Deep Learning Aerosol-Cloud Interactions from Satellite Imagery

Pierce Warburton, Kurtis Shuler, Lekha Patel

Sandia National Laboratories, PO Box 5800 MS 1202, Albuquerque, NM 87185
pdwarbu@sandia.gov, kwshule@sandia.gov, lpatel@sandia.gov

Abstract

Satellite imagery can detect a wealth of ship tracks, temporary cloud trails created via cloud seeding by the emitted aerosols of large ships, a phenomenon that cannot be directly reproduced by global climate models. Ship tracks are satellite-observable examples of aerosol-cloud interactions, processes that constitute the largest uncertainty in climate forcing predictions, and when observed are also examples of Marine Cloud Brightening (MCB), a potential climate intervention strategy. Leveraging the large amount of observed ship track data to enhance understanding of aerosol-cloud interactions and the potentials of MCB is hindered by the computational infeasiblity of characterization from expensive physical models. In this paper, we focus on utilizing a cheaper physics-informed advection-diffusion surrogate to accurately emulate ship track behavior. As an indication of aerosol-cloud interaction behavior, we focus on learning the spreading behavior of ship tracks, neatly encoded in the emulator's spatio-temporal diffusion field. We train a convolutional LSTM to accurately learn the spreading behavior of simulated and satellite-masked ship tracks and discuss its potential in larger scale studies.

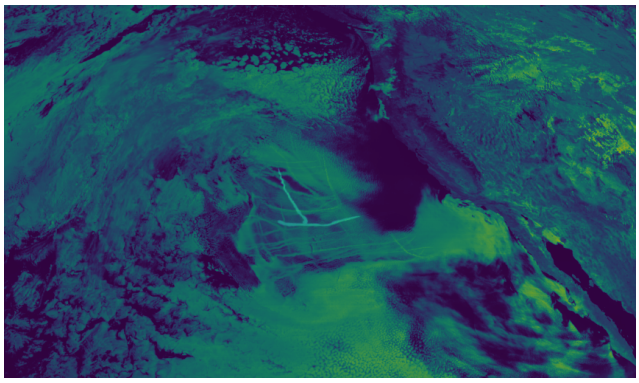

Figure 1: Ship tracks are long thin artificial cloud features caused by ship exhaust emissions and are observable in marine environments. Tracks highlight both aerosol-cloud interactions and cloud brightening, a phenomenon that could be artificially recreated to combat global warming. Image shown was observed on July 24 2019 at 18:00 off the coast of California via NOAA's GOES-R satellite imager.

Introduction

For decades, satellite imagery has been able to detect ship tracks, temporary cloud trails created via cloud seeding by the emitted aerosols of large ships traversing our oceans (see Figure 1), a phenomenon that global climate models cannot directly reproduce. Ship tracks are of interest because they are satellite-observable examples of *aerosol-cloud interactions*, processes that constitute the largest uncertainty in climate forcing predictions from global climate models (Stocker et al. 2013). Specifically, ship tracks are visible evidence of the ability of large amounts of aerosol emissions to perturb boundary layer clouds enough to alter the albedo of the atmosphere and thus significantly contribute to indirect radiative forcing (Capaldo et al. 1999; Eyring et al. 2010). In fact, bright ship track observations are examples of Marine Cloud Brightening (MCB) (Latham et al. 2012), a geoengineering technique that is currently being considered as a climate intervention strategy to halt the effects of global warming. While most analyses of cloud microphysical changes induced by large aerosol injections come from unrealistic simulations under pristine conditions (Blossey, Bretherton, and Mohrmann 2021), satellite-observed tracks form in complex dynamic environments that are challenging and expensive to physically replicate. Further, satellite retrieval products do not provide a holistic view of these changes at the high resolutions needed for truly understanding this behavior, motivating surrogate representations in analyzing and predicting the effects of ship tracks on the surrounding clouds, observed from satellite.

