

A staged deep learning approach to spatial refinement in 3D temporal atmospheric transport

M. Giselle Fernández-Godino ^a^{*,1}, Wai Tong Chung ^b, Akshay A. Gowardhan ^a,
Matthias Ihme ^b, Qingkai Kong ^a, Donald D. Lucas ^a, Stephen C. Myers ^a

^a Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, United States of America

^b Stanford University, 440 Escondido Hall, Stanford, CA 94305, United States of America

ARTICLE INFO

Keywords:

Atmospheric sciences
Geosciences
Plume transport
3D temporal sequences
Artificial intelligence
CNN
LSTM
Autoencoder
Autoregressive model
U-Net
Super-resolution
Spatial refinement

ABSTRACT

High-resolution spatiotemporal simulations effectively capture the complexities of atmospheric plume dispersion in complex terrain. However, their high computational cost makes them impractical for applications requiring rapid responses or iterative processes, such as optimization, uncertainty quantification, or inverse modeling. To address this challenge, this work introduces the Dual-Stage Temporal Three-dimensional UNet Super-resolution (DST3D-UNet-SR) model, a highly efficient deep learning model for plume dispersion predictions. DST3D-UNet-SR is composed of two sequential modules: the temporal module (TM), which predicts the transient evolution of a plume in complex terrain from low-resolution temporal data, and the spatial refinement module (SRM), which subsequently enhances the spatial resolution of the TM predictions. We train DST3D-UNet-SR using a comprehensive dataset derived from high-resolution large eddy simulations (LES) of plume transport. We propose the DST3D-UNet-SR model to significantly accelerate LES of three-dimensional (3D) plume dispersion by three orders of magnitude. Additionally, the model demonstrates the ability to dynamically adapt to evolving conditions through the incorporation of new observational data, substantially improving prediction accuracy in high-concentration regions near the source.

1. Introduction

Modeling atmospheric plume dispersion requires resolving the many scales of turbulence in fluid flows, a task for which large eddy simulations (LES) techniques are particularly suited. LES captures the dynamics of turbulent eddies, yielding high-fidelity insights into atmospheric dispersion phenomena. However, the computational cost of LES is prohibitive for real-time applications, such as emergency response, and for iterative processes like optimization and uncertainty quantification (UQ) that demand large-scale simulations. To address these constraints, machine learning-based models have been developed, offering a trade-off between computational efficiency and predictive accuracy.

In this context, we introduce the Dual-Stage Temporal 3D UNet Super-Resolution (DST3D-UNet-SR) model, a deep learning framework that significantly accelerates LES-based plume dispersion modeling while retaining high spatial resolution. DST3D-UNet-SR leverages a dual-stage neural network architecture: a temporal module (TM) and

a spatial refinement module (SRM). The TM, built on a U-Net architecture augmented with ConvLSTM layers, processes temporal sequences of low-resolution input data to predict plume evolution over time while capturing dependencies across time steps. Following this, the SRM enhances the spatial resolution of the TM outputs using a 3D U-Net architecture optimized with skip connections and transposed convolutions, enabling a fourfold increase in resolution. By combining these two specialized modules, the DST3D-UNet-SR significantly reduces computational costs by three orders of magnitude while retaining the fidelity of turbulence-resolving LES predictions.

2. Related work

Convolutional neural networks (CNNs) have been extensively utilized in atmospheric modeling due to their strong capability in capturing complex spatial patterns and dynamics. Camps-Valls et al. (2021) comprehensively reviewed CNN applications in Earth sciences, demonstrating their effectiveness in tasks such as climate projections and

* Corresponding author.

E-mail addresses: fernandez48@llnl.gov (M.G. Fernández-Godino), wchung@stanford.edu (W.T. Chung), gowardhan1@llnl.gov (A.A. Gowardhan), mihme@stanford.edu (M. Ihme), kong11@llnl.gov (Q. Kong), lucas26@llnl.gov (D.D. Lucas), myers30@llnl.gov (S.C. Myers).

¹ https://scholar.google.com/citations?user=ibl3c_gAAAAJ&hl=en.

air quality assessments. Specifically, Yu and Liu (2021) integrated satellite-derived NO₂ measurements (TROPOMI) with ground-level observations using CNNs, achieving significant improvements in spatial and temporal resolution compared to traditional interpolation methods. Fernández-Godino et al. (2023) further validated the applicability of CNNs in spatial deposition modeling, illustrating their effectiveness in reconstructing fine-scale atmospheric features.

Early CNN-based methods, such as the Super-Resolution Convolutional Neural Network (SRCNN) by Dong et al. (2015), established foundational frameworks for image-based spatial refinement tasks. Subsequently, super-resolution methods have evolved substantially, introducing various advanced architectures. For example, Jiang et al. (2020) developed the Dual-path Deep Fusion Network, demonstrating remarkable success in face image hallucination by integrating multi-level features. Xiao et al. (2023) proposed EDiffSR, leveraging diffusion probabilistic models to achieve high-quality super-resolution in remote sensing imagery. Additionally, Wu et al. (2024) introduced Modal Adaptive Super-Resolution, applying continual learning strategies to medical imaging tasks, thus highlighting the adaptability and effectiveness of super-resolution across diverse domains.

Temporal modeling plays a crucial role in forecasting tasks by analyzing time-dependent data to identify patterns and trends, thereby enabling accurate predictions of future events. Xiao et al. (2022) introduced the DeepCAMS model, employing a two-stage CNN approach combining video frame interpolation and spatial CNN-based super-resolution to downscale coarse-resolution aerosol data into high-resolution global PM_{2.5} estimates. While temporal frame interpolation enhances temporal resolution by creating intermediate data points, Long Short-Term Memory (LSTM) networks provide a sophisticated approach to modeling intricate temporal dependencies in sequential data. For example, Marcos Filho et al. (2024) demonstrated LSTM's capability to predict the complex temporal evolution of radioactive plumes in emergency scenarios, emphasizing their strength in capturing sequential dependencies. Similarly, our prior research (Fernández-Godino et al., 2021; Garcia-Cardona et al., 2022; Wang et al., 2021, 2024) demonstrated the strength of LSTMs in capturing sequential dependencies for forecasting applications.

Recognizing these complementary strengths, recent approaches have integrated CNNs and LSTMs for spatiotemporal modeling. For instance, the *Deep-Air* model by Tsokov et al. (2022) couples CNN-based spatial feature extraction with LSTM temporal modeling, successfully predicting air pollution across city scales. Similarly, Ding et al. (2021) employed a CNN-LSTM hybrid to forecast daily PM_{2.5} concentrations, demonstrating improved accuracy by simultaneously capturing spatial correlations and temporal dynamics. The Next Day Wildfire Spread dataset by Huot et al. (2022) offers a large-scale, multivariate benchmark for one-day fire propagation prediction that has demonstrably improved the training and evaluation of CNN-based segmentation models. Yet, these studies primarily handled two-dimensional (2D) spatial data and did not explicitly target spatial super-resolution.

Recent advancements have explored fully 3D spatiotemporal downscaling. Xiao et al. (2022) presented *DeepCAMS*, a two-stage CNN-based model that downscales coarse-resolution aerosol data first temporally, through video frame interpolation, and then spatially, via a CNN-based super-resolution step. Additionally, Geiss et al. (2022) introduced a unified approach employing full 3D CNNs to simultaneously handle latitude, longitude, and temporal dimensions, achieving high fidelity in downscaled atmospheric chemistry simulations. However, these fully integrated 3D CNNs entail substantial computational demands due to their high-dimensional input data.

In contrast, our proposed DST3D-UNet-SR model adopts a modular two-stage design, separating temporal forecasting from spatial refinement. Unlike the full 3D spatiotemporal CNN approach of Geiss et al. (2022) or the staged temporal and spatial interpolation of Xiao et al. (2022), our method explicitly avoids performing convolutions over time as a spatial dimension. Instead, temporal dependencies are

modeled using a ConvLSTM-driven Temporal Module (TM), while a separate Spatial Refinement Module (SRM), based on the 3D U-Net architecture presented by Chung et al. (2023), subsequently upsamples each prediction to high spatial resolution. This choice offers multiple practical advantages.

Firstly, our modular design allows each component to be optimized independently, facilitating flexibility and targeted improvements. The TM can incorporate newer sequence modeling techniques or updated observational data without retraining the SRM, and vice versa. Such modularity simplifies model maintenance and evolution, aligning with recommendations for scientific machine learning applications (Goodfellow et al., 2016).

Secondly, computational efficiency is greatly enhanced by reducing dimensionality at each modeling stage. The TM operates on lower-resolution data, significantly reducing computational load. Subsequently, the SRM processes higher-resolution outputs but only for single timesteps, which is memory-efficient compared to full 3D CNN processing. This staged approach substantially reduces complexity without sacrificing performance, a crucial consideration for real-time or iterative applications.

