The Sensitivity of Machine Learning Configuration for Predicting Temperature Profiles in the Planetary Boundary Layer

Kip F. Nielsen[1] and David A. Rahn[1]
[1]University of Kansas, Department of Geography & Atmospheric Science, Lawrence, KS, 66045, United States

Motivation

Traditional numerical weather models struggle with predicting the planetary boundary layer (PBL) in urban areas and during the morning transition. Accurately predicting this part of the atmosphere is crucial because of its downstream impacts, such as accurate air quality forecasts and the representation of convection-based processes. The rapid growth of information technology has increased the capability of machine learning, but several limitations and sensitivities arise when using it to predict the PBL. Most past work predicts components of the PBL, such as PBL height, using generic machine learning configurations. However, this work predicts vertical temperature profiles, in addition to testing a variety of pre-configured and custom loss functions, to learn more about the strengths of weaknesses of using machine learning.

Methods

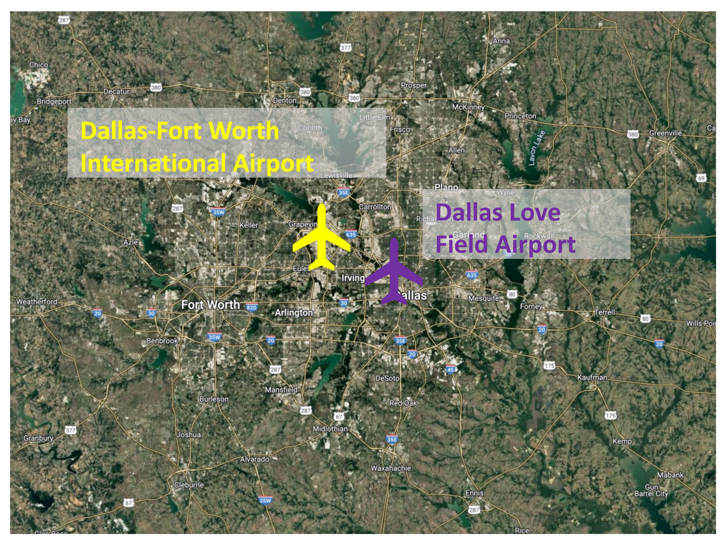
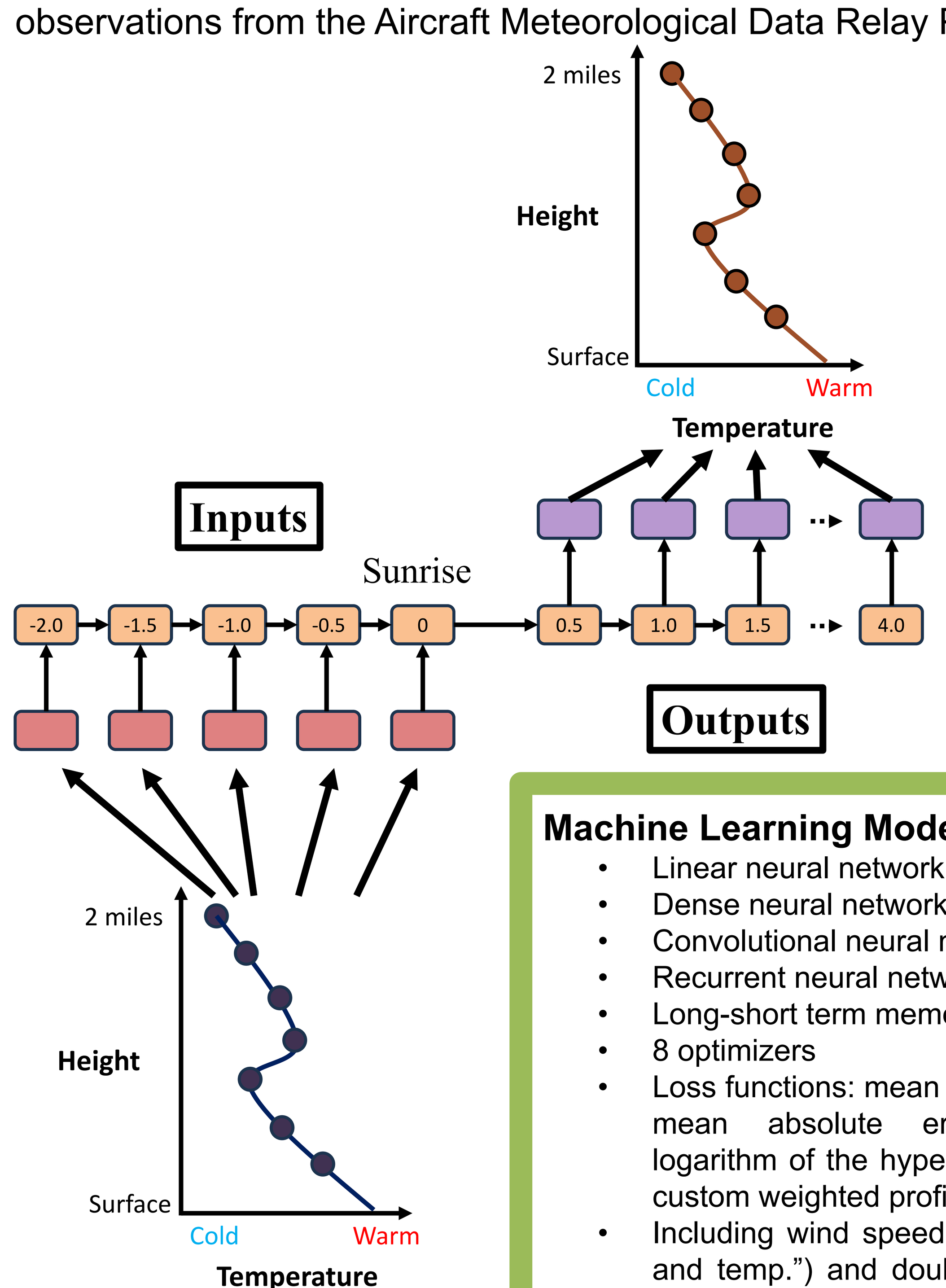