In particular, being able to predict how ship tracks disperse in space and time has many key impacts in the climate and atmospheric community. First, it aids analysis and understanding of cloud reaction from aerosol injection. For example, if the dispersion field is seen to differ in certain case-studies, this provides quantitative information on the potential to use aerosol injection for MCB. Slower dispersion parameters would indicate an atmospheric region that both promotes the brightening of clouds, and their longevity, that could inform climate intervention researchers on when and where to inject aerosols for limiting the solar radiation impact in that region. Second, understanding dispersion

across ship tracks, allows for a better understanding of its effects with the surrounding clouds, including exact spatial regions in which interactions are occurring more intensely than others. When quantifying with respect to the surrounding clouds, this would provide atmospheric researchers cases of the types of clouds, and their properties, which yield stronger interactions with injected aerosols. Since aerosol-cloud interactions are still poorly captured in global climate models, this would provide a greater understanding on the atmospheric and cloud conditions that should be studied further. Moreover, a fast Machine Learning (ML) framework to estimate dispersion parameters could be used to inform existing but expensive climate models' computational resource allocations, enabling enhanced understanding of the climate effects that stem from such interactions.

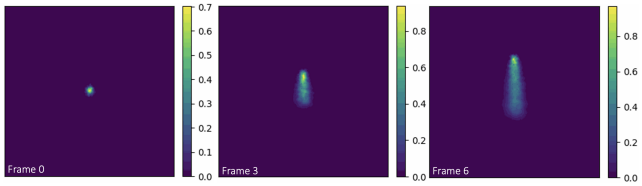

Figure 2: Cheap emulations of a ship track at ten minute intervals under a advection-diffusion model that shows normalized dispersion over space and time, as informed by Large Eddy Simulations (LES).

Model

To model the formation and behavior of ship tracks, we follow the reduced-order advection-diffusion stochastic differential parameterizaton of Patel and Shand 2022 that relates true aerosol emission trajectories with imaging observations captured by satellite over imaging window $\mathcal{X} \subset \mathbb{R}^2$ and time window $\mathcal{T} \subset \mathbb{R}_{\geq 0}$. This is conducted in a manner that models each aerosol parcel as a Lagrangian particle $\mathbf{x}_t$, at spatial coordinate $\mathbf{x} \in \mathcal{X}$ at time $t \in \mathcal{T}$ that follows the dynamics of the corresponding stochastic PDE. Specifically, after emission, each Lagrangian aerosol particle drifts according to an advection function $\mu(\mathbf{x}_t, t)$, which is described by the wind velocity at time t, and a *dispersion* function $D(\mathbf{x}_t, t)$ which describes the dispersion of the particle at spatial coordinate $\mathbf{x} \in \mathcal{X}$ at time $t \in \mathcal{T}$, within the surrounding clouds. As a generalization of the work of Patel and Shand 2022, this evolution is encoded in the following stochastic differential equation:

$$\mathrm{d}\mathbf{x}_t = \mu(\mathbf{x}_t, t)\mathrm{d}t + D(\mathbf{x}_t, t)\mathrm{d}W_t, \tag{1}$$

where $W_t \sim \mathcal{N}_2(\mathbf{0}, t\mathcal{I}_2)$ and $\mathcal{I}_2$ denotes the 2-dimensional identity matrix.

Since wind velocity can be captured by climate reanalysis, the advection function μ is seen as a known quantity in the above model, as informed by observed wind fields in applications to real data. However, since turbulence parameters are generally more challenging to accurately capture from high-resolution observational data, we formulate the spatio-temporal *dispersion* quantity as above, to supersede the frequently defined *diffusion* function. This modeling mechanism is used since *diffusion* in this context typically describes small-scale eddy diffusivities, whereas we would like to capture small scale diffusivities and larger-scale turbulent effects, as has been newly informed by LES simulations of ship tracks (McMichael et al. 2023+), to effectively describe aerosol spreading and corresponding interactions within the clouds.

To be able to analyze imaging observations of ship tracks, a large number of Lagrangian particles following Equation (1) can be simulated using the emission setup of Patel and Shand 2022, where the Euler-Mayurama approach can propagate Lagrangian particles from discrete time frame $n\Delta$ to $(n+1)\Delta$ via

$$\mathbf{x}_{(n+1)\Delta} = \mathbf{x}_{n\Delta} + \mu(\mathbf{x}_{n\Delta}, n\Delta)\Delta + \sqrt{\Delta}D(\mathbf{x}_{n\Delta}, n\Delta)\boldsymbol{\xi}, \tag{2}$$

where $\boldsymbol{\xi} \sim \mathcal{N}(0, \mathcal{I}_2)$. Pixel intensities can be subsequently calculated using a Gaussian density approximation of particle concentration in each pixel. Each pixel denotes an effective relative radiance along the track and is normalized to being in $[0, 1]$ for accurate comparability to satellite data, provided background clouds are removed from the images. However, by pixelating concentrations, information about each Lagrangian particle is lost, and renders learning the dispersion fields difficult from images alone.