Thirdly, our approach supports dynamic updating, a critical capability for sparse and evolving data scenarios common in atmospheric dispersion applications. Temporal inputs are represented as stacked spatial channels processed autoregressively by the ConvLSTM, allowing integration of new observational data at each prediction step without model retraining. This approach naturally accommodates real-time data assimilation, which would be computationally challenging in fully 3D CNN frameworks.

Additionally, the convolutional U-Net structure with skip connections ensures large-scale patterns are preserved and local details enhanced, addressing concerns about physical consistency and conservation noted in previous super-resolution studies (LeVeque, 2002; Wang et al., 2004).

DST3D-UNet-SR architecture uniquely integrates U-Net CNN super-resolution techniques and ConvLSTM temporal modeling into a coherent modular framework for efficient 3D plume dispersion modeling. This approach addresses critical gaps in previous research, which either limited their scope to two-dimensional fields or avoided super-resolution altogether (Dong et al., 2015; Lanaras et al., 2018; Liebel and Körner, 2016). Our modular design demonstrates that a staged deep learning model can efficiently produce accurate high-resolution 3D atmospheric transport predictions, offering significant practical benefits in computational speed, flexibility, and adaptability to new data.

3. Data

The simulations, referred in this work as *ground truth*, were originally designed to replicate similar conditions of the REACT (RElease ACTivity) experiment, conducted in October 2022 at the Nevada National Security Site (NNSS), where real-time xenon sensors monitored radiotracer releases (Wharton et al., 2023; Stave et al., 2024) (Fig. 9). Xenon is a reliable indicator of nuclear explosions, particularly underground nuclear tests. We collaborate with researchers who utilize data from the International Monitoring System or IMS, a network designed to detect and verify nuclear explosions globally to ensure compliance with the Comprehensive Nuclear-Test-Ban Treaty (CTBT).

A continuous release of Xe-133 was simulated, accounting for its half-life of 5.2 days. However, given that the simulation spans only five hours, the decay rate is negligible and does not impact the results within the considered timeframe. The xenon release occurs at the center of the complex terrain domain of 10 km × 10 km, i.e. ($x_{\text{release}}, y_{\text{release}}$) = (5 km, 5 km), as shown in Fig. 1 using a white star symbol. The gas is released continuously from a point source with a unit mass that can be scaled to cover specified source amounts. Xenon is used as a tracer in release experiments because it is a non-reactive, inert noble gas with

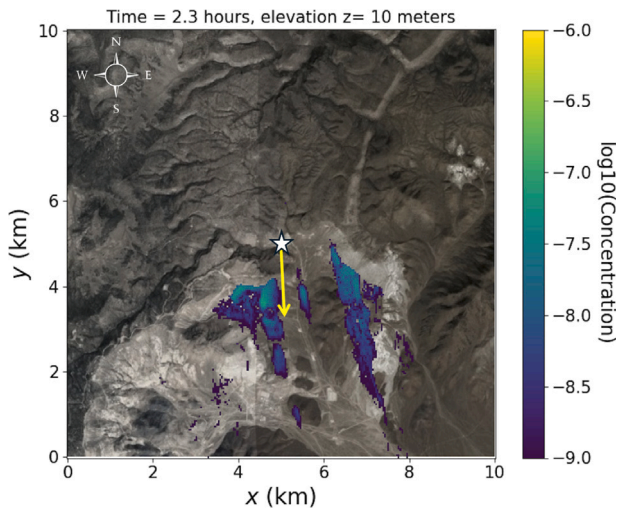


Fig. 1. An example of a xenon field at 10 m slice above the ground level with $w_s = 7.8$ m/s, $w_d = 357^\circ$ after 2.3 h following the start of the release. The white star marks the release location, while the yellow arrow indicates the wind direction. The complex terrain, characterized by mountains and valleys, is shown in gray scale.

a non-toxic nature that enables precise sensor measurements as later shown in Section 7.

The Aeolus (Gowardhan et al., 2021) dispersion code was used to perform the simulations. Aeolus is an advanced LES model specifically engineered to predict flow fields and the transport of gases and particles in areas characterized by urban environments and complex terrain, which poses unique challenges for atmospheric modeling. By employing the LES methodology, the model effectively solves the three-dimensional incompressible Navier–Stokes equations, allowing for a detailed representation of turbulent flow dynamics. The transport of tracers is intricately modeled using an in-line coupled Lagrangian dispersion approach, where each marker particles symbolizes a defined substance, derived from the source term applicable to both gaseous emissions and particulate matter. These marker particles are dynamically transported by the wind and turbulence calculated by the model, reflecting real-world atmospheric conditions. As they traverse the landscape, these particles interact with the surrounding terrain, leading to mass loss through processes such as gravitational settling, where heavier particles fall to the ground, and various non-gravitational deposition mechanisms applicable to gases, including adsorption and impaction, which further influence the distribution and concentration of gas and particles in the environment. This comprehensive modeling framework enables Aeolus to provide critical insights into tracer dispersion, aiding in environmental assessments and the development of effective mitigation strategies.

The dataset consists of 100 runs: 80 for training, 10 for validation, and 10 for testing. The simulations varied the parameters wind speed (w_s) and wind direction (w_d) selected using Latin Hypercube Sampling (LHS) (Stein, 1987), with w_s ranging from 1.5 to 10 m/s and w_d from 340° to 360° , representing the predominant wind conditions at this site. While inflow wind speed and direction are boundary conditions and were the primary sampled parameters, the complex terrain introduces significant stochastically driven turbulent transport in plume evolution due to nonlinear interactions with local topography. As a result, even simulations with similar inflow conditions can produce markedly different dispersion patterns, enhancing the effective diversity of the dataset. The complex terrain, characterized by mountains and valleys, presents a challenging problem. The ML model must learn the discontinuities in flow concentration. Fig. 1 presents a 2D slice of one of the concentration fields used for training.

The dataset was curated to ensure quality and relevance for training the ML model. From the 330 time steps extracted from Aeolus, each

one minute apart, we retained 33 time steps for training at 10 min intervals, resulting in a training duration of 5.5 h. Although the original test simulations span 330 time steps at 1 min intervals, we evaluate the model using only 33 time steps at 10 min intervals, consistent with the training resolution. The simulations were run using a resolution of $(z, y, x) = 200 \times 250 \times 250$ cells, each cell representing $20 \text{ m} \times 40 \text{ m} \times 40 \text{ m}$. Concentrations are recorded at the cell center. A height of $z = 0$ corresponds to ground level and is unrelated to sea level. Plume concentration data is presented logarithmically to facilitate visualization.

Initially, the simulation domain comprised a volume of $[z, x, y] = [4, 10, 10]$ km, which was then cropped to $[2, 5, 5]$ km for training purposes due to the plume's predominant movement towards the southeast quadrant and altitudes less than 2 km, influenced by wind velocity. At this stage, the dataset is resized into two distinct resolutions using the *resize* function from the Python package *skimage.transform*. The low-resolution data is formatted with dimensions [number of time steps, height, width, depth] = [33, 8, 32, 32], where the height, width, and depth correspond to the z , x , and y spatial dimensions, respectively. Each grid cell represents $250 \text{ m} \times 156.3 \text{ m} \times 156.3 \text{ m}$. The high-resolution data, on the other hand, is stored with dimensions [number of time steps, height, width, depth] = [33, 32, 128, 128], where the z , x , and y spatial dimensions correspond to grid cells of $62.5 \text{ m} \times 39.1 \text{ m} \times 39.1 \text{ m}$. Both datasets are stored in 32-bit format, with the low-resolution dataset totaling 0.1 GB and the high-resolution dataset totaling 6.6 GB.

The cropped domain was selected to focus on the region most impacted by plume dispersion, which is both scientifically significant and operationally relevant. The choice is motivated by the observation that plume concentrations rapidly dilute outside a radius of 5 km and an altitude of 2 km, making the extended domain less critical for prediction. This focus also reduces computational complexity and ensures that the training emphasizes high-gradient regions near the source, which pose greater modeling challenges.

The Aeolus LES model, as a data source, provides us with sufficient data cubes to train our model effectively. However, for cases with sparse data, data symmetries can be employed for data augmentation (Fernández-Godino et al., 2019) and performance enhancement. To predict gas dispersion, data can often be augmented for different wind directions by rotating the plume. However, in our case, this approach is not applicable due to asymmetries in the terrain.

4. DST3D-UNet-SR architecture

The DST3D-UNet-SR model is designed to predict the temporal and spatial resolution of 3D plume dispersion by integrating two specialized neural network architectures: a temporal module (TM) for predicting the transient evolution from low-resolution data and spatial refinement module (SRM) to enhance the spatial resolution. The separation into these two modules provides several advantages, including flexibility in design and training, modular upgrades, and targeted improvements. This modular approach allows each component to be independently optimized and updated, ensuring that improvements in one module do not impact the other. Therefore, it facilitates the incorporation of new techniques and advancements in either temporal prediction or spatial refinement without requiring a complete redesign of the entire ML model. This section details the architectures and training strategies of the TM and the SRM. The full DST3D-UNet-SR model contains approximately 4.15 million parameters (2.7 parameters per physics cell), with 3.2 million in the TM and 0.95 million in the SRM. A schematic of the two independent modules of the DST3D-UNet-SR model architecture is presented in Fig. 2, and a diagram illustrating the integration between TM and SRM is shown in Fig. 3.