Figure 1. Dallas-Fort Worth, TX, USA is selected because of its large urban area in the central Great Plains with no major topographic features or coastal influences with a high volume of air traffic (2010-2023) for observations from the Aircraft Meteorological Data Relay Program.



Machine Learning Models Tested

- Linear neural network (LNN)
- Dense neural network (DNN)
- Convolutional neural network (CNN)
- Recurrent neural network (RNN)
- Long-short term memory (LSTM)
- 8 optimizers
- Loss functions: mean square error (MSE), mean absolute error, Huber loss, logarithm of the hyperbolic cosine, and 3 custom weighted profiles
- Including wind speed as an input ("Wind and temp.") and doubling the depth and epochs ("Wind and temp. (x2)").

Weather Research and Forecasting (WRF) Model Setup

- 126x126 nested domain at 1km resolution
- Single-layer urban canopy model
- MYJ TKE boundary layer scheme
- Monin-Obukhov surface layer physics

Results

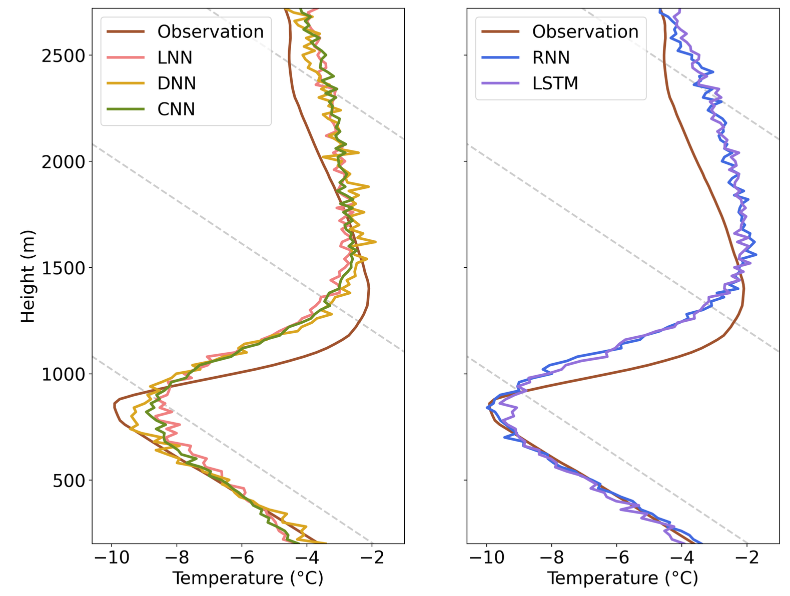


Figure 2. Example predicted profiles of the five model types forecasted 3 hours after sunrise on 20 January 2022. The observation (maroon) shows the actual profile.

