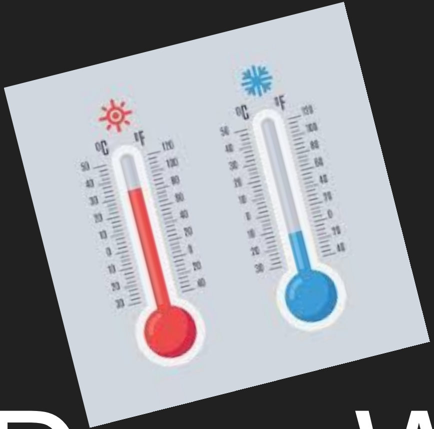


Deep Weather Prediction

Contributors

Miles Young

Blaine Mensa

Matthew Ng

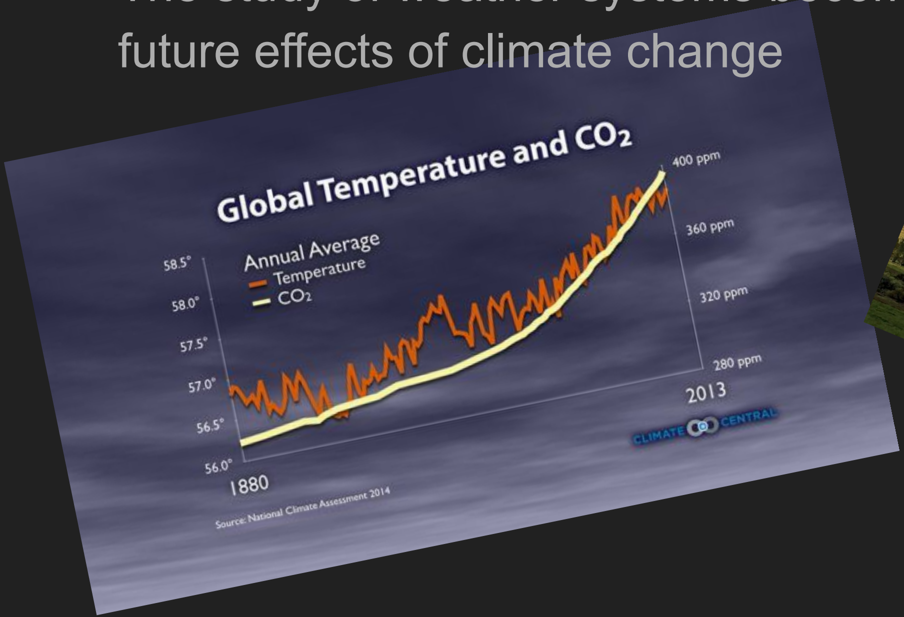
Alex Hoff

Project Introduction

- The focus of the project is weather prediction, specifically temperature
- We conducted a comprehensive literature analysis
- We trained multiple models of varying architecture with varying degrees of success
- We used multiple data sources primarily the NOAA daily summaries (sensor data)

Why is weather prediction important?

- Weather prediction is crucial for various purposes ranging from simple daily forecasting to detailed extreme weather prediction that can save lives
- The study of weather systems becomes increasingly important in studying the future effects of climate change