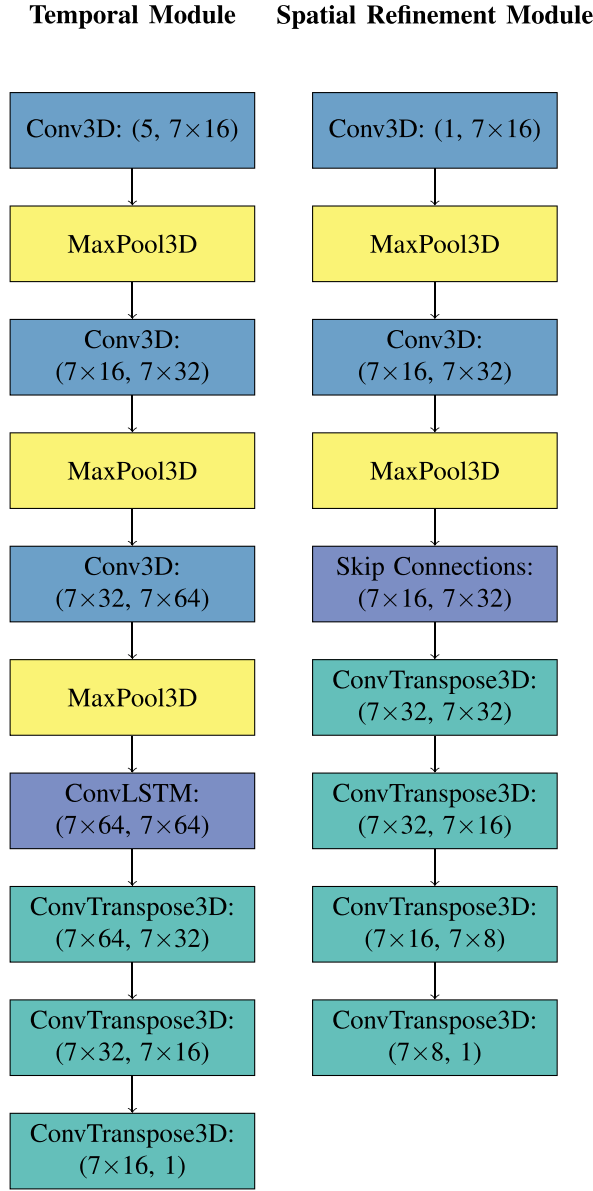


Fig. 2. Dual-Stage Temporal 3D UNet-SR model architecture.

4.1. Temporal module and training strategy

The temporal module (TM) in our DST3D-UNet-SR model is designed to capture and predict the evolution of atmospheric dispersion phenomena in 3D. This model employs a U-Net architecture for spatial features enhanced with ConvLSTM layers, enabling it to handle the spatiotemporal dependencies inherent in the data effectively. The TM inputs are sequences shaped as (batch size, 5, 8, 32, 32), representing a temporal window of five time steps, each with a spatial resolution of $8 \times 32 \times 32$. The encoder section of the model progressively reduces the spatial dimensions while increasing the number of channels, using three convolutional blocks followed by max-pooling and 20% dropout layers to prevent overfitting. The ConvLSTM layer at the bottleneck captures the temporal dependencies across the encoded representations. The decoder section then reconstructs the high-level features back to the original spatial dimensions, incorporating skip connections from the encoder to preserve spatial details. This process results in an output shaped as (batch size, 1, 8, 32, 32) representing a single time step, specifically the one that follows the five previous time steps provided as input,

allowing the model to predict the temporal evolution of the plume concentration autoregressively. Further details on the TM architecture are provided in [Appendix A.1](#).

The training strategy for the TM employs a sliding-window approach across the temporal dimension. Each training sequence consists of five consecutive time steps used to predict the next time step in the sequence. This ensures the model learns to capture temporal dependencies effectively. Various configurations were tested, and a window size of five time steps was found to capture sufficient temporal context while maintaining computational efficiency and model performance. The training process uses the mean squared error (MSE) loss function, and the parameters are updated using the Adam optimizer ([Kingma and Ba, 2014](#)), with a learning rate scheduler to adapt as validation loss plateaus. The TM model was trained independently of the SRM, on an Apple MacBook Pro equipped with an M2 Max chip and 32 GB of memory, utilizing the system's 12-core CPU, with a training time of approximately 1.2 min and 1000 epochs. This low training time is due to the reduced dimensionality of the TM input (from 128×128 to 32×32), which operates on dimensionally-compressed concentration patterns. This allows for efficient CPU-based training without compromising accuracy. Validation loss remained stable throughout training and closely matched training loss, suggesting no overfitting occurred despite the short training time.

During training, model generalization was monitored using a separate validation set of 10 simulations, distinct from the test set. Validation loss was used to adjust the learning rate via a scheduler and to prevent overfitting. Final evaluation was conducted on a held-out test set as described in Section 4. During inference, the TM operates in an autoregressive prediction framework. Initially, the model receives a five-time-step input of shape (1, 5, 8, 32, 32) consisting entirely of ground truth values to predict the subsequent time step, $t + 1$ of shape (1, 1, 8, 32, 32). This prediction is then incorporated into the input sequence for the next iteration, replacing the oldest time step while retaining the remaining ground truth values. As predictions progress, the reliance on ground truth data diminishes, with subsequent inputs increasingly comprising model-predicted values. By $t + 5$, the model operates in a fully predictive regime, where all input time steps are derived exclusively from prior predictions.

4.2. Spatial refinement module and training strategy

The spatial refinement module (SRM) in the DST3D-UNet-SR model focuses on enhancing the spatial resolution of under-resolved 3D atmospheric simulation data through a super-resolution approach ([Dong et al., 2015](#)). The module employs a U-Net architecture specifically designed for 3D data, enabling it to upscale coarse spatial resolutions to high-resolution outputs. The SRM processes input tensors of shape (batch size, 8, 32, 32), representing a low-resolution 3D spatial grid, and produces high-resolution outputs of shape (batch size, 32, 128, 128). Its encoder consists of convolutional layers that progressively reduce spatial dimensions while increasing feature channels. Each convolutional block is followed by batch normalization and LeakyReLU activation functions, with max-pooling layers for spatial downsampling. Skip connections between encoder and decoder layers ensure that spatial information is retained during the reconstruction process. The decoder employs transposed convolutional layers to upscale feature maps, achieving a fourfold enhancement in spatial resolution. While higher upsampling factors are feasible, they can exceed the model's capacity to capture detailed features effectively, as discussed in [Chung et al. \(2023\)](#). Further architectural details of the SRM are provided in [Appendix A.2](#).

The SRM is trained using supervised learning, with low-resolution data generated from high-resolution ground truth through average pooling. During training, the model receives low-resolution inputs and predicts high-resolution outputs, which are optimized to match the ground truth using the mean squared error (MSE) loss function. The Adam optimizer ([Kingma and Ba, 2014](#)) is used for parameter updates,

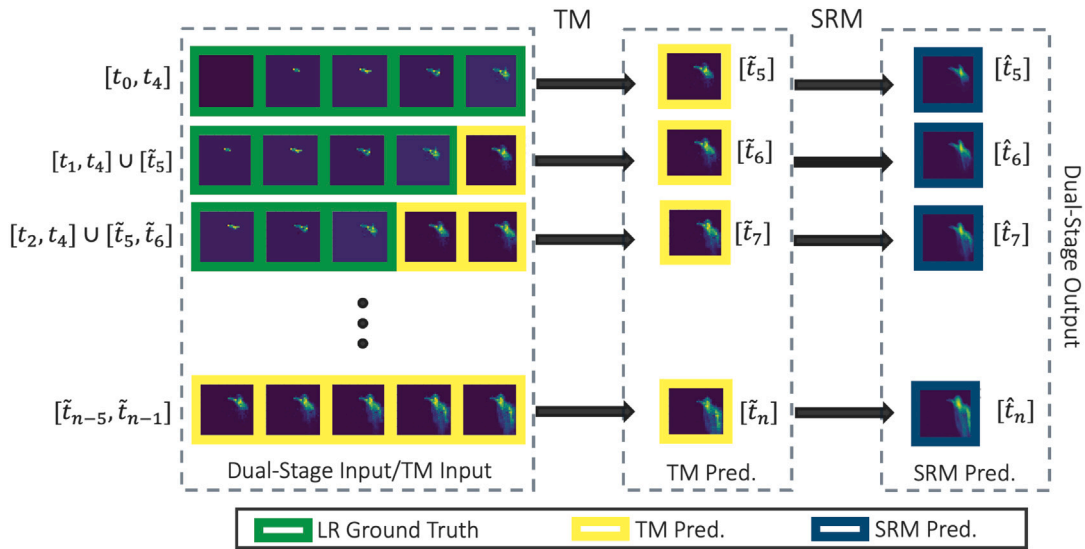


Fig. 3. DST3D-UNet-SR prediction workflow, showing the integration of the TM and the SRM. The TM receives sequences of low-resolution frames at selected time steps (green borders) and predicts future low-resolution frames (yellow borders), denoted as \tilde{t} . For each new prediction, the TM input includes the final four frames of the previous sequence concatenated with the latest predicted frame. These TM inputs and corresponding low-resolution predictions are then passed to the SRM, which refines them into high-resolution outputs (blue borders), denoted as \hat{t} . The process continues iteratively to produce a temporally and spatially refined forecast.

with a learning rate scheduler applied to reduce the rate when the validation loss plateaus. This training approach ensures convergence while capturing fine spatial details and preventing overfitting. The SRM was trained independently of the TM on LLNL's Lassen high-performance computing system, utilizing 4 NVIDIA Volta GPUs in parallel. The training process was completed in approximately 3 h over 100 epochs.