Dataset	LNN	DNN	CNN	RNN	LSTM
MSE of training	0.54	0.46	0.52	**0.35**	0.38
MSE of validation	0.60	0.63	**0.60**	0.64	0.62
MSE of testing	**0.74**	0.76	0.75	0.78	0.74

Table 1. MSE (i.e., loss, °C) from each of the models tested. In bold is the lowest MSE for each dataset. The LSTM is chosen for further analysis

Optimizer	MSE	Time to run (min.)[1]
Adadelta	1.96	149
Adagrad	0.99	53
Adam	0.92	60
Adamax	0.97	66
AdamW	0.88	78
Nadam	**0.87**	89
RMSprop	0.88	57
SGD	1.37	**39**

[1]48 CPU cores with 4 gb of memory per core (CPU efficiency of 80-97%).

Table 2. Testing dataset MSE (°C) and time to run as a function of different optimizers. In bold is the lowest MSE for each configuration. This and the rest of the results use a LSTM.

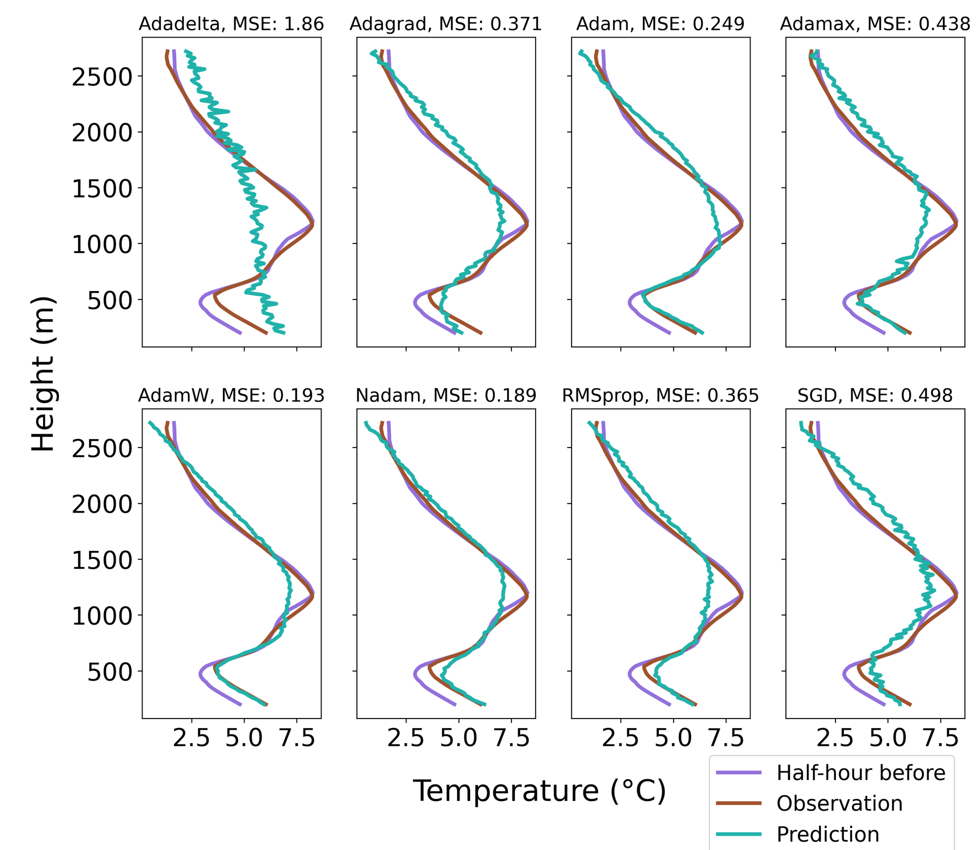


Figure 3. Example predicted profile based on the eight optimizers tested on 30 January 2018, at 3 hours after sunrise. The observation (maroon) and profile from half-hour before (purple) are also shown. AdamW is chosen for further analysis.

