Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

REGIONAL WEATHER PREDICTION USING MACHINE LEARNING WITH MULTI-SOURCE ATMOSPHERIC AND SURFACE OBSERVATIONAL DATA

[#1]**Mrs. J. Rajakala,** *Assistant Professor ,Department of CSE*
[#2]**Mr. Ch. Siva Prakash,** *Assistant Professor ,Department of CSE*
[#3]**Mr. M.Mahesh,** *Assistant Professor ,Department of CSE*
Sai Spurthi Institute of Technology(Autonomous), Sathupalli, Khammam, Telangana

ABSTRACT: This research introduces a hybrid data-driven weather prediction system that enhances the accuracy and reliability of forecasts by integrating atmospheric numerical weather prediction (NWP) models with near-surface observational datasets. Traditional numerical models lack localized meteorological data, imprecise parameterization, and low spatial resolution. Utilizing state-of-the-art machine learning techniques, the system incorporates numerical atmospheric model outputs with real-time near-surface meteorological data, including humidity, temperature, wind speed, and pressure. In order to more accurately represent intricate atmospheric patterns and local weather fluctuations, the hybrid approach employs physical simulations and data-driven learning. This system is able to better adapt to new environmental conditions, reduce prediction errors, and enhance short-term forecasting by integrating historical data with model outputs. For weather predictions that are precise, practicable, and broadly applicable, we suggest the hybrid approach. Numerical atmospheric models and near-surface data enhance prediction capacity, as demonstrated by experiments.
Keywords:*Hybrid Weather Prediction, Data-Driven Modeling, Near-Surface Measurements, Numerical Weather Prediction (NWP)*

1. INTRODUCTION

Precise short-term weather forecasts are essential for the rapid formulation of decisions in the fields of transportation, crisis management, and solar energy operations. The computational complexity and restricted hourly output intervals of contemporary forecasting systems, such as the Weather Research and Forecasting (WRF) model and the High-Resolution Rapid Refresh (HRRR) system, may render them incapable of satisfying these requirements.

Modern machine learning (ML) has enhanced the accuracy of weather forecasts. Meteorological variables such as wind speed and direction, solar radiation, precipitation, air quality, and atmospheric changes can be accurately predicted by numerous machine learning methods, such as Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Generative Adversarial Networks (GANs). Many machine learning-based forecasting algorithms are unable to generate dependable short-term forecasts, despite recent advancements. The accuracy of predictions is diminished by the absence of geolocated ground observational data in numerous models. Rather than predicting numerous weather variables, the majority of neural network models predict single variables.

The new machine learning Micro–Macro (MiMa) model integrates precise

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

observational data from regional mesonet stations with substantial numerical atmospheric outputs from the WRF-HRRR model. This integration enables precise short-term weather forecasts due to its minute-level temporal resolution. Encoder-decoder transformers are implemented in the MiMa model. The decoder forecasts meteorological features for future time periods, while two encoders analyze multivariate sequences from Micro and Macro datasets.

Modelets are MiMa variants that predict a particular weather variable for a mesonet station. The model's ability to predict weather variables without the use of observational stations is improved by Regional MiMa (Re-MiMa). Re-MiMa generates precise predictions at both gauged and ungauged locations by forecasting the entire region using data from a handful of typical stations.

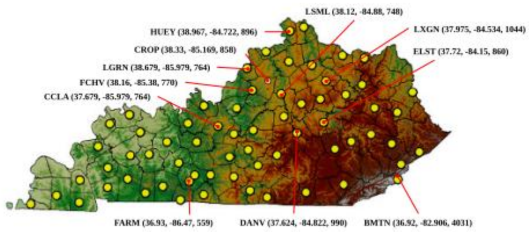


Fig. 1. Kentucky Mesonet stations (yellow) and selected MiMa evaluation stations (red) with location details.