4.3. Integration of temporal and spatial refinement modules

The combined workflow of the TM and SRM can be summarized as follows:

- (1) The TM uses an autoregressive process to make 10 min increment predictions. It receives a batch of low-resolution sequences, each with the shape (batch size, 5, 8, 32, 32). These sequences represent a sliding window of five consecutive time steps at 10 min intervals.
- (2) The TM processes these sequences and outputs a batch of predicted low-resolution frames, each of shape (batch size, 1, 8, 32, 32). This single frame represents the prediction for the next time step following the five previous time steps provided as input.
- (3) The next TM's prediction input consists of the last four time steps of the previous input concatenated with the latest TM's prediction. In other words, the next time step input is constructed by removing the first time step from the previous input and appending the prediction of the latest time step.
- (4) For the first five predictions, the TM includes low-resolution ground truth data in its input sequences. After the initial five predictions, all subsequent predictions are based solely on the model's previous predictions.
- (5) The TM's inputs and the predicted low-resolution frames from the fifth element of each sequence are then fed into the SRM.
- (6) The SRM processes these predicted frames and outputs high-resolution frames, each of shape (batch size, 32, 128, 128).

The workflow shown above is also depicted in Fig. 3.

5. Results

In this section, we assess the DST3D-UNet-SR model's effectiveness in predicting plume dispersion. The evaluation includes comparisons of

the predicted concentrations against ground truth values in the x-, y-, and z- directions.

To evaluate performance, we compare DST3D-UNet-SR against the high-resolution temporal model (HRTM), which shares the same architecture and parameter count as the TM but is trained directly on high-resolution volumetric data. This approach corresponds to direct spatiotemporal modeling strategies explored in recent literature (Stachenfeld et al., 2022; Liu et al., 2024). The HRTM thus provides a controlled baseline that isolates the architectural and workflow differences between the models.

Notably, the HRTM requires high-resolution input data, making it more computationally expensive, although it can achieve higher accuracy for early time steps. In contrast, DST3D-UNet-SR uses low-resolution input data and refines it to high-resolution predictions via its integrated SRM.

The HRTM was trained in parallel using four NVIDIA V100 GPUs (each with 16 GB of memory) on LLNL's Lassen high-performance computing system, with a total training time of approximately 7 h over 100 epochs. Detailed information about the HRTM can be found in Appendix B.

Figs. 5, 4, and 6 show the averaged concentration in the x-, y-, and z-planes, respectively, at various time steps for a single test case with $w_x = 5.7$ m/s and $w_d = 350.5^\circ$. These figures compare DST3D-UNet-SR and HRTM predictions (two top rows) with ground truth (third row). Variations are depicted across the other two dimensions within specified ranges, with appropriate spatial resolutions. Specifically, Fig. 4 shows variations across the x- and z-plane within ranges of 0 to 5 km and 0 to 2 km, respectively, with spatial resolutions of 0.04 km for y and 0.02 km for z vertical slices. Fig. 5 shows variations across the y- and z-plane within the same ranges and resolutions. Fig. 6 shows variations across the x- and y-plane within a range of 0 to 5 km, with a spatial resolution of 0.04 km for both dimensions.

In Figs. 5 to 6 we can see that qualitatively both, DST3D-UNet-SR and HRTM, perform well. The associated errors, presented as the difference between the predictions and ground truth give more insight. In high-resolution models like the HRTM, errors in early timesteps can rapidly propagate and grow because the model tries to predict both temporal evolution and fine spatial details simultaneously. This can result in exaggerated or over-predicted regions, particularly outside the main plume, as we see in the HRTM error row of Fig. 5, Fig. 4, and Fig. 6. The DST3D-UNet-SR's approach of handling temporal and

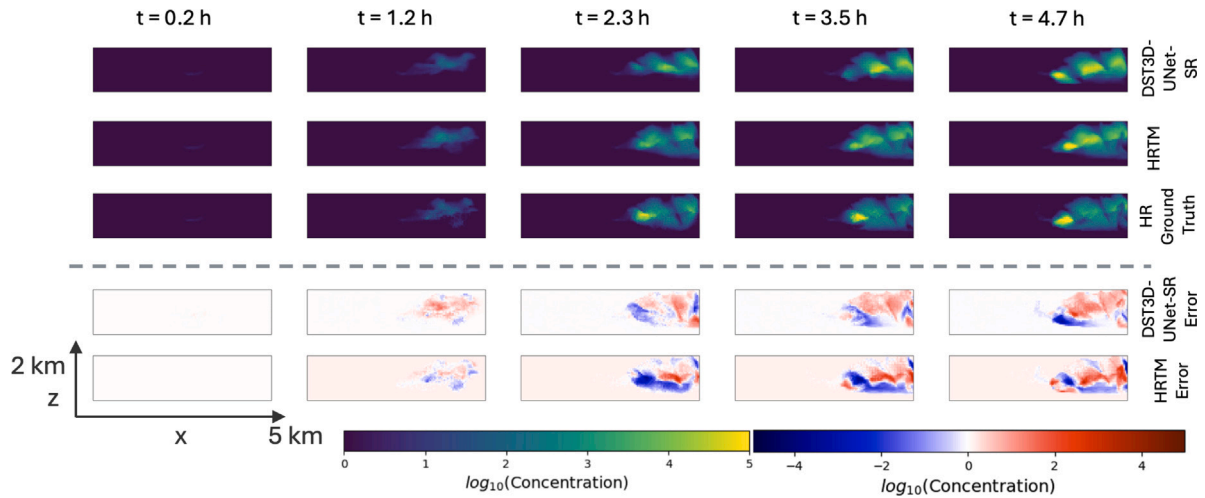


Fig. 4. Averaged concentration in the x - z plane at various time steps, comparing DST3D-UNet-SR and HRTM predictions (two top rows) with ground truth (third row). The errors are also shown in the figure (two bottom rows). The DST3D-UNet-SR error is calculated as the difference between DST3D-UNet-SR predictions and ground truth. Similarly, the HRTM error is calculated as the difference between HRTM predictions and ground truth. Variations are shown across the x and z dimensions within a range of 0 to 5 km and the z -dimension ranging from 0 to 2 km.

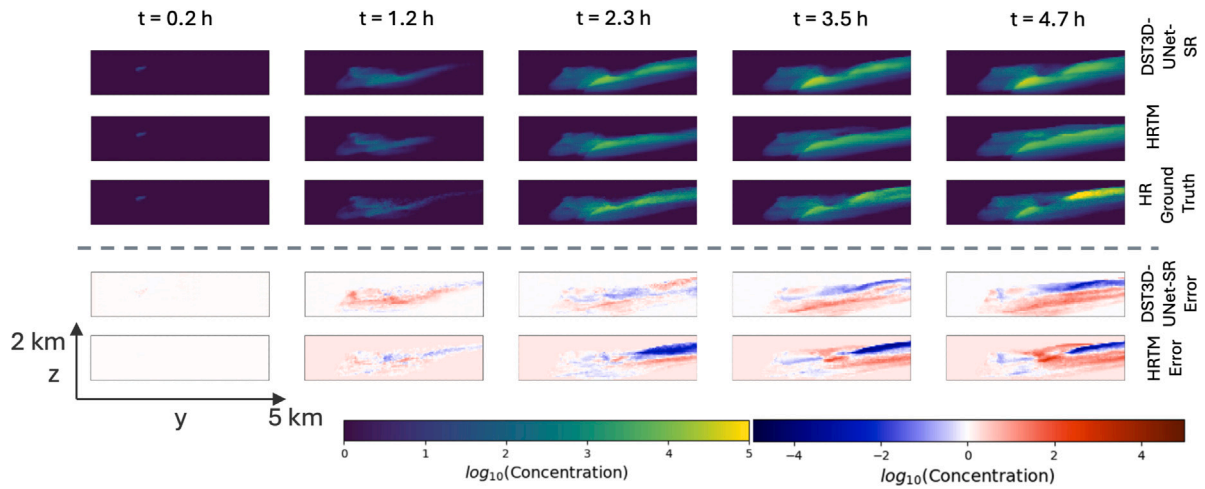


Fig. 5. Averaged concentration in the y - z plane at various time steps, comparing DST3D-UNet-SR and HRTM predictions (top two rows) with ground truth (third row). The errors are also shown in the figure (two bottom rows). The DST3D-UNet-SR error is calculated as the difference between DST3D-UNet-SR predictions and ground truth. Similarly, the HRTM error is calculated as the difference between HRTM predictions and ground truth. Variations are shown across the y and z dimensions within a range of 0 to 5 km and the z -dimension ranging from 0 to 2 km.

spatial resolution separately, starting with coarse temporal predictions and later refining them, allows it to avoid the common issues of noise and overemphasis on low-concentration regions.