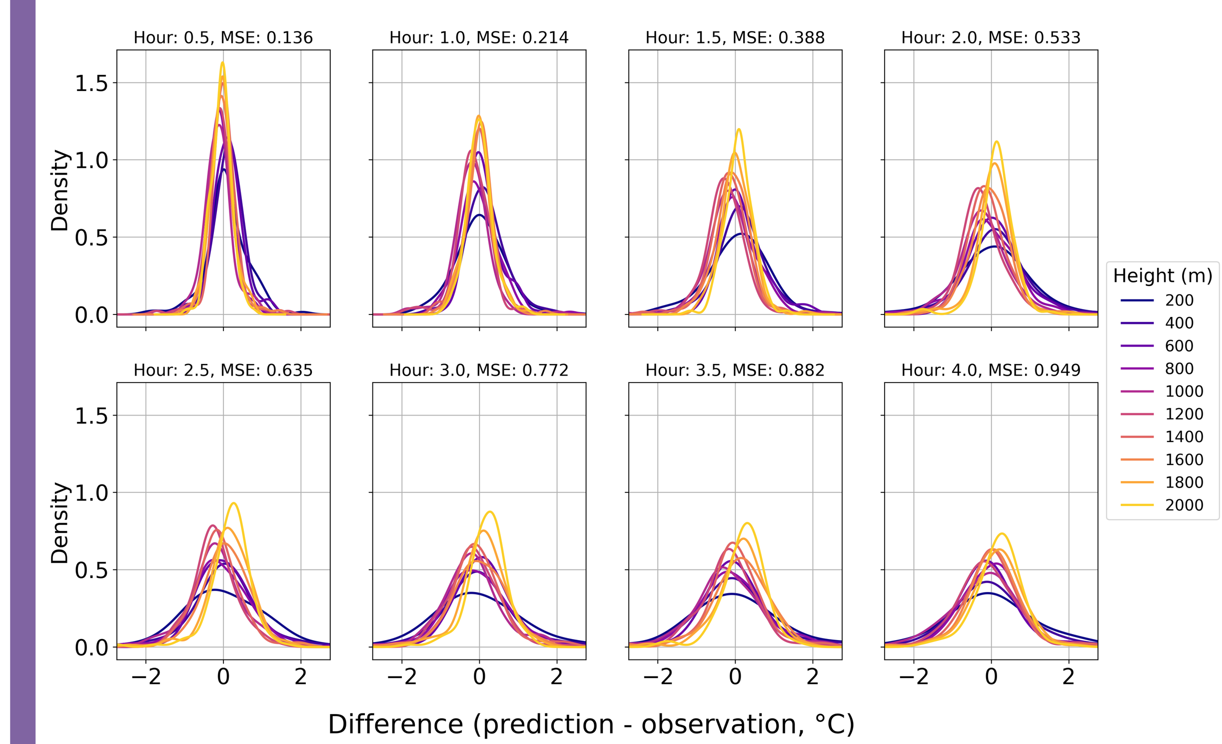


Figure 4. Probability density functions (PDFs) of the difference between the prediction and observation as a function of hour then height. The PDF at each height is the average between 100m below and 100m above.