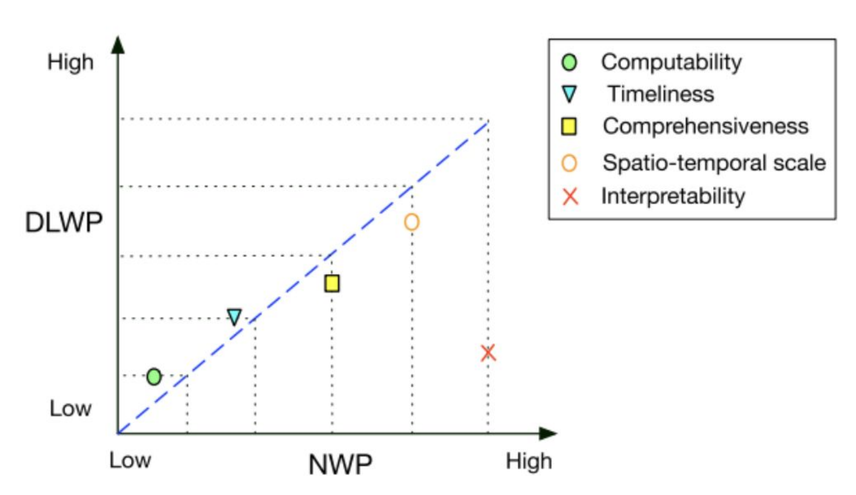


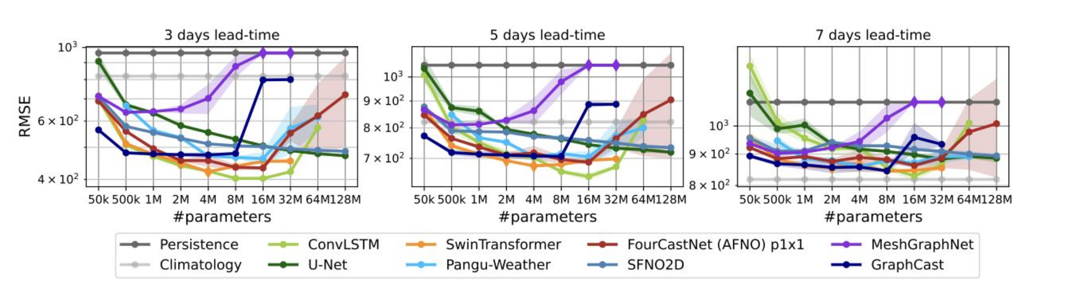
Traditional weather modeling

- Traditional weather models utilize domain expertise in physical study of the atmosphere
- Modern models are general circulation models (GCM) or numerical weather prediction (NWP) which are essentially a set of nonlinear equations, known as the primitive equations
- However, there are still several challenges in NWP. Firstly, due to the chaotic nature of atmosphere, the small differences in initial conditions have a huge impact on model results.
- They identify potential nonlinear correlations between these components through complex physical equations, such as atmospheric dynamics, to generate predictions within a wide spectrum of physical parameters

Shift towards Neural networks in weather modeling

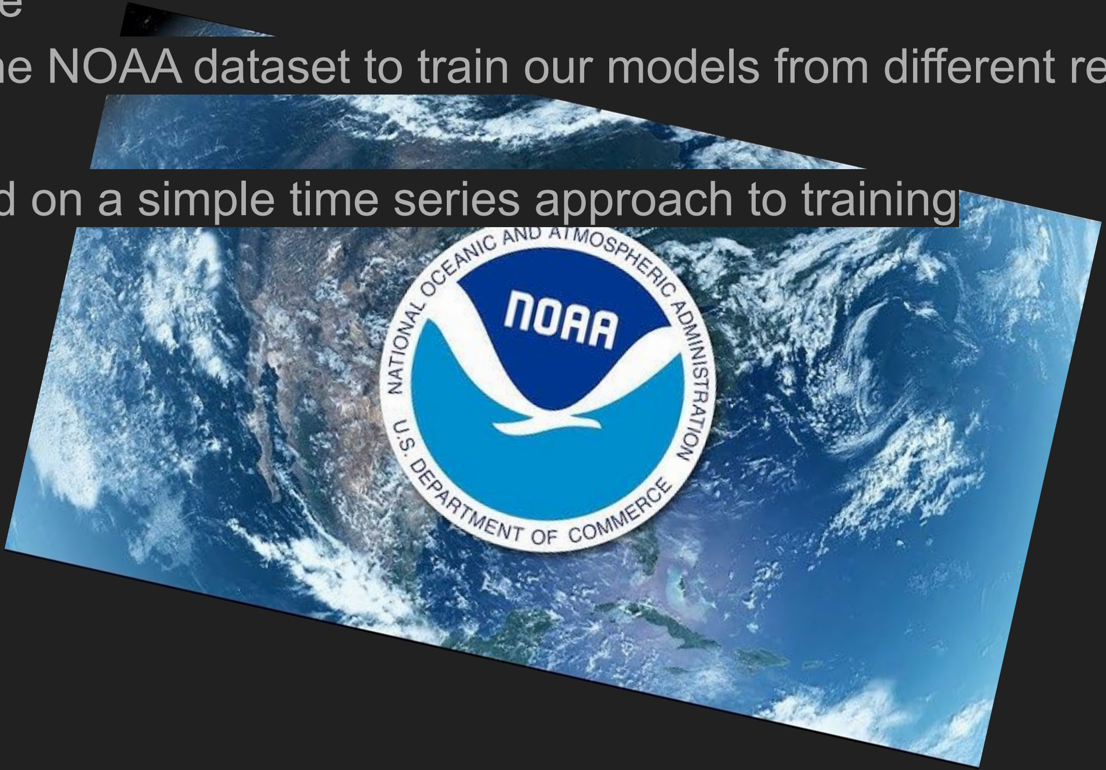
- The numerical approach (NWP) to solve the theory-based nonlinear equations is computationally expensive, which strongly relies on the capability of supercomputers.
- It has been shown that DLWP behaves as well as NWP, or even outperforms NWP in certain conditions
- Another obstacle is the effective integration of observational data from disparate sources, such as weather stations, radars, and satellites [8].