Additionally, Fig. 7 presents slices along the z -direction, captured 120 min after the gas release. These slices compare the DST3D-UNet-SR predictions (left column) with the ground truth (middle column). The right column displays the error, calculated as the difference between the predictions and the ground truth. This figure demonstrates variations across the x - and y -dimensions. In the z -direction, the vertical extent of 2 km is represented with slice intervals of 0.125 km.

6. Performance metrics

To quantify the performance of the models, we present four metrics: mean squared error (MSE), intersection over union (IoU), structural similarity index measure (SSIM), and conservation of mass (CM). Each of these metrics provides insight into different aspects of the model performance. While MSE and CM focus on the accuracy of the predicted values, IoU and SSIM assess the model's ability to accurately identify

the average plume location and outline. All metrics are computed across the entire volume and over time.

MSE measures the average squared difference between predicted and true values, indicating overall accuracy (Goodfellow et al., 2016). IoU measures the overlap between predicted and ground truth binary masks; in our data, a reasonable threshold that preserved the plume outline was $\log_{10}(\text{Concentration}) = 1$, reflecting the model's ability to identify the correct plume location (Jaccard, 1912). SSIM assesses the similarity between predicted and true images based on luminance, contrast, and structure, providing a measure of perceptual quality (Wang et al., 2004). In our 3D application, SSIM is computed slice-wise on each horizontal cross-section (x - y plane) of the volumetric field. The resulting 2D SSIM values are averaged over all slices at each time step to yield a single SSIM score per frame. SSIM has been used extensively for various 3D deep learning problems involving medical imaging (Man et al., 2023) and turbulent flows (Chung et al., 2023). CM calculates the difference in total mass between predictions and ground truth, ensuring mass conservation in the model (LeVeque, 2002).

Fig. 8 illustrates these performance metrics over time for both the DST3D-UNet-SR and the HRTM.

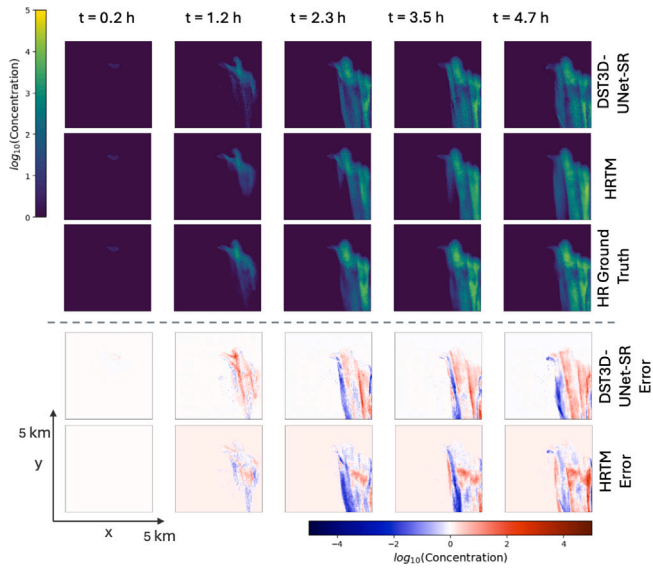


Fig. 6. Averaged concentration in the x-y plane at various time steps, comparing DST3D-UNet-SR and HRTM predictions (top two rows) with ground truth (third row). The errors are also shown in the figure (two bottom rows). The DST3D-UNet-SR error is calculated as the difference between DST3D-UNet-SR predictions and ground truth. Similarly, the HRTM error is calculated as the difference between HRTM predictions and ground truth. Variations are depicted across the x-direction and y-direction within ranges of 0 to 5 km.

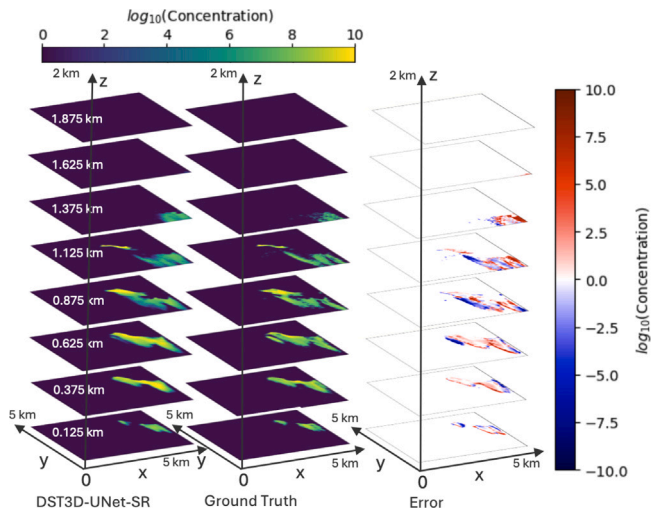


Fig. 7. Slices along the z-direction, 120 min post-gas release, illustrating DST3D-UNet-SR predictions (left) versus ground truth (middle). The error shown in the figure (right) is calculated as the difference between DST3D-UNet-SR predictions and ground truth. Variations are shown across the x and y dimensions within a range of 0 to 5 km and the z-dimension ranging from 0 to 2 km. The spatial resolution is 0.04 km for the x- and y-dimensions and 0.02 km for the z-dimension. Concentrations are recorded at the cell center. In the z-direction, a total vertical extent of 2 km is represented with slice intervals of 0.125 km.

The performance metrics averaged over 33 time steps across 10 test runs are summarized in Table 1.

Fig. 8 provides a detailed view of how these metrics evolve over time showing that the DST3D-UNet-SR model outperforms the HRTM overall, except during the initial five time steps when the HRTM benefits from high-resolution ground truth data. However, once the models rely on predictions alone, the DST3D-UNet-SR demonstrates superior performance across multiple metrics.

The DST3D-UNet-SR model effectively captures the 3D nature of plume dispersion through its convolutional operations, while the LSTM

Table 1

Performance metrics averaged over 33 time steps across 10 test runs. The arrow direction indicates whether higher or lower metric values represent better performance. Note that the HRTM performance is overall inferior, despite the first 5 time steps being high-resolution ground truth for this model.

Model/Metric	MSE ↓	IoU ↑	SSIM ↑	CM ↓
DST3D-UNet-SR	1.2 ± 0.4	0.62 ± 0.08	0.84 ± 0.03	0.3 ± 0.3
HRTM	1.5 ± 0.4	0.62 ± 0.09	0.51 ± 0.02	0.6 ± 0.8

layers accurately learn and predict the temporal evolution of the plume for up to five hours. The separation of temporal and spatial refinement helps avoid the error amplification that is more prominent in the HRTM model. This advantage is particularly evident in the consistently lower MSE values and higher SSIM scores of the DST3D-UNet-SR compared to the HRTM.

The MSE plot reveals two distinct slopes: an initial sharp rise as both models transition from relying on ground truth to making predictions, followed by a more gradual slope where the DST3D-UNet-SR's errors grow slower than the HRTM's, indicating better control over error propagation. Despite similar spikes in MSE around the 2-h mark, the DST3D-UNet-SR maintains a lower overall error due to better generalization across time steps.

In the initial stages of dispersion, the concentration gradients are larger, making it more challenging to capture the plume's exact location. In this phase, even small shifts in the predicted plume's position can lead to a significant reduction in overlap, causing the IoU to drop sharply. Nevertheless, the IoU scores indicate that both models perform similarly in terms of spatial overlap with the ground truth, although DST3D-UNet-SR maintains slightly better performance over time.

For the SSIM metric both models show a similar abrupt decline as they struggle to maintain the correct spatial patterns and texture of the plume during the initial transition. Rapid changes in concentration gradients and plume morphology can lead to distortions in the predicted structure, reducing the SSIM score. However, the decline for the DST3D-UNet-SR is less severe. Both models' SSIM values plateau after one hour, but the DST3D-UNet-SR significantly outperforms the HRTM.

