## Development of a Machine Learning System to Emulate WRF Forecasts for Thunderstorm Nowcasting in Support of the North Dakota Cloud Modification Project

Marwa Majdi, Lynnlee Rosolino, David Delene, Darin Langerud

## Abstract:

Short-term forecasting of convective storms is a critical component of operational weather modification programs. This work presents the development of a machine learning (ML) system designed to emulate high-resolution Weather Research and Forecasting (WRF) model outputs using radiosonde observations from Bismarck, ND. The objective is to provide fast, efficient, and reliable forecasts of thunderstorm-relevant variables to support seeding operations across western North Dakota as part of the North Dakota Cloud Modification Project (NDCMP). The NDCMP is a long-standing operational weather modification program that conducts summertime cloud seeding to enhance precipitation and reduce hail damage over western North Dakota. Accurate and timely forecasts of storm development are essential for planning and executing seeding missions effectively.

The ML framework consists of two separate deep learning models: one trained on 00Z radiosonde observations to forecast hours 1–12, and the other trained on 12Z soundings to forecast hours 13–24. Each model predicts hourly values of convective parameters, including Convective Available Potential Energy (CAPE), Convective Inhibition (CIN), bulk wind shear, radar reflectivity, vertical motion, and divergence, variables that typically require running computationally expensive WRF simulations. By training on storm-only cases from multiple convective seasons, the models are tuned to recognize atmospheric environments conducive to thunderstorm development.

The ML system dramatically reduces forecast latency compared to traditional numerical weather prediction, offering near-instantaneous emulation of up to 24-hour WRF forecasts. It is particularly suited for rapid nowcasting needs during active cloud seeding operations. Results demonstrate strong agreement between ML-predicted fields and actual WRF outputs across training and testing datasets, supporting the model's operational potential as an emulated forecast tool. This approach enables more efficient planning and decision-making for cloud seeding programs, especially in data-sparse or time-constrained environments.