Experiment

- We trained multiple models to predict the average maximum daily temperature
- We used the NOAA dataset to train our models from different regions in the bay area
- We focused on a simple time series approach to training

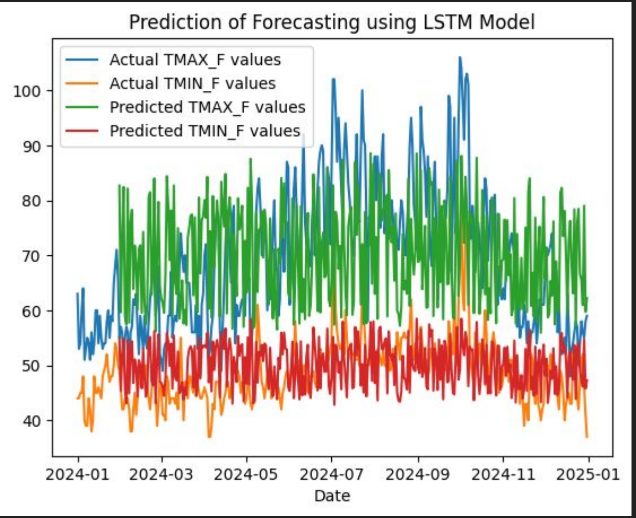


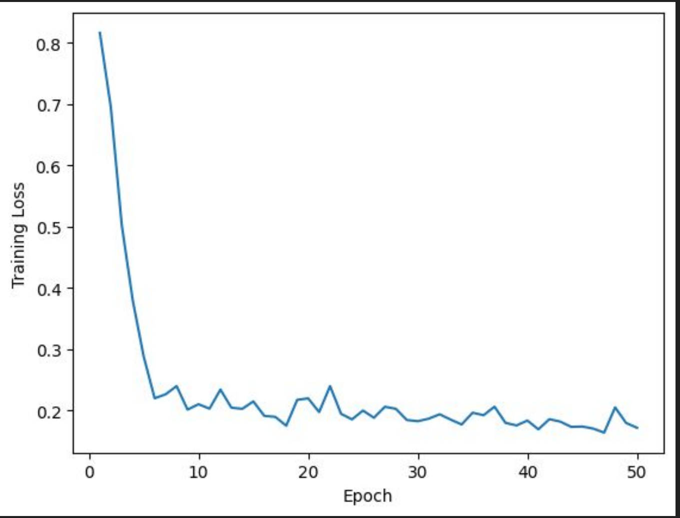
Experiment details

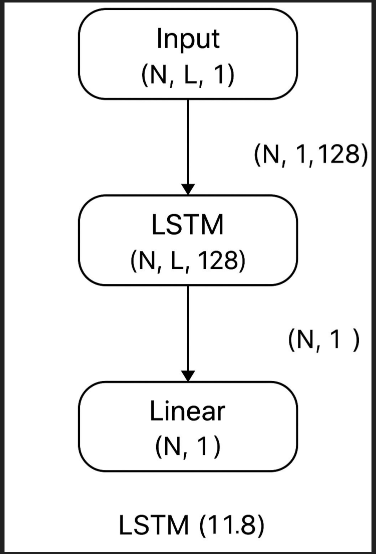
- Data preprocessing involves taking data (in our case about 10 years of daily averages) and splitting into train, test, validation (with no temporal overlap). Then the data was processed into sequences (windows) and standardized.
- Four example models were trained with varying degrees of success

