Advances in Thunderstorm and Lightning Prediction: An Extensive Review

Pria Bharti*[1], Dilip Kumar Choubey[2] and Kanchan Bala[3]

{Pria.240301010@iiitbh.ac.in, dkchoubey.cse@iiitbh.ac.in, kanchanbala237@gmail.com}

[1, 2]Department of Computer Science & Engineering, Indian Institute of Information Technology Bhagalpur, Bihar, India.
[3]Department of Computer Science & Engineering, Gaya College of Engineering, Gaya, DSTTE, Patna, Bihar, India.

Abstract. Lighting and thunderstorms is one of the well renowned phenomenon noted in the history of mankind that impacted numerous life loss, financial loss and destruction of infrastructure. The underlying literature depicted changes in conventional methods of prediction to customary new edge technologies by incorporating Artificial Intelligence. The study predicted the transformation of prediction models and methodology from traditional models to highly sophisticated automation techniques around the world. Deep learning / Machine learning increases the accuracy of pinpointing the exact location of an event by using more accurate data rather than broad-scale prediction as offered by the traditional approach. Major challenges include data quality, model transparency & significant computational demands. Future work will prioritize hybrid modelling and improved data assimilation methods to enhance probabilistic forecasts through 2025.

Keywords: Thunderstorm, Lightning, Long Short-Term Memory Networks, Machine Learning, Deep Learning, Numerical Weather Prediction.

1 Introduction

Thunderstorms are severe and temporary in nature and is caused when hot, moist air rises from earth's surface cools and forms cloud precipitation. The basic characteristics of a thunderstorm are strong winds, very heavy rain that might result in flooding, wildfires, and infrastructure destruction. Thunderstorms are further categorised into four groups: single-cell convection, multicell convection, and supercell storms [13]. Lightning is one of the most dengerous phenomenon linked with natural electrical discharges caused by the presence of opposite charges in the atmosphere, which are suddenly neutralized, generating a giant spark [1]. Thunderstorms and lightning are critical in balancing earth's electrical environment, as such phenomena result in refilling earth nitrogen level that is one of the essential element for plant growth. Precise prediction of thunderstorms and lightning is critical to prevent mass destruction of infrastructure, aviation, and human safety. Traditional or historical approaches were heavily based on quantitative measures derived from the study of large-scale weather patterns, using data from particular times and places for analysis. In recent years, it has been observed that the extensive use of data-driven approaches has led to the increased use of multiple, larger observational datasets. In the traditional approach, thunderstorm prediction relied heavily on stable quantitative data such as LI (Lightning Index) & CAPE, along with observations of upper-

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362529

air conditions and atmospheric instability. During the 20th century, forecasting and prediction were dominated by synoptic patterns & empirical rules, especially in countries like the US, due to the invention of NWP (Nowcast Weather Prediction) models, such as WRF (Weather Research and Forecasting), and variants that simulate thunderstorms at high resolution on a grid. However, such predictions range widely and are often pinpointed between micro and synoptic (large) scales, which means different things depending on scientific discipline [2]. In traditional approaches, LPI (Lightning Potential Index) is integrated into NWP using vertical wind, hydrometer, or mixing ratio estimates to estimate striking rates for lightning [2]. Sometimes multiple probabilistic methods, such as European Center for Medium-Range Weather Forecasting (ECMWF) / GEFS (Global Ensemble Forecast System)/ICON (Icosahedral Nonhydrostatic), were combined to improve occurrence forecasts [1]. Further, it was understood that a combination of radar-based derived and Doppler data absorption improves nowcasting, as observed in NOAA's WoF (Warning on Forecast). However, these approaches lack a fine-tuned opinion for individual flashes or explosive system growth.

2 Literature review

2.1 Challenge in Prediction

The specific characteristics like short life span, the butterfly effect and focus only on very small geographical area leads to multiple hurdles in thunderstorm forecasting. Issues such as scattered or irregular data distributed in remote areas can exacerbate or amplify the errors in forecasted thunderstorm data. Similarly, lightning models also face concerns that lead to inaccurate strike prediction and an inappropriate match between accuracy & lead time.