Under the simplified case where $D(\mathbf{x}, t) = \sigma$ and σ is a constant dispersion parameter, Patel and Shand 2022 present a unsupervised statistical learning framework to learn σ. However, the authors discuss huge computational difficulties when attempting to learn interaction behavior from large scale observational data. These challenges motivate the use of computationally efficient deep learning architectures, however, such models require training data that can accurately capture the dynamics of real ship tracks.

To do so, we consider the LES from McMichael et al. 2023+, and can compute an approximate dispersion field that can be calculated by applying the Euler-Mayurama scheme backwards in time. Specifically, since LES contains aerosol concentration in a spatio-temporal field, along with corresponding wind fields, dispersion fields can be estimated through the covariance matrix of the backward wind-advected aerosol field evaluated at each time step. Due to the computational burden of running LES to extract dispersion fields, we use our studies to inform the dispersion behavior from that observed from LES.

Under the advection-diffusion parameterization, such simulations can be used to study the parameterization or form of D. Specifically, using Equation 2, we define the *wind differenced* estimator of $\mathbf{x}_{(n+1)\Delta}$ as

$$\hat{\mathbf{x}}_{(n+1)\Delta} = \mathbf{x}_{n\Delta} + \mu(\mathbf{x}_{n\Delta}, n\Delta)\Delta, \tag{3}$$

and find

$$\Delta^{-1/2}(\mathbf{x}_{(n+1)\Delta} - \hat{\mathbf{x}}_{(n+1)\Delta}) \sim \mathcal{N}(0, D(\mathbf{x}_{n\Delta}, n\Delta)).$$

Here, an estimate for $D(\mathbf{x}_{n\Delta}, n\Delta)$ can be obtained by considering the empirical covariance matrix of the wind-differenced Lagrangian particle positions.

Preliminary analysis has shown a spatially-varying dispersion field that is effectively constant across the ship track along specific time periods, but which increases somewhat linearly along the track close to the emission source, where interactions with older aerosols are heightened, and which eventually decays to a constant value as spreading subsides. To establish proof of concept in utilizing an ML model, in the simulations that follow, we choose $D(\mathbf{x}, t) \propto \begin{pmatrix} \sigma_x + c_{x_t} x & 0 \\ 0 & \sigma_y + c_{y_t} y \end{pmatrix}$, where σ_x, σ_y are constants and c_{x_t}, c_{y_t} are time-varying sequences of constants. Here, the time index informs the strength of linear dispersion, expected to decay towards zero as t grows large, as observed by LES through eventual stabilization in mixing (McMichael et al. 2023+). Further σ_x and σ_y denote baseline constant dispersion values along each axis. We additionally note that the track can be rotated across the grid so that across and along–track effects can be visualized. The inclusion of higher-resolved dispersion fields and cross-dispersion behavior is left for future work.

An emulation example of a ship track under this parameterization is shown in Figure 2, where frames are 10 minutes apart.

Architecture

We now describe a novel physics-constrained ML model that can be used to learn the effective dispersion values from emulated video data, taken from the imaging framework of Patel and Shand 2022, using the dispersion setup described above.

To analyze the potential of supervised learning tools to this problem, we aim to learn dispersion values across a ship track, from 2D images over time. In this manner, we label ship track pixels with their relative dispersion values over an entire video to be learned. Since time is typically sampled at 10-30 minute intervals, we learn the effective dispersion field sampled at each frame to reflect temporal changes. To do so, we construct a convolutional long short-term memory framework (ConvLSTM), combining the temporal analysis of a bidirectional LSTM with three convolutional layers including two fully connected dense layers, that can be trained on the emulated data (shown in Figure 3). When labeling each video, two channels corresponding to a tensor describing the dispersion in the x and y directions are used at each time point of image creation; this is labeled with the computed values from D. Leveraging the bidirectional LSTM's internal weight adjustment through time, allows for the network to fully capture the temporal spreading dynamics, which can be challenging with other ML frameworks. Further, the loss function was uniquely crafted to capture the diffusive dynamics seen by LES, penalizing poorer predictions at the track edges and ensuring positive dispersion estimates per pixel. Learning rates, number of dense ConvLSTM layers, bidirectional layers, convolutional layers, including formations of how the forward and backward LSTMs are combined, optimizers and loss function parameters were all tuned to build the best predictive model during training.