Experiment details

- The first model attempted was a simple LSTM based model
- The model was trained on both maximum and minimum temperature with a resulting MSE of ~130 on the test set (1 year daily data)

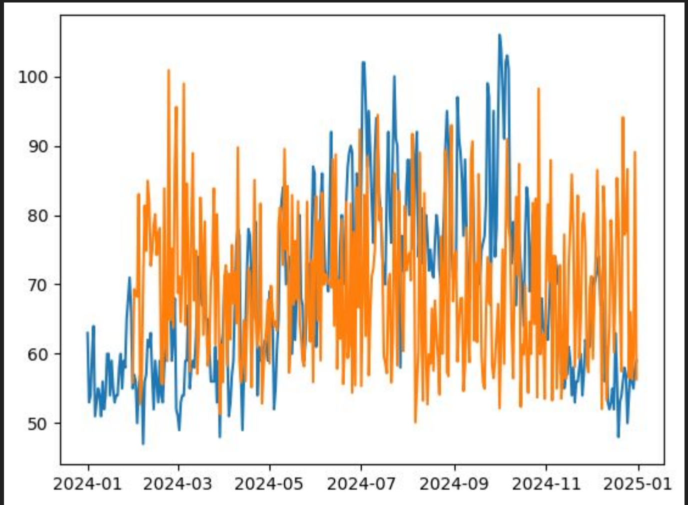


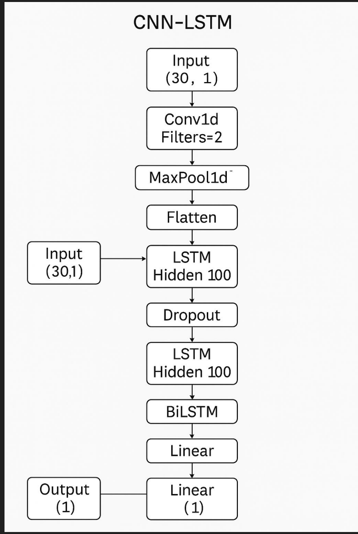




Experiment details

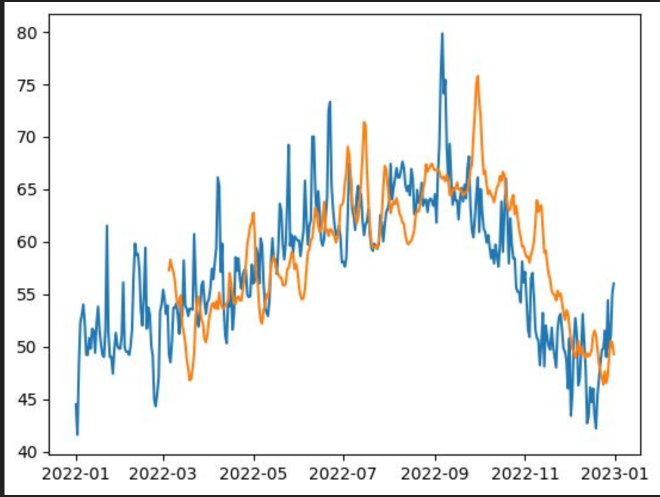
- The next model trained was based on research suggesting using 1D CNN on sequence input followed by RNNs
- While the complexity of the model was significantly greater, but its performance was actually degraded reaching MSE ~183