With respect to CM, the DST3D-UNet-SR again shows superior and more robust performance in maintaining the physical properties of the plume over extended periods. The spikes in CM seen in both models around the 1-h mark highlight the models' responses to dynamic plume behavior. While the DST3D-UNet-SR also experiences a spike, it is less pronounced than that of the HRTM. This is because the SRM controls the spread of the plume by focusing on enhancing spatial details in regions where the temporal behavior is already well-predicted. The dual-stage process allows the model to mitigate large errors that could arise from incorrect predictions in peripheral regions.

In terms of computational efficiency, the DST3D-UNet-SR model demonstrates a significant advantage over the HRTM, achieving single-time-step computation times nearly five times faster (0.35 ± 0.01 s for DST3D-UNet-SR, compared to 1.58 ± 0.01 s for HRTM), with a total parameter count of ≈ 4.15 million. This efficiency corresponds to a speedup of three orders of magnitude relative to the original simulation runtime (2.25 h), compared to the two orders of magnitude improvement achieved by the HRTM. These attributes make DST3D-UNet-SR model an excellent choice for rapid response scenarios or iterative processes requiring a large number of simulations, where both speed and accuracy are critical.

7. Validation against simulation sensor data

To further assess the predictive accuracy of the DST3D-UNet-SR and HRTM models, we compared their outputs to the sensor data generated by Aeolus simulations described in Section 3. Although direct comparison with experiments is not possible due to the sparsity of field data, the agreement of simulation conditions with the REACT

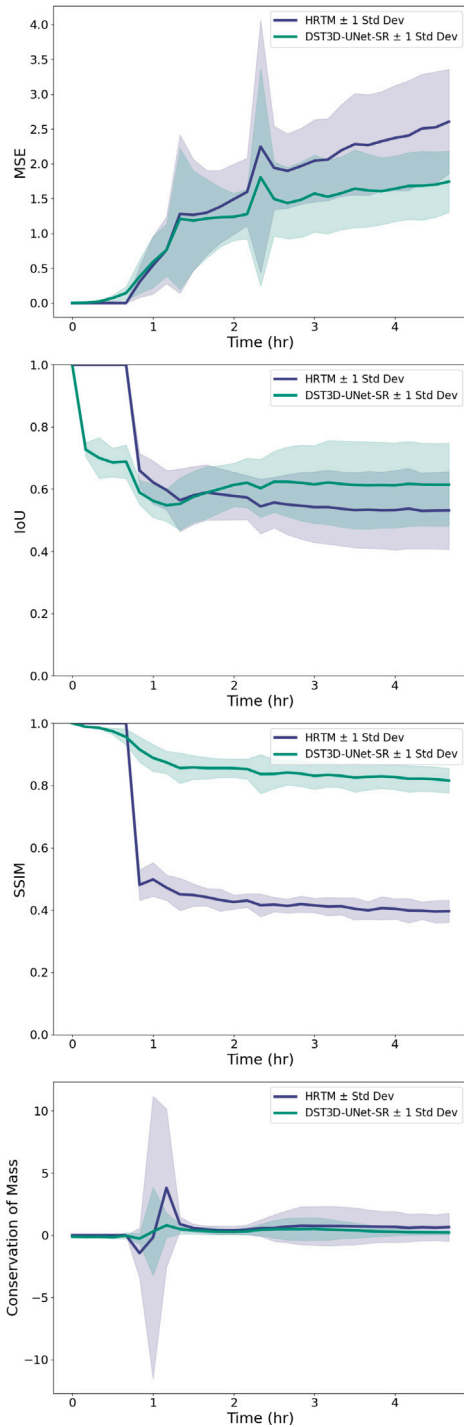


Fig. 8. Performance metrics over time for DST3D-UNet-SR and HRTM. The metrics include MSE, IoU, SSIM, and CM. The continuous line represents the mean performance across 10 test simulations, while the shaded regions denote ± 1 standard deviation from the mean.

experiment provides a robust reference point for evaluating model accuracy.

In these evaluations, sensor locations within the Aeolus simulation grid were identified and treated as ground truth references (see Fig. 9). Since sensor height was not accounted for, concentration predictions were averaged across the vertical (z) dimension. To evaluate the models' predictive accuracy at the sensor sites, concentration values were compared at discrete locations over time rather than across the

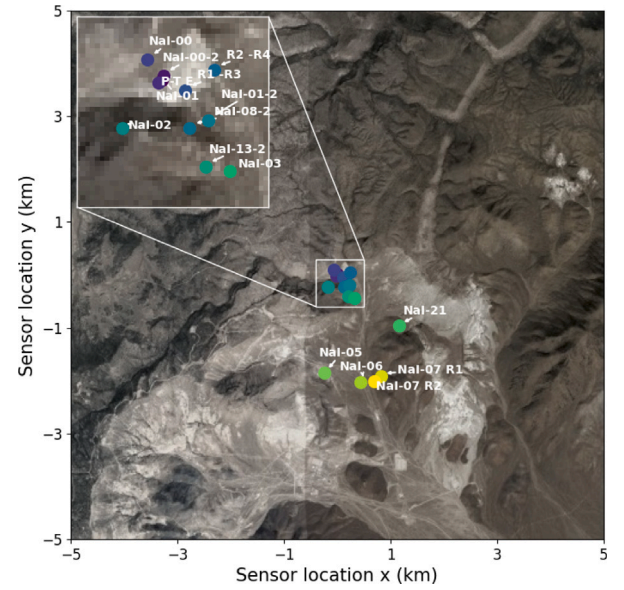


Fig. 9. Real-time xenon sensors were deployed to monitor radiotracer releases at the Nevada National Security Site (NNSS) during the REACT (Release ACTivity) experiment in October 2022.

entire field, enabling a focused analysis of model performance under conditions replicating actual sensor placements during the experiment. Notably, as the models were originally trained on Aeolus simulations, no additional training or new simulations were required for this study.

Fig. 10 shows the z -averaged $\log_{10}(\text{concentration})$ detected by each sensor over time, reported as the mean across the 10 test runs. The gray dashed line, approximately 2.5 h after the release, acts as a qualitative threshold indicating the transition from reliable to less reliable model predictions. The figure includes ground truth values based on Aeolus simulations (purple), predictions from DST3D-UNet-SR (blue), and predictions from HRTM (green). Additionally, the updated DST3D-UNet-SR (yellow) represents predictions generated during testing by reintroducing full DuoStage model inputs at three specific time steps: 1 h, 1.5 h, and 2.5 h after release. These inputs reflect full volumetric fields rather than single-point sensor data, and were selected to evaluate the model's responsiveness to periodic information updates during later stages of plume development. The update times were chosen to span the transition from the model's peak performance period into the pseudo-steady-state phase (as marked by the gray dashed line in Fig. 10). While additional or fewer updates could be explored, this choice provided a practical trade-off between model responsiveness and operational feasibility. A systematic study of update frequency and its impact on model accuracy is left for future work. This approach demonstrates the model's potential for real-time refinement without formal retraining or assimilation.

For sensors closest to the source (0 km to 0.54 km), both models initially agree closely with the ground truth during the first two hours. Between 2 and 2.5 h, the DST3D-UNet-SR remains more accurate, closely following the ground truth, while the HRTM starts to under-predict. Beyond this point, DST3D-UNet-SR begins to slightly over-predict concentrations relative to the ground truth, while the HRTM continues to under-predict. This discrepancy is likely due to sensor concentrations being recorded at discrete spatial points, causing the HRTM to generally under-predict concentrations inside the plume while over-predicting concentrations outside of it, as discussed in Section 5. However, while the HRTM captures the concentration plateau reached after 2.5 h, DST3D-UNet-SR predictions continue to increase smoothly over time.