Results

During the training procedure, 11 frames corresponding to 2 hours of a ship track formation were used. In total over 10,000 video emulations with variations of $\sigma_x, c_{x_t}, \sigma_y$ and c_{y_t}, dispersion as described via LES, and constant dispersion, were trained and tested on 100 variants (whose prediction metric was averaged across the testing data). These also varied the starting position of the ship tracks, and the corresponding wind fields, utilizing variations of the wind fields extracted from LES in some trained examples. A description of different models tested and their respective training validation losses and testing errors, computed by finding the mean squared error of predictions per pixel across the video, is shown in Table 1 below. For completeness, the constant–in both space and time– dispersion case was also tested by our model, and was important in deciding the final model for learning variable dispersion. Further, the singly varied dispersion accounts for simplified dispersion fields as described in Patel and Shand 2022, and time-limiting track behaviors as suggested by LES.

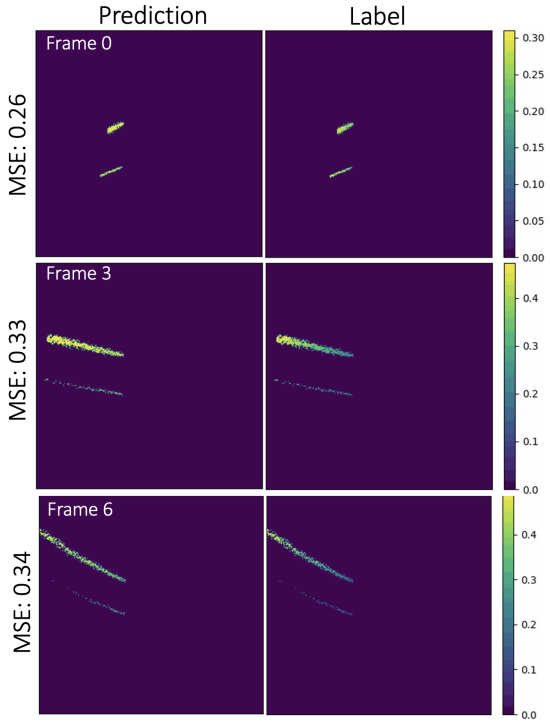

Figure 3: A trained bi-directional convolutional LSTM framework predicting two ship track dispersions over space and time over ten minute intervals.

Overall, we find good performance of the selected model in our studies, using the mean squared error metric per pixel over each frame and time-step. The bi-directionality of the LSTM makes a significant impact in more accurately predicting the corresponding dispersion fields, and in particular with the inclusion of extra dense layers in the corresponding CNN. The testing of different learning rates was also crucial in reducing the MSE over the final models chosen. Addition-

Scenario	Models	Metric (MSE)		
		Channel 1	Channel 2	Overall
1	Unidir	0.318	-	0.318
	Bidir	0.2409	-	0.2409
	3-bidir	**0.2042**	-	**0.2042**
2	3-bidir+av	0.1729	0.1729	0.1729
	3-bidir+mult	0.1383	0.1383	0.1383
	3-bidir+sum	0.1402	0.1361	0.13815
	3-bidir+concat	**0.0603**	**0.0617**	**0.061**
3	Uni+concat	0.312	-	0.312
	Bidir+concat	0.1762	-	0.1762
	3-bidir+concat	**0.1428**	-	**0.1428**
4	–"–+LR(10^{-1})	0.2199	0.2767	0.2483
	+LR(10^{-2})	**0.2134**	**0.2265**	**0.21995**
	+LR(10^{-4})	0.298	0.3272	0.3126
	+dropout	0.273	0.31	0.2915