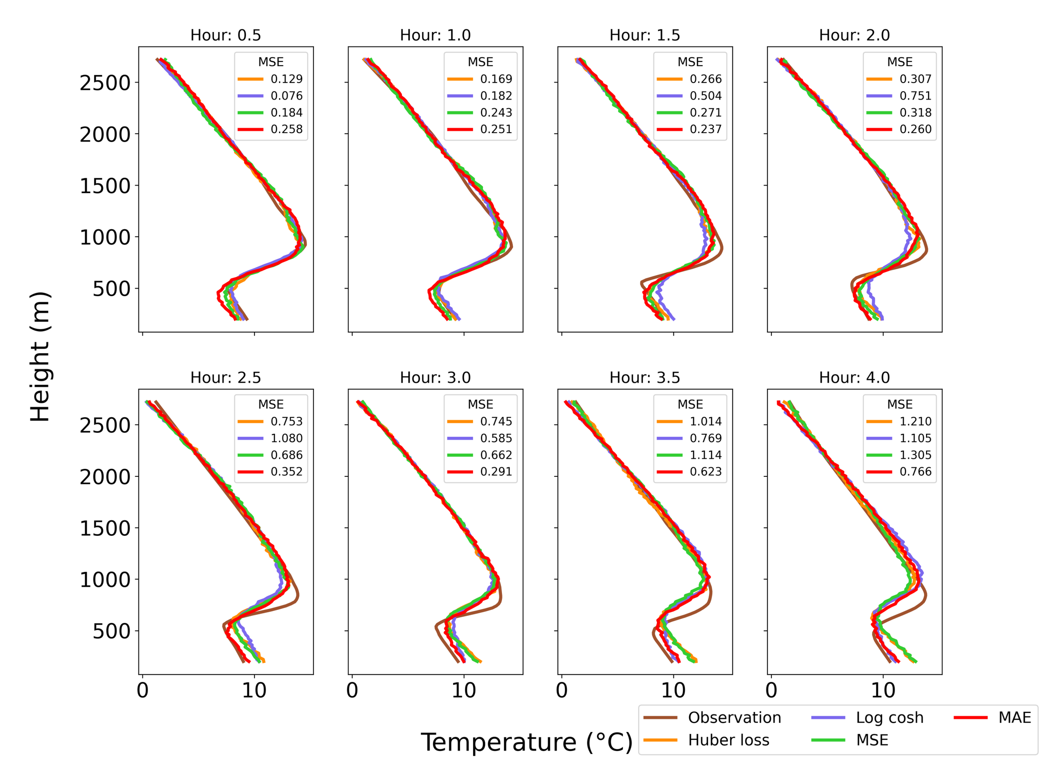
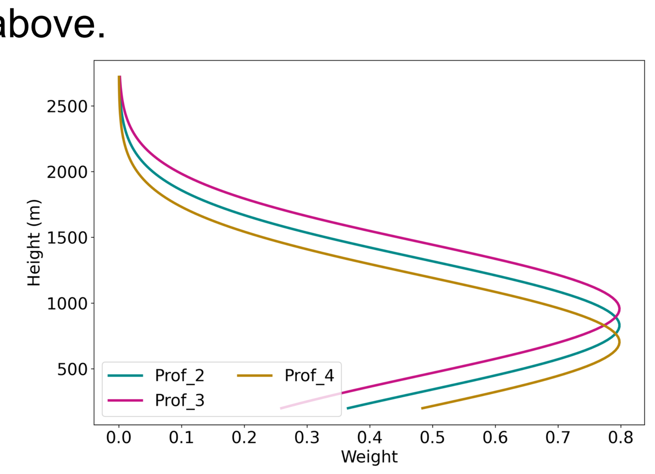


Figure 5. Example predicted profile based on the four standard loss functions on 19 January 2017.



