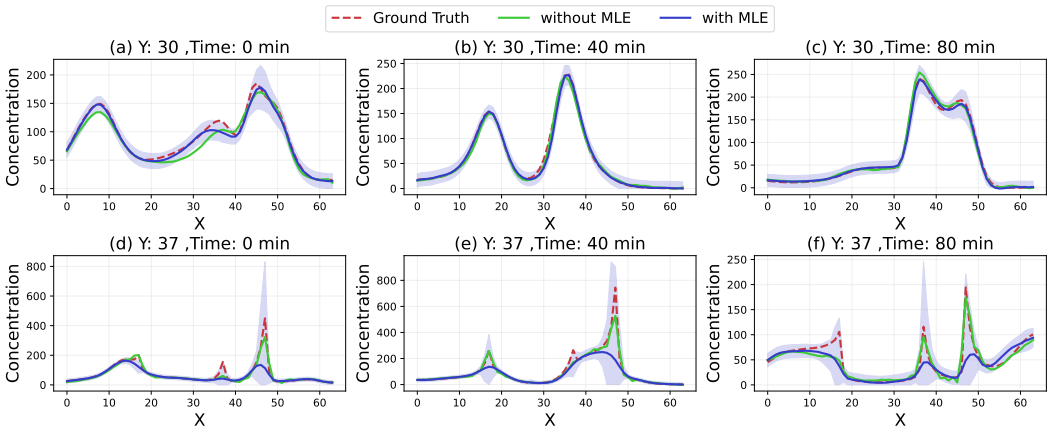

Figure 4: Analyzing the empirical coverage of the 95% Prediction Interval (shown as the shaded region) for the model trained on Gaussian Negative Log-likelihood (NLL) loss for a particular spatial slice of $y = 30$ and $y = 37$ (a line passing through 3 pollution sources, that can be identified using the peaks in the ground truth) on an example case.

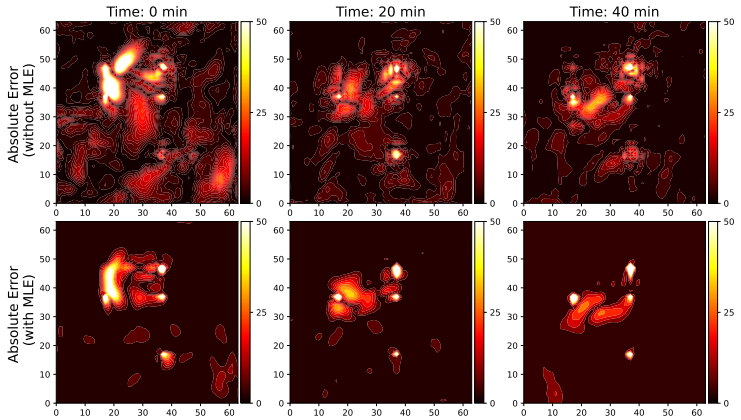

Figure 5: Comparing the Absolute Errors of the model trained with and without the Gaussian Negative Log-likelihood (NLL) loss at different time intervals on an example test case.

An example showing side-by-side comparisons of the predicted concentration fields are shown in Figure 6.

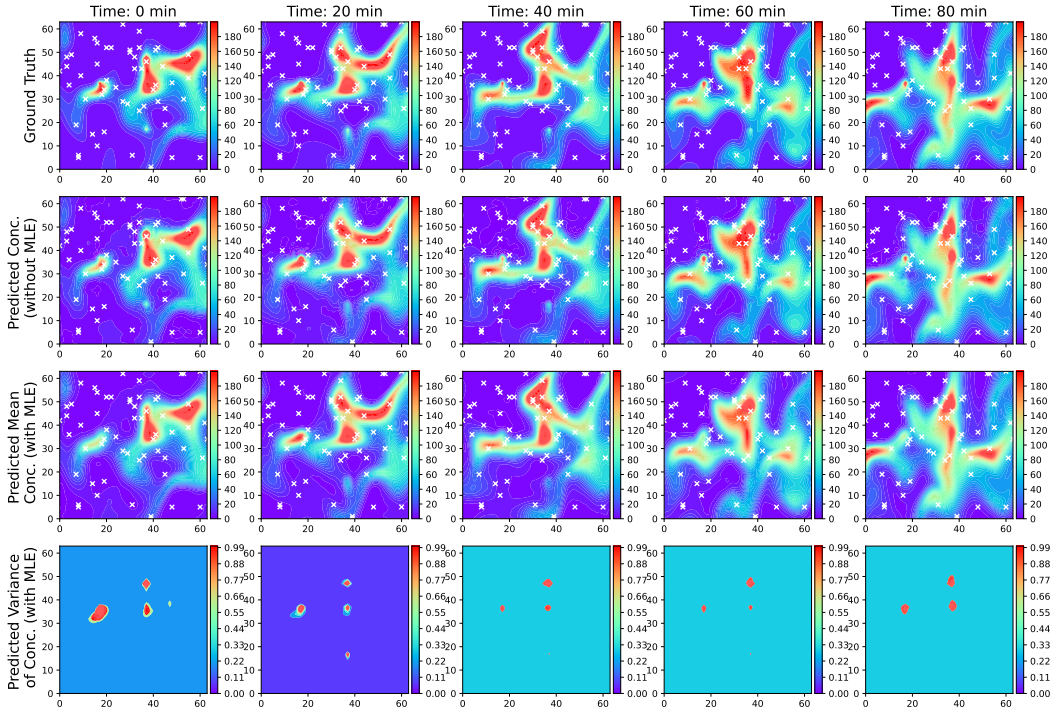

Figure 6: Demonstrating the reconstruction of the global field from sparse sensor measurements on a representative test example. The white crosses denote the position of the sensors. Top row: shows the ground truth concentration fields at various time stamps, Middle row: the predicted mean of the concentration fields, Bottom row: the predicted variance of the concentration fields