Table 1: Performance of different Conv LSTM models, on dispersion estimation using two channels (for x and y respectively) and under four scenarios. Scenario 1 denotes constant dispersion learned in 1 channel (axis) only; i.e. that $c_{x_t} = c_{y_t} = 0$ for all t with σ_x learnt and σ_y known. Scenario 2 denotes constant dispersion learned in both channels (axes); i.e. that $c_{x_t} = c_{y_t} = 0$ for all t with both σ_x, σ_y learnt. Scenario 3 denotes variable dispersion learned in 1 channel (axis) only, with the other fixed at a constant value; i.e. that $c_{x_t} \neq 0, c_{y_t} = 0$ for all t with σ_x learnt and σ_y known. Scenario 4 denotes variable dispersion learned in both channels; i.e. that $c_{x_t} \neq 0, c_{y_t} \neq 0$ for all t with σ_x, σ_y both learnt. Three different Convolutional LSTM models were tested: a unidirectional (Unidir) LSTM with 1 layer (input ship track image, one hidden CNN layer and output dispersion field), Bidirectional (Bidir) LSTM with one hidden CNN layer and Bidirectional LSTM with 3 hidden CNN layers (3-bidir). Associated with the bidirectional LSTM, various modes to combine the forward and backward LSTM were tested: average (av), multiplication (mult), summation (sum) and concatenation (concat). In addition, various learning rates (LR) were tested, with the exclusion (dense) or inclusion of dropout. Over all tests, the mean squared error (MSE) between the input video labels and dispersion predictions, applied per pixel and per frame over time, was used as the performance metric of interest.

ally, though not presented, we varied the number of frames in the image for the models to learn from and found that the LSTMs' performances improve drastically with an increased number of frames. For computational reasons, however, the studies shown are of a fixed number of frames, though we highlight improved performance with increased temporal data.

While dispersion predictions are in general accurately depicted by the models, their accuracy can be improved, even for a fixed number of frames. The biggest challenge when training convolutional LSTMs over additional channels (dimensions) is ensuring the architecture treats both channels equally in its final predictions. When varying the dispersion in both x and y directions prediction over multiple channels

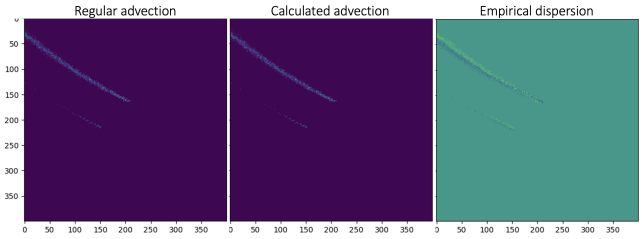

Figure 4: Deep learning training can consider regular emulated tracks (left), calculated advection tracks from the previous frame (middle), or wind-differenced tracks (difference between left and middle) showing empirical dispersion (right).

results in a higher MSE overall. To combat that in future studies, low dimensional representations of the dispersion field could be constructed so as to reduce the architectural loss between multiple channels of interest.

In these studies, the loss function used when training the convolutional LSTMs has a penalization parameter to penalize dispersion predictions outside of track areas by a factor of 100. However, when considering just the track itself in the image, this could be the reason why predictions are not as accurate as they perhaps could be. Looking at lower dimensional representations of tracks themselves, again, could combat this issue.

While considering lower dimensional representations of tracks themselves could combat this issue, we can additionally study physics-informed training data that is *wind-differenced*, i.e. by differencing image concentrations with the wind velocity, via Equation 3 and utilizing the empirical dispersion estimate

$$\Delta^{-1/2}(\mathbf{x}_{(n+1)\Delta} - \hat{\mathbf{x}}_{(n+1)\Delta}),$$

for training in both x and y channels.

Figure 5 shows the different training images that could be used in addition to the regular emulation previously used, for deep learning aerosol dispersion. This is a first step to understanding physics-informed machine learning algorithms, that typically further constrain the underlying loss function to deal with physical dynamics, and their potential to learn aerosol dispersion rates from imagery.

We now consider wind-differenced tracks showing empirical dispersion in the training of the ConvLSTM with the most accurate architecture (three layer bidirectional ConvLSTM with dropout and batch normalization) as shown in Table 1. Using this architecture, we create additional tensor inputs to the network which update the videos to inform the model of the physically observed effects of the dispersion field around the emulated tracks. We find a substantial improvement in predictive power of the dispersion rates using the wind-differenced tracks, as opposed to regular or pure-advection driven emulations, as is shown in Figure 5.