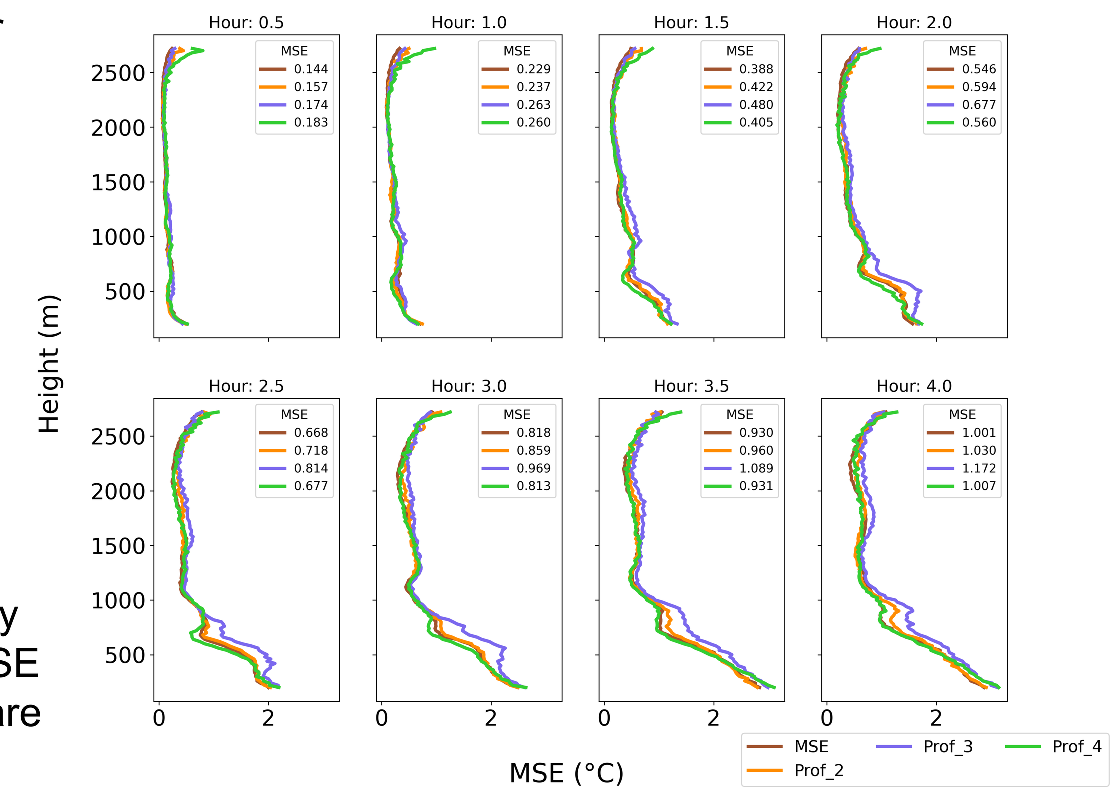


Figure 6. Custom weighted profiles (top) that are multiplied by the profile of MSE for a customized loss function. Average MSE profiles (°C) using these loss functions as a function of hour are also shown (right). Prof_4 is chosen for further analysis.