2.2 Advances in Deep Learning and Machine Learning

Deep Learning (DL) and Machine Learning (ML) have transformed thunderstorm and lightning prediction by processing large datasets from radars, satellites, and NWP outputs. Reviews highlight DL's superiority in handling spatiotemporal data over traditional methods. The SALAMA feedforward neural network post-processes NWP data for thunderstorm occurrence prediction, outperforming random forests [1]. Ensemble DL models using radar videos predict severe events with lead times up to 60 minutes [3]. In convective weather, U-Nets and CNNs enable spatially explicit nowcasting of hail and wind, as in ProbSevere systems [4]. Advantages and uses on satellite images by applying U-Net with ResNet-v2 results in short lead times [5]. By combining models such as RNN and CNN results in handling sequential data for forecasts up to 24 hours [1], [6]. Hybrid models, such as ADSNet (Adaptive Difference Siamese Network) and LM-ResNet (Language Modelling Residual Network), integrate attention mechanisms and multiple data sources to achieve superior performance. WoFSCast, an AI-NWP hybrid, accelerates storm-scale predictions 10 times faster. Surveys note DL's efficiency in extreme weather, including thunderstorms, using datasets such as GOES-R and Himawari-8 [5]. Comparative studies show that gradient boosting and CNN-LSTM excel at regional forecasting, achieving accuracies of up to 88%. Several articles that use data mining and soft computing techniques to predict lightning and thunderstorms have been summarized [14]. Further, a table summarizes the tools used, methods, issues, advantages, future work, and existing work.

Table 1. A summary of Techniques, Tool, and Dataset used to forecast Thunderstorm and Lightning.

S.NO	Paper Reference No.	Data Set	Tool Used	Techniques Used	Issues	Advantages	Accuracy
1	[3]	Instability (Wind pressure, Surface Pressure) Moisture, Cloud Cover	Python	SALAMA (Supervised Machine Learning Model) for binary classification of thunderstorms	1. Short lead times (0–2 h). 2. Omits certain predictors and convective inhibition.	1. Delivers reliably calibrated thunderstorm probabilities with superior forecast skill (up to 11 h lead). 2. Scales linearly with spatial extent and effectively manages ensemble forecast uncertainty.	Brier Skill Scores (BSS) over the NWP baseline 11 h lead time.
2	[4]	Radar video Sequence, CAPPI, Spatiotemporal sequences	-------	LRCN: CNN + LSTM	1. Performance is influenced by seasonal heterogeneity and region-specific training. 2. The quality of radar and lightning data constrains generalization.	1. The system functions as a probabilistic and binary warning tool, employing value-weighted scores to address class imbalance. 2. Adapt to Liguria's complex orography, and support decision-making for extreme rainfall events.	(wTSS strategy): 0.68(±0.0.04), wCSI 0.10(±0.02); TSS 0.8247, wTSS 0.742, CSI 0.2164, wCSI 0.1425

3	[7]	Wind, Moisture, Pressure, Composite Radar Reflectivity,	GraphCast, Python, WoFS, WRF-ARW	GraphCast framework, Graph Neural Network (GNN)-based model, spatial Storm Structure Evaluation, Spectral analysis	1. Modest blurring in predictions. 2. No substantial loss of information at smaller scales out to 2 hr despite MSE loss. 3. Potential degradation 4. Trained on data that may not fully resolve all storm-scale dynamics long-term.	1. Matches or exceeds WoFS forecast skill but runs much faster. 2. Environmental regions match 70%–80%.	70%–80%
4	[8]	Composite Reflectivity, Potential Temperature, Water Vapor Mixing Ratio, Grid Resolution	----------	SALAMA 1D, NWP, LINET (LIghtning NETwork)	1. No systematic long-term testing. 2. Limited Dataset Size and Diversity. 3. No overfitting validation Set.	1. Functional prediction from convection-sanction ICON-D2-EPS ensemble technique for Europe, paired with lightning observations.	SALAMA 1D-EPS consistently has a higher BSS than SALAMA 1D up to 11 h.
5	[15]	Wind Speed, Location, Time (Latitude and longitude) from National Center for Environmental Information (NCEI)	----------	SVM (Support Vector Machine), Random Forest (RF), ANN (Artificial Neural Network), GBM (Gradient Boosting Model), Stacked Generalized Linear Model (GLM ensemble)	1. Lack of temporal validation. 2.Computational/Implementation gaps. 3. No causal analysis	1. Probabilistic output provides uncertainty. 2. Improved skill/reliability. 3. Combines multiple models' strengths. 4. Robust integration of diverse data sources.	95%