The updated DST3D-UNet-SR simulations, incorporating additional data points, show a smaller gap between predictions and the ground

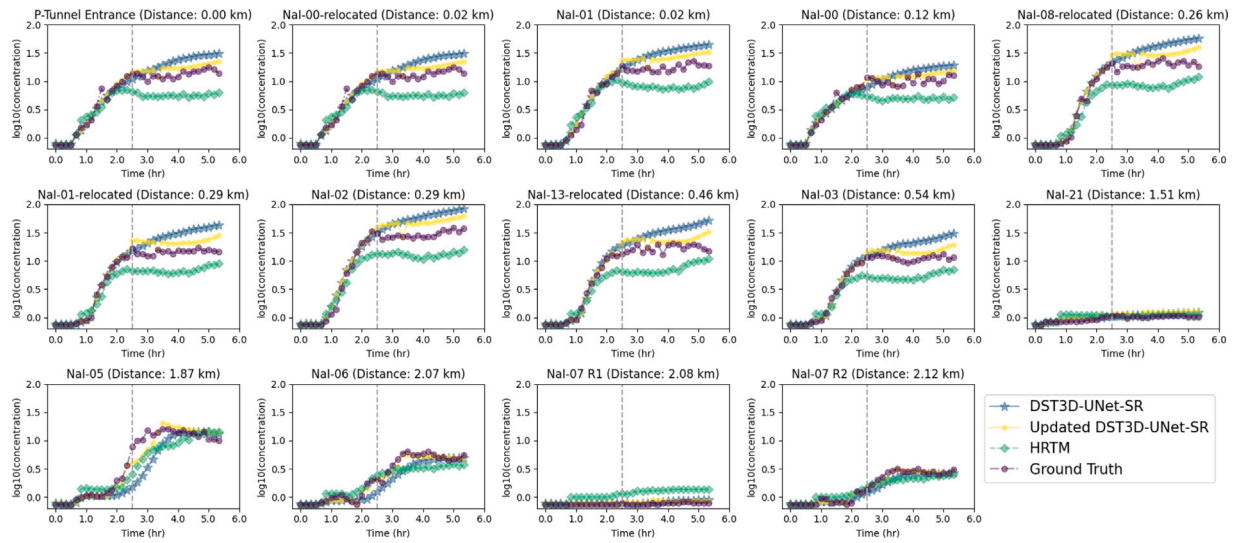


Fig. 10. Validation of DST3D-UNet-SR and HRTM performance against Aeolus simulations (Ground Truth). The plot shows the z-averaged $\log_{10}(\text{concentration})$ detected by each sensor as a function of time, reported as the mean across the 10 test runs. The vertical gray dashed line qualitatively marks the separation between areas of good and poorer performance, approximately 2.5 h after the release. Updated DST3D-UNet-SR refers to the predictions generated by DST3D-UNet-SR when it is provided with three additional data points as inputs at specific time intervals.

truth. The updates also enable the model to capture the pseudo-steady-state concentration plateau reached after 2.5 h, as shown by the Aeolus simulations. While these updates improve predictions, particularly in high-concentration regions near the source, slight over-predictions persist. These updated predictions are included to illustrate the potential for improving model performance with new data. However, a detailed study into updating methods and their broader impacts is beyond the scope of this paper and will be addressed in future work.

For sensors located farther away (1.51 km to 2.12 km), both models exhibit comparable performance, effectively capturing late-time behavior with greater accuracy. This improved performance at larger distances is likely attributed to the plume becoming more diffuse, resulting in smoother concentration gradients that are easier to predict. Additionally, averaging over larger areas at these distances diminishes the impact of smaller-scale inaccuracies, enabling both models to align more closely with the ground truth. For these distant sensors, the updated DST3D-UNet-SR predictions demonstrate enhancements relative to the original DST3D-UNet-SR and HRTM, though the improvement is less pronounced compared to sensors closer to the source.

Because only a small fraction of physical sensors ever recorded plume hits, we validate using synthetic data from Aeolus LES by comparing model concentration forecasts to simulated sensor readings at discrete spatial and temporal locations. This setup provides a realistic assessment of the model's operational performance in the absence of dense in-situ observations. Future work will explore direct validation against experimental datasets when available.

8. Conclusion

The DST3D-UNet-SR model outperforms the baseline HRTM across all key metrics, including lower mean squared error (MSE) and conservation of mass (CM) deviations, alongside superior structural similarity index measure (SSIM) and intersection over union (IoU) scores. These results highlight its capacity to accurately replicate the spatiotemporal dynamics of atmospheric plume dispersion while adhering to fundamental physical conservation laws.

A significant advancement of the DST3D-UNet-SR is its computational efficiency, achieving a fivefold reduction in single-time-step computation time compared to HRTM and a 45% reduction in training time (5.5 to 3 h). This balance between computational speed and predictive fidelity underscores its suitability for applications requiring

rapid response, such as real-time plume dispersion modeling, iterative optimization, and uncertainty quantification.

We demonstrated that the model can dynamically incorporate observational updates during inference, enhancing predictive accuracy at later stages, particularly near the source where concentrations are highest. This capability highlights the potential of DST3D-UNet-SR to adapt to evolving conditions in real-world applications. Future work will explore the scalability and effectiveness of this adaptive approach across diverse scenarios.

Regarding additional comparisons with other deep-learning based super-resolution models, we recognize that these are important directions but fall outside the scope of the current work, which focuses on the dual-stage design tailored for 3D atmospheric transport. We plan to include such comparisons in future studies that extend the present framework across domains.

This study introduced DST3D-UNet-SR, a dual-stage deep learning model for accelerating LES-based predictions of 3D plume transport by up to three orders of magnitude. By independently learning temporal and spatial refinement, the model provides accurate reconstructions of high-resolution atmospheric fields from low-resolution inputs. We confirm via synthetic sensor tests the model's practical utility in emulating critical near-source dynamics. Future work will explore adaptation of the model to different physical settings and its integration with real-time forecasting tools to support emergency response and decision-making.

CRediT authorship contribution statement

M. Giselle Fernández-Godino: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Wai Tong Chung:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization. **Akshay A. Gowardhan:** Software, Conceptualization. **Matthias Ihme:** Writing – review & editing. **Qingkai Kong:** Writing – review & editing, Supervision, Conceptualization. **Donald D. Lucas:** Writing – review & editing, Supervision, Investigation, Conceptualization. **Stephen C. Myers:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Lawrence Livermore National Laboratory's LivChat based on OpenAI's GPT-4 in order to ensure that the writing meets academic standards and is free of errors. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Institutional release number LLNL-JRNL-2001564. This Low Yield Nuclear Monitoring (LYNM) research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, NNSA, PNNL, and SNL.

Appendix A. Dual-Stage Temporal 3D UNet-SR model architecture

The Dual-Stage Temporal 3D UNet-SR architecture is designed to enhance the temporal resolution and spatial detail of atmospheric dispersion simulations. The model consists of two stages: a low-resolution-based temporal module (TM) and a high-resolution-based spatial refinement module (SRM). Further details are included in [Appendices A.1](#) and [A.2](#).

A.1. Temporal module

The TM is constructed using a convolutional autoencoder architecture, incorporating ConvLSTM layers at its bottleneck to capture temporal dependencies within the input data. Its architecture comprises 3,214,401 trainable parameters. It operates on input data of shape (batch size, 5, 8, 32, 32), representing a temporal window of five time steps, each with a spatial resolution of $8 \times 32 \times 32$.

Encoder:

- **enc1:** Conv3d(5, 7×16 , 3, padding = 1) followed by BatchNorm3d, ReLU activation, MaxPool3d(2), and Dropout3d(0.2).
- **enc2:** Conv3d(7×16 , 7×32 , 3, padding = 1) followed by BatchNorm3d, ReLU activation, MaxPool3d(2), and Dropout3d(0.2).
- **enc3:** Conv3d(7×32 , 7×64 , 3, padding = 1) followed by BatchNorm3d(2), ReLU activation, MaxPool3d(2), and Dropout3d(0.2).

ConvLSTM:

- **convlstm:** Conv3d(7×64 , 7×64 , 3, padding = 1).

Decoder:

- **dec1:** ConvTranspose3d(7×64 , 7×32 , 2, stride = 2) followed by BatchNorm3d, ReLU activation, and Dropout3d(0.2).
- **dec2:** ConvTranspose3d(7×32 , 7×16 , 2, stride = 2) followed by BatchNorm3d, ReLU activation, and Dropout3d(0.2).
- **dec3:** ConvTranspose3d(7×16 , 1, 2, stride = 2) followed by ReLU activation.

The TM generates single time-step predictions with an output shape of (batch size, 1, 8, 32, 32), which are recursively fed back into the model in a sliding window framework consisting of five time steps.

A.2. Spatial refinement module

The SRM processes the TM output, reshaping it to (batch size, 8, 32, 32) and produces high-resolution output of shape (batch size, 32, 128, 128). Its architecture has 951,873 trainable parameters.

Encoder:

- **enc1:** Conv3d(1, 7×16 , 3, padding = 1) followed by BatchNorm3d, LeakyReLU activation, and MaxPool3d.
- **enc2:** Conv3d(7×16 , 7×32 , 3, padding = 1) followed by BatchNorm3d, LeakyReLU activation, and MaxPool3d.

Adjust Channels for Skip Connections:

- **adjust_channels:** Conv3d(7×16 , 7×32 , 1) followed by BatchNorm3d and LeakyReLU activation.

Decoder:

- **dec1:** ConvTranspose3d(7×32 , 7×32 , 3, stride = 2, padding = 1, output_padding = 1) followed by BatchNorm3d and LeakyReLU activation.
- **dec2:** ConvTranspose3d(7×32 , 7×16 , 3, stride = 2, padding = 1, output_padding = 1) followed by BatchNorm3d and LeakyReLU activation.
- **dec3:** ConvTranspose3d(7×16 , 7×8 , 3, stride = 2, padding = 1, output_padding = (1, 1, 1)) followed by BatchNorm3d and LeakyReLU activation.
- **dec4:** ConvTranspose3d(7×8 , 1, (1, 3, 3), stride = (1, 2, 2), padding = (0, 1, 1), output_padding = (0, 1, 1)) followed by LeakyReLU activation.