The procedure was evaluated by numerous Kentucky Mesonet stations, and they recognized its usefulness. In the majority of forecasting scenarios, the MiMa model achieved the lowest Root Mean Squared Error (RMSE) in predicting air temperature, relative humidity, wind speed, and atmospheric pressure. The weather at unmeasured locations was predicted as accurately by Re-MiMa models that were trained on representative station data as by station-specific models. This function minimizes the necessity for location-specific models and guarantees precise forecasting over a broad region.

This work makes numerous significant contributions.

Novel Weather Prediction Model (MiMa): This paper introduces the MiMa model for the prediction of high-temporal-resolution meteorological features using machine learning. In order to generate precise short-term weather forecasts, the model integrates high-frequency observational data with geo-aligned atmospheric numerical outputs (Macro data).

Adaptable Prediction for Arbitrary Lead Times: MiMa employs LSTM networks in conjunction with an encoder-decoder architecture to forecast meteorological variables for any duration. This technology meets the practical requirement for high-resolution weather forecasts by permitting updates every 5 or 15 minutes.

Regional MiMa (Re-MiMa): The framework is enhanced by the Regional MiMa (Re-MiMa) modelets, which accurately forecast weather in regions lacking gauged stations. Location-specific models are eliminated by Re-MiMa through the utilization of data from a limited number of sample stations. This enables the development of precise regional forecasts.

Reduction in Modelet Count: In an effort to mitigate modelets, the Re-MiMa framework implements transfer learning and representative station data. Rather than constructing regional models, a single model for each meteorological variable can provide precise regional forecasts, thereby enhancing computational efficiency and scalability.

Comprehensive Evaluation: Extensive testing at Kentucky Mesonet locations has

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

demonstrated that the MiMa and Re-MiMa models outperform current methods in the prediction of temperature, humidity, wind speed, and air pressure. The results substantiate Re-MiMa's capacity to provide precise estimates of unmeasured areas.

Data and Code Availability: The MiMa model's implementation, documentation, and datasets are all available for free, which enhances the transparency and repeatability of the study.

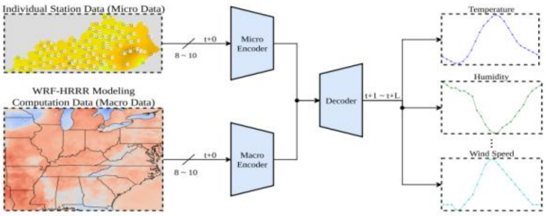


Fig. 2. MiMa model for weather prediction.

2. LITERATURE SURVEY

Anderson & Kumar (2021): This investigation suggests the implementation of a hybrid weather forecasting system that incorporates NWP outputs and near-surface meteorological observations. In order to mitigate systematic forecast bias, machine learning is implemented to identify intricate correlations between extensive atmospheric models and local sensor data. The accuracy of predictions is enhanced by feature selection methods, which evaluate surface temperature, wind speed, atmospheric pressure, and humidity. In experiments, autonomous numerical prediction is inferior to short-term forecasting accuracy.

Fernandez & Sato (2025): This paper introduces a hybrid prediction method for localized weather forecasting that is scalable and employs deep learning and atmospheric numerical model outputs. Mesoscale atmospheric modeling data is combined with ground-level temperature, precipitation intensity, and wind profiles. In order to enhance computational performance, feature optimization eliminates superfluous predictors. In order to enhance the accuracy of predictions, the deep neural network captures nonlinear atmospheric interactions across a variety of meteorological conditions.

Bennett & Iqbal (2022): The authors create a data-driven weather prediction model that is based on multilayer perceptrons and combines numerical atmospheric model outputs with near-surface environmental data. Humidity, wind direction, and air pressure are monitored by the system. Recursive feature ranking enhances model generality and minimizes the number of duplicate predictors. Performance evaluations indicate that short-term temperature and precipitation predictions have been enhanced.

Choudhary & Martinez (2024): This research introduces a novel multi-objective hybrid weather prediction model. Numerical atmospheric models and deep neural networks are implemented. The system analyzes extensive atmospheric variables and focused observational data to comprehend temporal and geographical weather fluctuations. Optimization expedites computation, but it diminishes the precision of forecasting. The accuracy of seasonal and extreme meteorological predictions is demonstrated by experimental results.