6	[16]	MRMS radar reflectivity, Visible Reflectance, Shortwave Infrared Reflectance, Longwave Infrared Brightness Temperature	Python	Deep learning (U-Net CNN architecture), watershed algorithm, track linking (trackpy)	1. High rate of true positive events without lightning (62.5%). 2. Inability to distinctly separate lightning-producing from non-lightning storms using radar thresholds.	1. Object-based (not pixel-based) evaluation better matches physical event understanding. 2. Lead times to both radar reflectivity and lightning can be bulk analyzed.	69.4% of tracks are true positives 69.4%. Median lead times
7	[17]	1872 3D-composite reflectivity dataset files from thirteen noticing days during the Korean summer of 2018. manually verified 768 tracks comprise of 5384 convective cells.	--------	Support Vector Regression (SVR), Non-linear,Autoregressive Time-series models, Linear Regression (LR), RF, ANN.	1. Challenges in predicting sudden/stochastic storm motion. 2. Insufficient historical data for some storms. 3. Problem of splits/mergers in storm tracks not fully addressed. 4. Danger of overfitting if inadequate data splitting	1. Flexible for different sampling intervals and nowcasting ranges. 2. Can capture non-linear/abrupt storm movements. 3. Superior RMSE performance over Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN). 4. More precise future location prediction when enough temporal features are available.	The best RMSE (root mean squared error) for next-location prediction is often achieved with a combined ANN/RF method.
8	[18]	Rapid Update Cycle (RUC) model data and Multiyear Reanalysis of Remotely Sensed Storms	Scikit-learn	LR, RF, GB forests, SVM	1. High subjectivity in predicting marginal supercells and linear hybrid models. 2. Computational cost of processing azimuthal shear and sounding predictors. 3. Missing data handling.	1. Reduces cognitive load for manual storm-mode identification. 2. Enables climatological analyses. 3. Potential for operational use	GB forecast: 0.77 ± 0.02 Linear SVM: 0.70 ± 0.02

		(MYRORSS)				with real-time data.	
9	[21]	GOES-16, High-Resolution Rapid Refresh (HRRR) forecasts, NOAA GLM, Earth with total lightning	Python, Scikit-learn	Neural Networks, Postprocessing of HRRR (High-Resolution Rapid Refresh), Probabilistic forecasting	1. Reduced skill overnight. 2. Over forecasting of probabilities. 3. Limited use of larger flash thresholds.	1. Improves upon existing non-ML baselines. 2. Enhances operational SPC (Storm Prediction Center) thunderstorm forecast. 3. Finer-scale predictions.	BSS 0.42 (overall)
10	[25]	NEXRAD radar via Multi-Radar/Multi-Sensor System (MRMS), GOES-16 Advanced Baseline Imager (ABI) and GLM, ECMWF IFS forecasts, ASTER DEM data from April to September 2020.	Light for GB, Python	GB trees	1. Feature importance contradictory to exclusion studies. 2. Limited generalizability. 3. No benefit from DEM (overfitting). 4. Data redundancy, topography difference.	1. Radar is the most important. 2. Satellite alternative for no-radar areas. 3. Lightning is useful for lightning nowcasts.	GB achieves 85–92%

Some researchers [11-14] also compared and analyzed several existing ML models for diabetes and thunderstorm prediction, accordingly.

3 Research Gap

3.1 Diversity in Scales and Regions: Different data sources, like satellite and radar, with different resolutions, might create contrary results distinct from what the actual models think and conclude. These issues prevent models from functioning effectively across diverse regions. Which concerns arise from complex terrains and different places / areas. There is no standard or method for testing and validating models across multiple scales & regions, which makes it challenging to transfer reference data across regions or terrains.

3.2 Understanding model's decisions & reliance on prediction: In a probabilistic based prediction, such as GBM & RF sometimes ignore why a model made a decision, which might

lead to overforecasting / false alarm, or unreliable assessment. In real-time operation, as such models might also lead to computational limitations.

3.3 Scaling operations responsibility: The model mentioned above has not been fully tested over the long term period, although models are very well tested and adopted over small pilot projects. These projects lack ethical checks, such as bias screening. Due to such hurdles, these models cannot be smoothly integrated into large systems like the Warn-on-Forcast. They cannot reliably support policies meant to protect against climate-related risks.

3.4 Data heterogeneity and Predictor exclusion: Current approaches rely on region / season-specific inputs (e.g., radar, NWP, satellite) with redundancies and exclusions (e.g., convective inhibition, DEMs), leading to overfitting, seasonal biases, and poor handling of data quality issues, without robust augmentation for underrepresented global contexts.

3.5 Limited Lead Time and Temporal Integrity: Models like SALAMA and Graph Cast achieve reliable short-term (0–11 h) probabilistic forecasts, and they struggle with extended horizons (>12 h) due to unmodeled stochastic storm dynamics, splits / mergers, and overnight degradation, lacking causal inference and longitudinal testing across diverse ensembles.