Real Data Study

In order to analyze the effect of the ML model in understanding real ship tracks, we use the most accurate model ana-

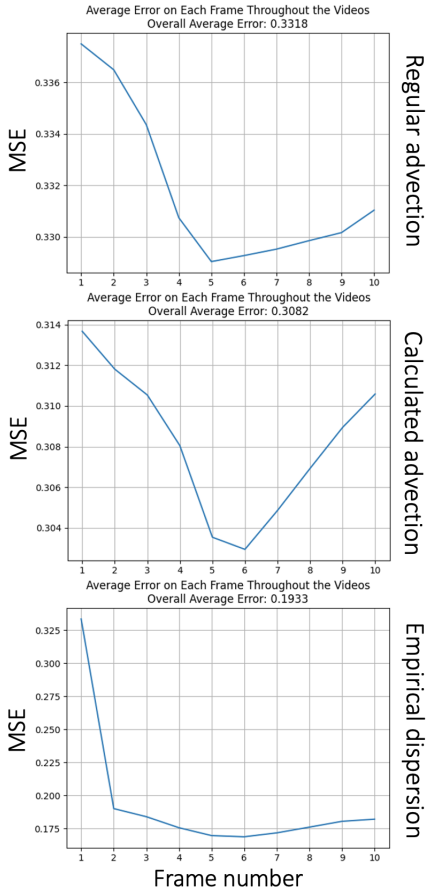

Figure 5: Deep learning prediction errors over a video of 10 frames using a three layer bidirectional ConvLSTM, trained using regular emulated tracks (top), calculated advection tracks from the previous frame (middle), or wind-differenced tracks (difference between left and middle) showing empirical dispersion (bottom).

lyzed previously (bidirectional LSTM with 3 hidden CNN layers) to study the dispersion of hand labeled ship tracks taken off the coast of California from NOAA's GOES-R imager (see Figure 1). To highlight the potential impact of the ML model, we study three different case-studies of ship tracks at 15 minute increments. The first, is a hand-labeled track that does not appear to disperse heavily during the time frame we analyze. The second, is another hand-labeled track that shows some dispersion along the track. Both tracks are highlighted in Figure 1. After using the same model but trained with multiple ship tracks in the same region of interest, the last shows several ship tracks' dispersivities during the same time period. In all studies, we plot the point-wise absolute dispersion values given by $\sqrt{D_x(\mathbf{x},t)^2 + D_y(\mathbf{x},t)^2}$, from outputted longitude and latitude dispersion predictions $D_x(\mathbf{x},t)$ and $D_y(\mathbf{x},t)$, at positions $\mathbf{x}$ and time t, respectively.

In the first track studied, shown in Figure 6, the track head, shown in the north west corner of the track, is seen to coincide with the movement of its corresponding ship. In doing so, dispersion values are predicted higher than along the track, describing stronger aerosol interaction with the track in that region, identified by the track's overlap with three other ship tracks as seen in Figure 1. Along the track, the minimal spreading is echoed by constant dispersion predictions. Finally, higher dispersion is predicted at the end of the track, where its linearity is lost over time. A frequent feature of ship tracks, this may be attributed to higher turbulence in the area. In this region, the model predicts a higher dispersion, reflecting stronger turbulence-influenced aerosol-cloud interactions with the surrounding clouds.

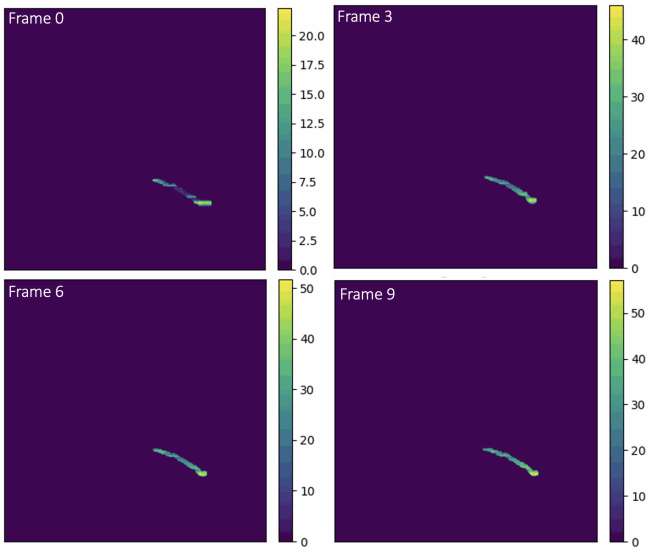

Figure 6: Predicting dispersion values from a hand labeled ship track observed from NOAA's GOES-R imager on July 24 2019, observed at 15 minute intervals between 6 and 7pm.

The second track, shown in Figure 7, highlights a varying dispersion field as predicted by the model. In comparison to the first track studied, the track's shape is largely non-linear, with higher predicted dispersion values reflecting varying aerosol-cloud interactions in creating non-uniform regions of the track. In particular, the model captures higher dispersion at the end of the track, which is seen to disperse faster, both across and along the track over time. In addition, the predicted dispersion values are consistent between both tracks, which are observed as similarly changing track widths in the same satellite image.