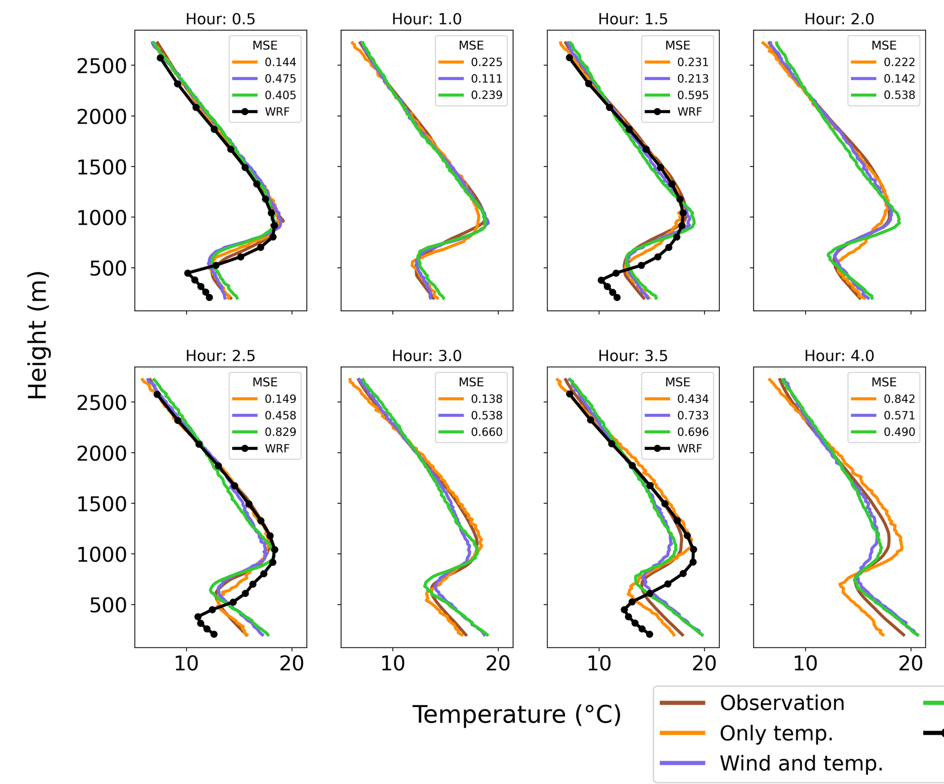


Figure 7. An evolution of predicted profiles using only temperature as an input (orange), wind and temperature as inputs (purple), wind and temperature as inputs but with double the number of epochs and units (green), and a WRF simulation (black and dotted) compared with the actual profiles (maroon) on 23 December 2021.

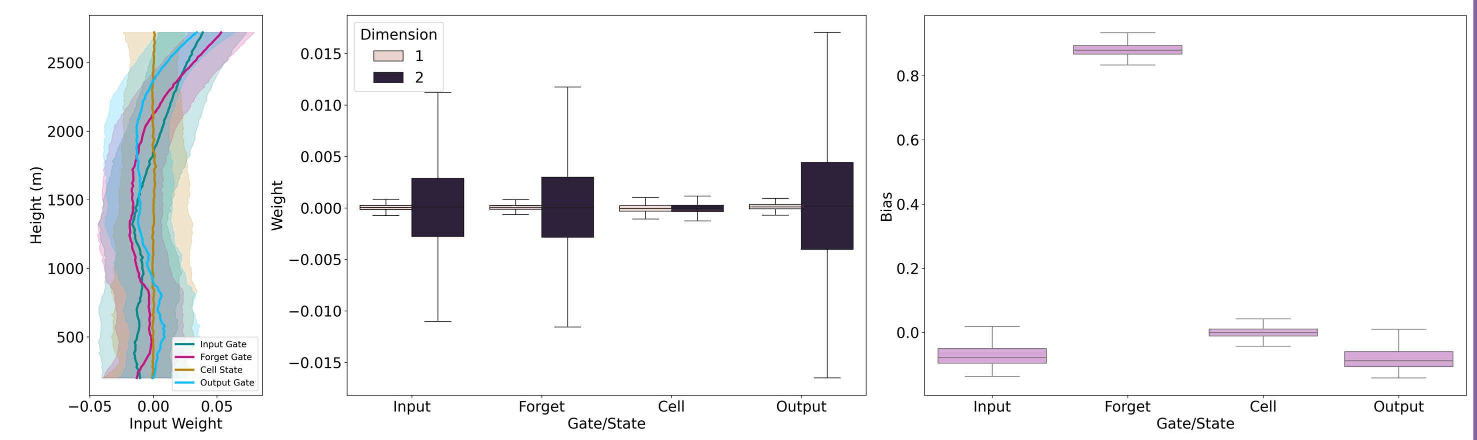


Figure 8. Post-processing input weights in the LSTM as a function of height for the input gate, forget gate, cell state and output gate (left), as well as the weight (center) and bias (right) from the hidden state of the prior cell.

Conclusions

- The LNN had the lowest MSE on the testing dataset, and the RNN and LSTM were the least generalizable (i.e., largest spread between datasets), but the LSTM best captured features of the rapidly evolving PBLs.
- The top three optimizers were only separated by 0.012, but AdamW best captured the evolution of the PBL.
- Mean absolute error was the most accurate loss function with an overall MSE of 0.538°C and a correlation coefficient of 0.958. Prof_4 had the lowest MSE of 0.604°C and a correlation coefficient of 0.953.
- The inclusion of wind speed as an input did not significantly improve the prediction but was generally better sequentially throughout the predicted time series and with inversions at the top of the PBL.
- Compared to WRF, the LSTM generally does better sequentially and below the temperature inversion at the top of the PBL.
- This work highlights the capability in using only temperature to predict the evolution of temperature profiles when the PBL undergoes rapid changes (i.e., during the morning transition), and the value of testing various configurations to back out sensitivity of the model.

Acknowledgements

This project was supported by the Madison and Lila Self Graduate Fellowship and the Department of Geography and Atmospheric Science at the University of Kansas.