4 Discussion and Future Direction

The aim of all such improvements and improvisation in models and methods is to enhance forecaster's faith in prediction. It was noted that enhancement in hybrid ML / NWP by comprising live data feed fused with multimodal data has given room for such trust on forecasters [8], [5]. An improved stereotype data across regions will positively impact unpredicted climate changes across multi-level terrestrial [5]. These days the prediction on thunderstorms and lightning are intricate processes and respective research works in this subject area also depict multiple concerns with improved ways to mitigate the same. Based on reviewed literature it is imperative to understand that these works provide enough guidance on critical areas of improvement along with way on how to enhance the existing model soundness and robustness of contemporary modelling approaches. However, for lighting there is more scope for extensive research and study by using advanced methodologies with data fusion. In the absence of clear signals of initial convective development, a reliable forecast remains uncertain.

Table 2. Proposed Suggestion on Existing Work

S. No.	Reference No.	Existing Work	Future Work
1	[2]	In this paper researcher compare and substantiate the simulated and monitored result of NMM and ARW techniques which is core of WRF for mesoscale system. Which identify NMM technique is well transmitted for squall line with more potential.	Limited data for model evaluation couldn't concluded appropriately. Depth study on NMM and ARW will give precise understanding about the models.
2	[9]	Current weather research hurdles in terms of data integration, mesoscale techniques and framework of convective processes that focuses on advancement in thunderstorm prediction by integration of ML, radar & satellite.	Future work will focus on improving weather & climate prediction by enhancing high-resolution details that simulate storm and atmospheric processes using ML techniques and multiple types of observational data to strengthen predictions.
3	[10]	An explicit mass-based lightning simulation, integrated with a weather prediction model, is assessed in the investigation, which reveals improved flash-rate forecasting across Japan relative to observational estimation techniques, wherein the dual-moment mass microphysics method surpasses the single-moment method.	To improve model performance, future work will combine additional factual data with microphysical parameters to enhance the accuracy of LF (lightning frequency) and spatial predictions.
4	[13]	To improve the prediction accuracy of thunderstorms and lightning, this paper surveys different techniques of soft computing and data mining, including fuzzy logic. It also focuses on mesoscale phenomena and challenges.	To enhance the lead time and accuracy of thunderstorm and lightning prediction, future work will focus on combining hybrid soft computing models with advanced data assimilation techniques and real-time experimental data.
5	[28]	The work focuses on developing a Random Forest ML technique to predict thunderstorm happening in the Upper West, Ghana region using meteorological variables such as Convective Available Potential Energy	Future work suggests technique as a "groundwork" for sophisticated systems, addressing limitations such as poor extreme-event

		(CAPE), Total Column Rain Water (TCRW) and Total Column Water (TCW).	prediction and a lack of spatial / temporal integration.
6	[30]	The work aims to understand how thunderstorm gusts form and to improve prediction using models and AI, especially in the Beijing-Tianjin-Hebei area, where the landscape makes forecasting difficult.	Emphasizes validation amid small samples (~54 days) and complex gust causes, positioning TG-net as operational at the Institute of Urban Meteorology for ongoing refinement.
7	[31]	This work focuses on highly localized, region-based weather change detection with a mobile-adaptable alternative that is more widely applicable than convection-based weather portals.	Future iterations could incorporate lightning-specific sensors (e.g., electromagnetic detectors) and hybrid ARIMA-CNN models for convective event nowcasting, bridging local IoT with global networks like the National Lightning Detection Network (NLDN)
8	[32]	To predict daily weather parameters such as humidity, wind speed, temperature and pressure for Indian cities (Bangalore, Mumbai, Kolkata, Chennai, Delhi), this system proposed a method that uses an ANN (Artificial Neural Network) and provides accuracy against Global Data Assimilation System (GDAS) benchmarks.	By combining ANNs with RNNs (Recurrent Neural Networks), future work will focus on longer horizons (>192 hours) and multi-parameter predictions. Research can also link to climate studies and validate against emerging events, such as post-COVID emission changes.
9	[34]	The paper addressed the issue of imbalanced data for lightning prediction. Earlier dynamics and thermodynamic parameters taken from the KLAPS model also mitigate under sampling. To improve, RF technique for lightning use NWP analysis data (KLAPS).	The proposed method is highly dependent on NWP accuracy. Future researches may use from integrating refined remote sensing data to improve for small period predictions (1-2h lead time). Enhancing the temporal window will also improve the model performance.

10	[36]	Traditional nowcasting is excellent for short-term but not for beyond 3 hours. Paper emphasizing nowcasting limitations and the shift to NWP/ML hybrids for longer horizons in pan-European Air Traffic Flow Management (ATFM).	Enhancing the use of NWP datasets in future research can involve designing a model that captures the full ensemble distribution rather than treating it in isolation, and that also analyzes temporal dependencies in the dataset using time-series modelling, such as RNNs.