Last, we test the model's ability in predicting multiple tracks within the same image, as is common when studying ship tracks (see Figure 1). Multiple ship tracks observed on July 24 2019 were observed in the same satellite image off the coast of California. After being labeled (Warburton, Shuler, and Patel 2023+), all tracks' dispersivities were analyzed using the model, and are shown in Figure 8. Here, it is seen that the model is able to pick up track outlines within the first 15 minutes, and highlight tracks that both persist and disperse, frequently joining other tracks, via higher dispersion estimates. Tracks that are not seen to persistent between

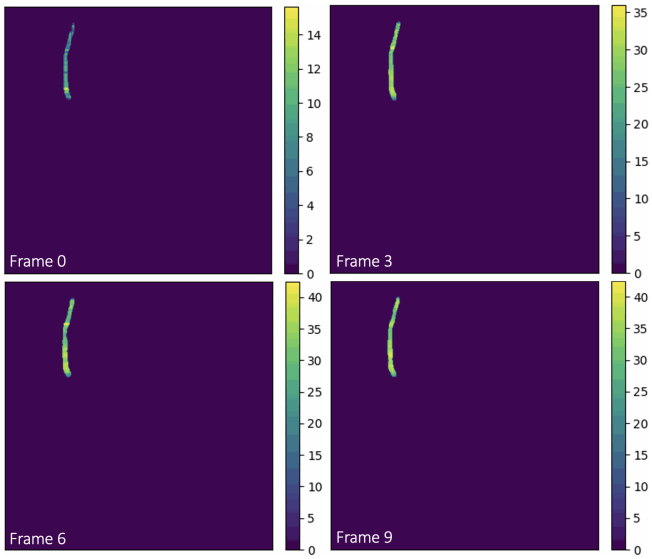

Figure 7: Predicting dispersion values from a hand labeled ship track observed from NOAA's GOESR imager on July 24 2019, observed at 15 minute intervals between 6 and 7pm.

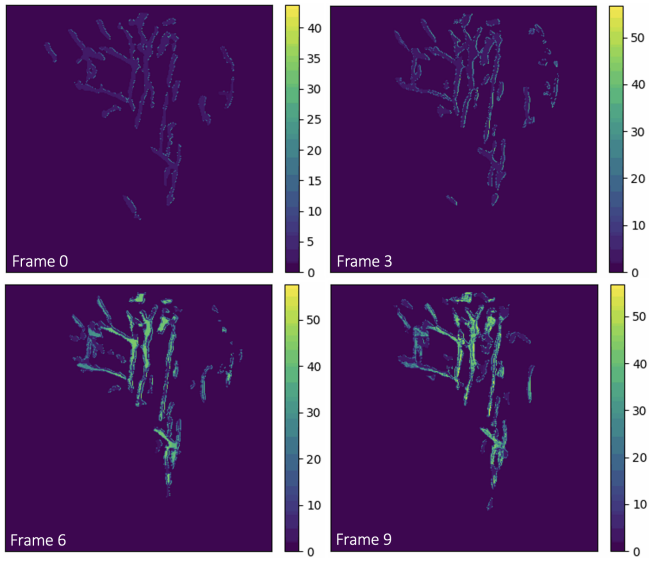

Figure 8: Predicting multiple dispersions from machine learned masks of ship tracks observed from NOAA's GOESR imager on July 24 2019, observed at 15 minute intervals between 6 and 7pm.

images, or that diffuse slowly over the considered time period, have lower estimated values. In particular, this case study highlights the power of the ML algorithm in determining spatial-temporal regions of high ship track activity, and specific tracks that rapidly disperse to form new linear features of atmospheric interest.

The architecture that is used, ingests video inputs to learn *where* ship tracks disperse more heavily. However, the advection fields are not yet leveraged in these studied models, and could potentially further the impact of this work in the atmospheric sciences. As is seen from the simulated studies, under the same scenarios above, we therefore expect the results of this model to show a marked improvement highlighting the more powerful incorporation of the underlying physics in estimating the dispersion fields. This however, is left for future work.