Appendix B. High-resolution temporal module architecture

The high-resolution temporal model (HRTM) employs a convolutional autoencoder architecture, similar to the TM, with ConvLSTM layers integrated at the bottleneck to effectively capture temporal dependencies within the input data. The HRTM is composed of 4,166,274 trainable parameters. The model operates on input data of shape (batch size, 5, 32, 128, 128), representing a temporal window of five time steps, each with a spatial resolution of $32 \times 128 \times 128$.

Encoder:

- **enc1:** Conv3d(5, num_layers \times 32, 3, padding = 1) followed by BatchNorm3d, ReLU activation, MaxPool3d, and Dropout3d.
- **enc2:** Conv3d(num_layers \times 32, num_layers \times 64, 3, padding = 1) followed by BatchNorm3d, ReLU activation, MaxPool3d, and Dropout3d.
- **enc3:** Conv3d(num_layers \times 64, num_layers \times 128, 3, padding = 1) followed by BatchNorm3d, ReLU activation, MaxPool3d, and Dropout3d.

ConvLSTM:

- **convlstm:** Conv3d(num_layers \times 128, num_layers \times 128, 3, padding = 1).

Decoder:

- **dec1:** ConvTranspose3d(num_layers \times 128, num_layers \times 64, 2, stride = 2) followed by BatchNorm3d, ReLU activation, and Dropout3d.
- **dec2:** ConvTranspose3d(num_layers \times 64, num_layers \times 32, 2, stride = 2) followed by BatchNorm3d, ReLU activation, and Dropout3d.
- **dec3:** ConvTranspose3d(num_layers \times 32, 1, 2, stride = 2) followed by ReLU activation.

Like the TM, the HRTM produces single time-step predictions with an output shape of (batch size, 1, 32, 128, 128), which are recursively fed back into the model in a sliding window framework consisting of five time steps.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Camps-Valls, G., Tuia, D., Zhu, X.X., Reichstein, M., 2021. Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science and Geosciences. John Wiley & Sons.
- Chung, W.T., Akoush, B., Sharma, P., Tamkin, A., Jung, K.S., Chen, J.H., Guo, J., Brouzet, D., Talei, M., Savard, B., Poludnenko, A.Y., Ihme, M., 2023. Turbulence in focus: Benchmarking scaling behavior of 3D volumetric super-resolution with blastnet 2.0 data. *Adv. Neural Inf. Process. Syst. (NeurIPS)* 36.
- Ding, Y., Zhang, X., Zhang, Y., Ren, Y., Wei, H., Jia, D., 2021. A hybrid CNN-LSTM model for predicting PM2.5 concentration using spatiotemporal correlations between cities. *Sci. Total Environ.* 792, 148413.
- Dong, C., Loy, C.C., He, K., Tang, X., 2015. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (2), 295–307.
- Fernández-Godino, M.G., Balachandrar, S., Haftka, R.T., 2019. On the use of symmetries in building surrogate models. *J. Mech. Des.* 141 (6), 061402.
- Fernández-Godino, M.G., Lucas, D.D., Kong, Q., 2023. Predicting wind-driven spatial deposition through simulated color images using deep autoencoders. *Sci. Rep.* 13 (1), 1394.
- Fernández-Godino, M.G., Panda, N., O'Malley, D., Larkin, K., Hunter, A., Haftka, R.T., Srinivasan, G., 2021. Accelerating high-strain continuum-scale brittle fracture simulations with machine learning. *Comput. Mater. Sci.* 186, 109959.
- García-Cardona, C., Fernández-Godino, M.G., O'Malley, D., Bhattacharya, T., 2022. Uncertainty bounds for multivariate machine learning predictions on high-strain brittle fracture. *Comput. Mater. Sci.* 201, 110883.
- Geiss, A., Silva, S.J., Hardin, J.C., 2022. Downscaling atmospheric chemistry simulations with physically consistent deep learning. *Geosci. Model. Dev.* 15 (17), 6677–6694.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.
- Gowardhan, A.A., McGuffin, D.L., Lucas, D.D., Neuscamman, S.J., Alvarez, O., Glascoe, L.G., 2021. Large eddy simulations of turbulent and buoyant flows in urban and complex terrain areas using the aeolus model. *Atmos.* 12 (9), 1107.
- Huot, F., Hu, R.L., Goyal, N., Sankar, T., Ihme, M., Chen, Y.-F., 2022. Next day wildfire spread: A machine learning dataset to predict wildfire spreading from remote-sensing data. *IEEE Trans. Geosci. Remote Sens.* 60, 1–13.
- Jaccard, P., 1912. The distribution of the flora in the alpine zone. 1. *New Phytol.* 11 (2), 37–50.
- Jiang, K., Wang, Z., Yi, P., Lu, T., Jiang, J., Xiong, Z., 2020. Dual-path deep fusion network for face image hallucination. *IEEE Trans. Neural Netw. Learn. Syst.* 33 (1), 378–391.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., Schindler, K., 2018. Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS J. Photogramm. Remote Sens.* 146, 305–319.
- LeVeque, R.J., 2002. Finite Volume Methods for Hyperbolic Problems, vol. 31, Cambridge University Press.
- Liebel, L., Körner, M., 2016. Single-image super resolution for multispectral remote sensing data using convolutional neural networks. In: XXIII ISPRS Congress, Commission VII. International Society for Photogrammetry and Remote Sensing (ISPRS), pp. 883–890.
- Liu, X.-Y., Zhu, M., Lu, L., Sun, H., Wang, J.-X., 2024. Multi-resolution partial differential equations preserved learning framework for spatiotemporal dynamics. *Commun. Phys.* 7 (1), 31.
- Man, C., Lau, V., Su, S., Zhao, Y., Xiao, L., Ding, Y., Leung, G.K., Leong, A.T., Wu, E.X., 2023. Deep learning enabled fast 3D brain MRI at 0.055 tesla. *Sci. Adv.* 9 (38), eadi9327.
- Marcos Filho, A., Santos, M.C., Pereira, C.M., 2024. Exploring LSTM-based prediction for radioactive plume atmospheric dispersion in nuclear power plant emergencies: A preliminary study. *Nucl. Eng. Des.* 419, 112982.
- Stachenfeld, K., Fielding, D.B., Kochkov, D., Cranmer, M., Pfaff, T., Godwin, J., Cui, C., Ho, S., Battaglia, P., Sanchez-Gonzalez, A., 2022. Learned coarse models for efficient turbulence simulation. In: Proceedings of the International Conference on Learning Representations. ICLR.
- Stave, S., Bertschinger, K., Emmons, S., Becker, E., Fast, J., Siciliano, E., Zalavadia, M., 2024. Real-Time Xenon Sensor Analysis Report PNNL-35939. Tech. Rep., Pacific Northwest National Laboratory.
- Stein, M., 1987. Large sample properties of simulations using Latin hypercube sampling. *Technometrics* 29 (2), 143–151.
- Tsokov, S., Lazarova, M., Aleksieva-Petrova, A., 2022. A hybrid spatiotemporal deep model based on CNN and LSTM for air pollution prediction. *Sustain.* 14 (9), 5104.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, Y., Fernández-Godino, M.G., Gunawardena, N., Lucas, D.D., Yue, X., 2024. Spatiotemporal predictions of toxic urban plumes using deep learning. *arXiv preprint arXiv:2406.02582*.
- Wang, Y., Oyen, D., Guo, W., Mehta, A., Scott, C.B., Panda, N., Fernández-Godino, M.G., Srinivasan, G., Yue, X., 2021. Stressnet-deep learning to predict stress with fracture propagation in brittle materials. *Npj Mater. Degrad.* 5 (1), 6.
- Wharton, S., Brown, M.J., Dexheimer, D., Fast, J.D., Newsom, R.K., Schalk, W.W., Wiersema, D.J., 2023. Capturing plume behavior in complex terrain: an overview of the nevada national security site meteorological experiment (METEX21). *Front. Earth Sci.* 11, 1251153.
- Wu, Z., Zhu, F., Guo, K., Sheng, R., Chao, L., Fang, H., 2024. Modal adaptive super-resolution for medical images via continual learning. *Signal Process.* 217, 109342.
- Xiao, Y., Wang, Y., Yuan, Q., He, J., Zhang, L., 2022. Generating a long-term (2003–2020) hourly 0.25° global PM2.5 dataset via spatiotemporal downscaling of CAMS with deep learning (DeepCAMS). *Sci. Total Environ.* 848, 157747.
- Xiao, Y., Yuan, Q., Jiang, K., He, J., Jin, X., Zhang, L., 2023. EDiffSR: An efficient diffusion probabilistic model for remote sensing image super-resolution. *IEEE Trans. Geosci. Remote Sens.* 62, 1–14.
- Yu, M., Liu, Q., 2021. Deep learning-based downscaling of tropospheric nitrogen dioxide using ground-level and satellite observations. *Sci. Total Environ.* 773, 145145.