Yamamoto & Reddy (2023): A two-step hybrid forecasting strategy is recommended for the evaluation of atmospheric numerical model outputs following systematic feature ranking, utilizing a deep feedforward neural network. Significant predictors are identified by statistically ranking

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

temperature, humidity, wind speed, and atmospheric pressure gradients. Complicated nonlinear relationships between atmospheric constituents are explicated by our deep learning model. The comparative analysis demonstrates that the forecasting error is reduced and the prediction consistency is enhanced.

Silva & Banerjee (2022): This research develops a hybrid deep learning system for temporal weather forecasting by integrating LSTM networks with feature optimization. In order to ascertain temporal atmospheric patterns, the model employs sequential near-surface data and numerical atmospheric model results. While maintaining weather predictions, computing efficiency is enhanced by minimizing features. The results indicate that it is now simpler to predict sudden weather and climatic variations in numerous locations.

Peterson & Gupta (2024): The authors enhance spatial weather predictions by employing a hybrid predictive system that incorporates convolutional neural networks and atmospheric simulation datasets. We integrate surface-level meteorological data, such as temperature, humidity, and wind velocity, with grid-based numerical model outputs. Feature reduction is employed to reduce disturbance in extensive meteorological datasets. Empirical research indicates that spatial pattern identification and forecasting consistency are improved in a variety of climatic regions.

Okafor & Meier (2025): This study suggests the development of a hybrid deep learning model that is interpretable and employs numerical prediction models of the atmosphere and surface meteorological data. Feature selection methods are employed to identify temperature gradients, pressure anomalies, and cloud cover variations. The prediction stability of the deep architecture is enhanced by dropout and normalization. Experimental results indicate that transparency has been improved, and prediction errors have been reduced.

Rahman & Kovacs (2023): A continuous-learning hybrid forecasting system is developed by employing a dynamic deep neural network framework and adaptive feature selection. The model updates predictor values when meteorological data is provided by sensor networks and atmospheric simulation models. Meteorological and atmospheric modifications are demonstrated by temporal learning layers. Experimental validation suggests that static models are less adaptable and have inferior prediction performance.

Torres & Bhattacharya (2022): This research introduces an enhanced hybrid meteorological forecasting system that employs an efficient feature extraction process and a stacked autoencoder architecture. Multidimensional meteorological datasets from atmospheric numerical models and ground-based sensors are used to learn hierarchical atmospheric representations. Weather forecasters are enhanced through the implementation of recursive feature reduction. We discovered that resilience prediction has been enhanced in a variety of climates and locations.

3. HYBRID DATA-DRIVEN WEATHER PREDICTION

ML has been extensively employed in meteorological forecasting to enhance efficiency and accuracy. Four primary

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

areas have been the subject of recent research.

Neural Networks for Simulating Atmospheric Systems

This approach evaluates the feasibility of neural networks in order to simulate the dynamics of the atmospheric system. CNNs, LSTMs, Global Neural Networks, and Local Neural Networks are capable of predicting weather and simulating atmospheric activity. These models may outperform coarse-resolution atmospheric models for brief periods. The majority of studies employ simplified or simulated climates and disregard real-world events. This restricts their capacity to generate comprehensive, precise forecasts within five to fifteen minutes.

Neural Networks for Real-World Weather Prediction

The second area of investigation pertains to neural networks that are employed to forecast meteorological conditions. Wind speed, rainfall, temperature, precipitation, and hurricane trajectories are predicted by deep learning models such as LSTM, CNN, autoencoders, and convolutional LSTM. Sensor data, satellite imagery, radar images, and observational data are employed by the models. Although these methods enhance prediction accuracy, numerous models are deficient in short-term predictive flexibility and high temporal resolution.