Conclusion

This paper presents the first study that applies a novel deep learning algorithm to observational imaging data of aerosol-cloud interactions, that can be constrained using physical dynamics. By studying ship tracks, examples of aerosol-cloud interactions and an observational study of the effect of Marine Cloud Brightening, we present a cheap emulator, capable of simulating ship tracks, and their interactions with the surrounding clouds with a reduced-order stochastic differential equation. Although aerosol-cloud interactions are highly complex physical phenomena, we use real studies of Large Eddy Simulations of ship tracks to inform the dispersion field of our emulator, whose advection field is captured by wind velocities that are generally measurable. In doing so, thousands of cheap video emulations of ship tracks can be conceived and whose parameters can be varied to train a bidirectional convolutional-LSTM, capable of estimating the dispersion field over time in ship track regions. Further, the algorithm's accuracy is demonstrated to improve by succinct physics-informing of dispersion fields via empirically computed wind-differenced dispersion values of the ship tracks themselves. The dispersion field as developed in this model provides a quantitative means to understand aerosol-cloud interactions across various parts of the track, and therefore surrounding clouds. Predictions that show dense aerosol-cloud interactions could inform the allocation of computational resources in climate models. Further, comparisons of estimated fields between different atmospheric and cloud conditions informs climate intervention researchers the conditions under which using aerosol injection for MCB can effectively be used to combat localized climate change.

In future work, we aim to increase the prediction of our model via studying more LES, and training our emulator directly with the extracted spatial-temporal fields that LES produce. Though expensive, LES can be used to simulate many realistic cloud dynamics, and our framework would benefit from understanding the variety of potential cloud responses in line with the estimated dispersion fields. While it is important to understand how the dispersion field varies across tracks, which can be visualized immediately using our Convolutional LSTM architecture, there may be computational and training benefit in using lower-dimensional representations of the fields, such as via principal components or autoencoders for accelerated learning. All of these aspects are being considered for future work using our constructed framework.

While the analyses of tracks formed under different atmospheric conditions show the power of ML in furthering

scientific understanding of aerosol-cloud interactions, applications of this work are not limited to this exemplar. For example, this methodology could *directly* be used to quantify the spread of, for instance, methane leaks, plane condensation trails and wildfires on the surrounding climate.

Acknowledgements

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly-owned subsidiary of Honeywell International, Inc., for both the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

Blossey, P. N.; Bretherton, C. S.; and Mohrmann, J. 2021. Simulating Observed Cloud Transitions in the Northeast Pacific during CSET. *Monthly Weather Review*, 149(8): 2633 – 2658.

Capaldo, K.; Corbett, J. J.; Kasibhatla, P.; Fischbeck, P.; and Pandis, S. N. 1999. Effects of ship emissions on sulphur cycling and radiative climate forcing over the ocean. *Nature*, 400(6746): 743–746.

Eyring, V.; Isaksen, I. S. A.; Berntsen, T.; Collins, W. J.; Corbett, J. J.; Endresen, O.; Grainger, R. G.; Moldanova, J.; Schlager, H.; and Stevenson, D. S. 2010. Transport impacts on atmosphere and climate: Shipping. *Atmospheric Environment*, 44(37): 4735–4771.

Latham, J.; Bower, K.; Choularton, T.; Coe, H.; Connolly, P.; Cooper, G.; Craft, T.; Foster, J.; Gadian, A.; Galbraith, L.; Iacovides, H.; Johnston, D.; Launder, B.; Leslie, B.; Meyer, J.; Neukermans, A.; Ormond, B.; Parkes, B.; Rasch, P.; Rush, J.; Salter, S.; Stevenson, T.; Wang, H.; Wang, Q.; and Wood, R. 2012. Marine cloud brightening. *Philos Trans A Math Phys Eng Sci*, 370(1974): 4217–62.

McMichael, L.; Schmidt, M.; Wood, R.; Blossey, P.; and Patel, L. 2023+. Exploring ship track spreading rates with a physics-informed Langevin particle parameterization. Forthcoming.

Patel, L.; and Shand, L. 2022. Toward data assimilation of ship-induced aerosol-cloud interactions. *Environmental Data Science*, 1: e31.

Stocker, T.; Qin, D.; Plattner, G.-K.; Tignor, M.; Allen, S.; Boschung, J.; Nauels, A.; Xia, Y.; Bex, V.; and Midgley, P. 2013. IPCC, 2013: Summary for Policymakers. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

Warburton, P.; Shuler, K.; and Patel, L. 2023+. Database of global ship tracks between 2006-2007. Forthcoming.