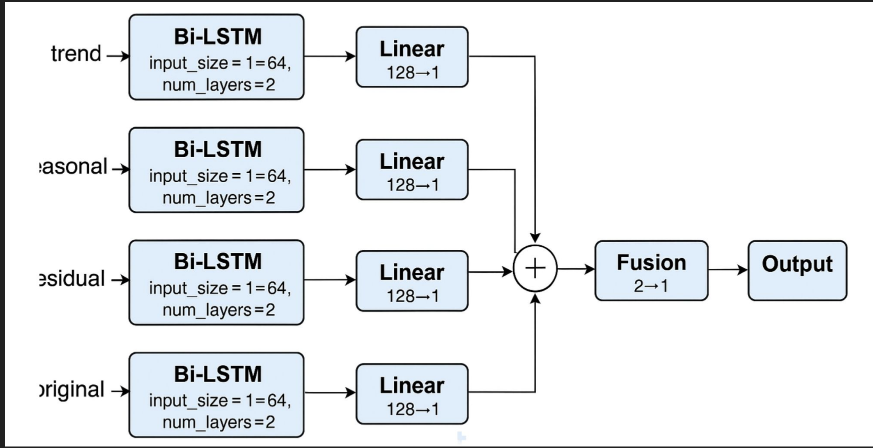




Experiment details

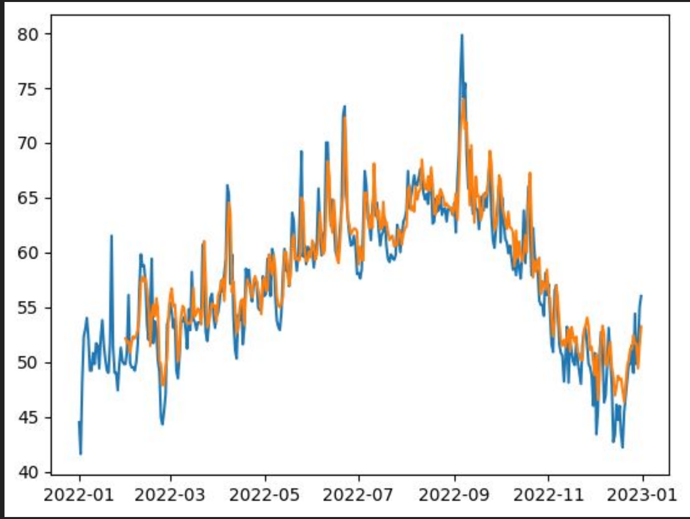
- The final model involves not only a different architecture but different data preprocessing
- Data is decomposed into trend, seasonal and residual components using a statistical method rolling STL
- The model processes the decomposed series along with the original separately and then fuses the result together to generate predictions
- The model achieved our best performance training on 10 years data resulting in MSE ~25





Experiment details

- Finally, we reimplemented the simple LSTM model, but trained it on more predictor features at a different weather station (Oakland international airport NOAA station)
- This resulted in vastly improved accuracy reaching mse of ~6 confirming our intuition that more variables are necessary in weather prediction
- Variables included ['TEMP', 'DEWP', 'SLP', 'STP', 'WDSP', 'MAX', 'MIN', 'PRCP', 'YEAR', 'MONTH', 'DAY']



Discussion

- The performance in the majority of models was severely limited and experimentation with hyperparameters and minor changes to the model architectures had limited effect
- Despite following relevant research in the area, we have come to the conclusion that more emphasis must be put towards developing models that align closer with physical understanding of atmospheric processes
- The potential for models trained on univariate temperature data is very limited

Future work

- Future models should be trained on as many relevant variables as possible (wind, cloud coverage, pressure conditions, humidity, etc)
- Spatial data in combination with temporal data is promising and is commonly used in the field. Attempting to incorporate radar imaging would likely have a positive impact on performance
- Deeper study to physics informed modeling is another promising area

Works cited

- Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K., & Wang, X. (2021). Deep learning-based weather prediction: A survey. Big Data Research, 23, 100178. https://doi.org/10.1016/j.bdr.2020.100178
- Chen, S., Yin, H., Zhang, H., Li, Z., Shen, Z., Yan, S., & Wang, W. (2024). Foundation models for weather and climate data understanding: A comprehensive survey. arXiv preprint arXiv:2401.09566. https://doi.org/10.48550/arXiv.2401.09566
- Karlbauer, M., Maddix, D. C., Ansari, A. F., Han, B., Gupta, G., Wang, Y., Stuart, A., & Mahoney, M. W. (2024). Comparing and contrasting deep learning weather prediction backbones on Navier-Stokes and atmospheric dynamics. arXiv preprint arXiv:2407.14129
- Li, B., & Qian, Y. (2024). Weather prediction using CNN-LSTM for time series analysis: A case study on Delhi temperature data. Applied and Computational Engineering, 92(1), 121–127. https://doi.org/10.54254/2755-2721/92/20241738
- Zhang, K., Huo, X., & Shao, K. (2023). Temperature time series prediction model based on time series decomposition and Bi-LSTM network. Mathematics, 11(9), 2060. https://doi.org/10.3390/math11092060