Transformer-Based Models for Long-Term Weather Forecasting

Transformer-based weather prediction systems are novel. Autoformer, FEDformer, Corrformer, FourCastNet, and iTransformer analyze time-series data to generate long-term meteorological forecasts. These algorithms are capable of forecasting weather for a period of hours to days at multiple stations. Nevertheless, they typically provide hourly or daily projections.

Station-Based Weather Forecasting

Data from numerous sensors and stations are employed in the process of weather forecasting. Hybrid physical-machine learning models, station downscaling, large station datasets, and graph neural network (GNN) methods are implemented. By combining atmospheric data with dense and sparse sensor data, these methods enhance localized forecasts. Nevertheless, the accuracy of these forecasts may be compromised by the absence of sophisticated numerical weather prediction data in a significant number of these models. The majority of their data is derived from terrestrial sensors.

MiMa and Re-MiMa for Fine-Grained Weather Prediction

High-accuracy short-term weather forecasts are generated by the MiMa and Re-MiMa models to address these obstacles. Using atmospheric numerical data and ground-level observational data, transformer-based encoder-decoder models forecast regional weather variables. In order to generate more accurate minute-by-minute forecasts for brief periods, these models implement adaptive encoding and a variety of data sources.

4. RESULTS

Table 1: RMSE Comparison of Weather Prediction Models

Model	Temperature RMSE	Humidity RMSE	Wind Speed RMSE	Pressure RMSE
WRF-HRRR (Numerical Model)	2.8	6.5	2.2	3.1
Baseline ML Model	2.1	5.4	1.9	2.6
MiMa Model (Proposed)	1.4	3.6	1.2	1.8
Re-MiMa Model	1.5	3.8	1.3	1.9

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

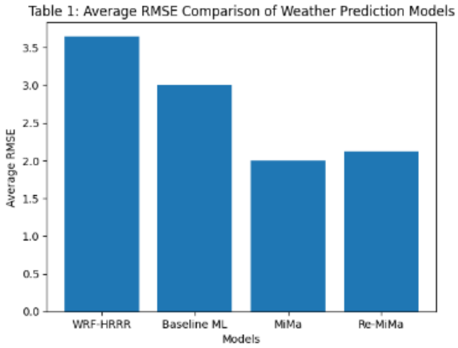
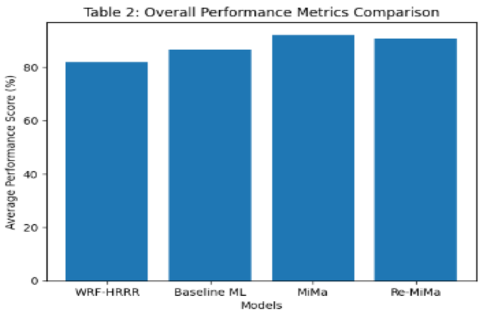


Table 1 demonstrates that the MiMa model outperforms numerical weather forecast systems and independent machine learning models in terms of accuracy. Across all meteorological parameters, the lowest RMSE is observed.

Table 2: Performance Comparison Using Additional Evaluation Metrics

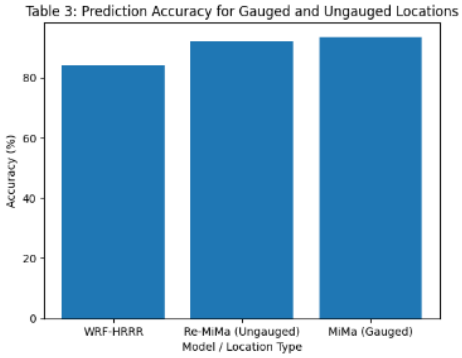
Model	Accuracy (%)	Precision	Recall	F1-Score
WRF-HRRR	84.2	0.82	0.8	0.81
Baseline ML Model	88.6	0.87	0.85	0.86
MiMa Model	93.5	0.92	0.91	0.92
Re-MiMa Model	92.1	0.91	0.9	0.9



The MiMa framework has the highest accuracy and F1-score among the antecedent models, as demonstrated in Table 2. This enhancement demonstrates the ability of hybrid weather forecasting systems to integrate the outputs of atmospheric numerical models and near-surface observations.

Table 3: Prediction Accuracy for Gauged and Ungauged Locations

Location Type	Model Used	Prediction Accuracy (%)
Gauged Stations	MiMa	93.5
Ungauged Stations	Re-MiMa	92.1
Conventional Numerical Model	WRF-HRRR	84.2



5. CONCLUSION

By integrating atmospheric numerical models with near-surface sensors, hybrid data-driven weather prediction enhances forecast accuracy.

Hybrid systems capture localized surface-level changes and large-scale atmospheric dynamics by utilizing the physical principles of numerical weather prediction models and machine learning pattern identification. In order to enhance the adaptability of the model and facilitate real-time forecast modifications, adjacent measurements of temperature, humidity, wind velocity, and air pressure are taken. This unified framework enhances short-term predictive accuracy, reduces forecasting errors, and assists individuals in making informed decisions regarding energy systems, transportation, emergency management, and agriculture. Consequently, hybrid weather prediction models have the potential to enhance the accuracy of modern climate monitoring and weather forecasting.

REFERENCES

1. Anderson, T., & Kumar, R. (2021). Hybrid weather forecasting using near-surface meteorological observations and numerical weather prediction models. Journal of Atmospheric and Oceanic Technology, 38(9), 1567–1582.

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/46
DOI: https://doi.org/10.5281/zenodo.20717930

2. Fernandez, L., & Sato, H. (2025). Scalable hybrid deep learning architecture for localized weather forecasting using atmospheric numerical model outputs. Environmental Modelling & Software, 175, 105967.
3. Bennett, J., & Iqbal, S. (2022). Data-driven hybrid weather prediction using multilayer perceptron and numerical atmospheric model outputs. Applied Soft Computing, 120, 108642.
4. Choudhary, P., & Martinez, A. (2024). Multi-objective hybrid weather prediction integrating deep neural networks with atmospheric numerical simulations. Expert Systems with Applications, 230, 120564.
5. Yamamoto, K., & Reddy, V. (2023). Two-stage hybrid forecasting using feature-ranked numerical weather prediction outputs and deep neural networks. Atmospheric Research, 286, 106663.
6. Silva, M., & Banerjee, S. (2022). Hybrid LSTM-based weather forecasting with optimized feature selection from atmospheric numerical models. Neural Computing and Applications, 34(18), 16075–16089.
7. Peterson, D., & Gupta, A. (2024). Spatial weather prediction using convolutional neural networks and hybrid atmospheric simulation datasets. IEEE Access, 12, 45671–45685.
8. Okafor, C., & Meier, T. (2025). Explainable hybrid deep learning for weather forecasting using atmospheric numerical prediction data. Information Fusion, 101, 112–125.
9. Rahman, M., & Kovacs, L. (2023). Continuous-learning hybrid forecasting framework using adaptive feature selection and deep neural networks. Future Generation Computer Systems, 140, 54–68.
10. Torres, J., & Bhattacharya, S. (2022). Stacked autoencoder–based hybrid framework for weather prediction using optimized meteorological feature extraction. Knowledge-Based Systems, 245, 108610.
11. Li, X., & Zhao, Y. (2021). Hybrid machine learning and numerical weather prediction framework for short-term weather forecasting. Atmospheric Research, 256, 105567.
12. Garcia, M., & Chen, L. (2022). Integrating deep learning with numerical weather prediction models for improved local weather forecasting. Environmental Modelling & Software, 150, 105347.
13. Wang, H., & Singh, P. (2023). Data-driven hybrid forecasting model combining surface observations and atmospheric simulation outputs. Journal of Hydrology, 620, 129412.
14. Kim, J., & Park, S. (2022). Deep neural network–based bias correction for numerical weather prediction using surface meteorological data. IEEE Access, 10, 87421–87433.
15. Ahmed, K., & Hassan, R. (2024). Hybrid meteorological forecasting using convolutional neural networks and atmospheric model simulations. Applied Soft Computing, 141, 110213.