5 Conclusion

On the basis of thunderstorm and lightning nature and characteristics above review shows that even great enhancement in radar, stallite & model based approach the exact forcasting remain constrained due to limitation of masoscale representation, data integrity and data scarcity. Even though there is advancement in ML/DL that has improved prediction of thunderstorm and lightning by enhancing risk assessment and nowcasting but the basic gap that exist in the scarcity of data its interpretation and application of data in wide range of regions and terristerials. However, it is also notable that due to interactive methodology coupled with high resolution nonhydrostatic NWP systems with diverse observational database and ML based inferences. Based on reference literature it also observed that polishing and enhancing microphysical schemes , data assimilation stretegies and evolving hybrid statistical dynamic framework would markedly improved both storm scale & lightning prediction capabilities. Ensuring emerging method translate effectively into real time forcasting atmosphere future research should compute long- horizon temporal modelling and operational feasibility, also apply some stretegies for handling data imbalance. Continuing research will further refine tools to mitigate the impact of these hazardous phenomenon.

References

[1] Wang X., Hu K., Wu Y., Zhou W.: A survey of deep learning-based lightning prediction. Atmosphere. 14(11):1698. doi: 10.3390/atmos14111698 (2023)

[2] Litta A.J., Ididcula S.M., Mohanty U.C., Prasad S.K.: Comparison of thunderstorm simulations from WRF-NMM and WRF-ARW models over East Indian region. Advances in Meteorology. 951950. doi: 10.1155/2012/951950 (2012)

[3] Yousefnia K.V., Bölle T., Zöbisch I., Gerz T.: A machine-learning approach to thunderstorm forecasting through post-processing of simulation data. Quarterly Journal of the Royal Meteorological Society. 150(763):3495-3510. doi: 10.1002/qj.4777 (2024)

[4] Guastavino S., Piana M., Tizzi M., Cassola F., Iengo A., Sacchetti D., Solazzo E., Benvenuto F.: Prediction of severe thunderstorm events with ensemble deep learning and radar data. Scientific Reports. 12(1):20049. doi: 10.1038/s41598-022-23306-6 (2022)

[5] McGovern A., Chase R.J., Flora M., Gagne D.J. II, Lagerquist R., Potvin C.K., Snook N., Loken E.: A review of machine learning for convective weather. Monthly Weather Review. 150(12):2895-2928. doi: 10.1175/MWR-D-22-0010.1 (2022)

[6] Verma S., Srivastava K., Tiwari A., Verma S.: Deep learning techniques in extreme weather events: a review. arXiv preprint arXiv:2308.10995 (2023)

[7] Flora M.L., Potvin C.K.: WoFSCast: a machine learning model for predicting thunderstorms at watch-to-warning scales. Geophysical Research Letters. 52(10):e2024GL112383. doi: 10.1029/2024GL112383 (2025)
[8] Yousefnia K.V., Bölle T., Metzl C.: Increasing NWP thunderstorm predictability using ensemble data and machine learning. arXiv preprint arXiv:2502.13316. 2025.
[9] Lin J., Qian T., Bluestein H.B., Ditlevsen P., Lin H., Seiki T., Tochimoto E., Barnes H., Bechtold P., Carr F.H., Freitas S.R., Goodman S.J., Grell G., Han J., Klotzbach P., Roh W., Satoh M., Schubert S., Zhang G., Zhu P.: Current challenges in climate and weather research and future directions. Atmosphere-Ocean. 60(3-4):506-517. doi:10.1080/07055900.2022.2079473 (2022)
[10] Tomioka T., Sato Y., Hayashi S., Yoshida S., Iwashita T:. Advantage of bulk lightning models for predicting lightning frequency over Japan. Progress in Earth and Planetary Science. 10:60. doi: 10.1186/s40645-023-00592-w (2023)
[11] Choubey D.K., Paul S.: Classification techniques for diagnosis of diabetes disease: a review. International Journal of Biomedical Engineering and Technology. 21(1):15-39. (2016)
[12] Choubey D.K., Paul S., Bhattacharjee J.: Soft computing approaches for diabetes disease diagnosis: a survey. International Journal of Applied Engineering Research. 9:11715-11726 (2014)
[13] Bala K., Choubey D.K., Paul S.: Soft computing and data mining techniques for thunderstorms and lightning prediction: a survey. Proceeding in International Conference on Electronics, Communication and Aerospace Technology ICECA. 2017:42-46 (2017)
[14] Bala K., Choubey D.K., Paul S., Lala M.G.N.: Classification techniques for thunderstorms and lightning prediction: a survey (2018)
[15] Tirone E., Pal S., Gallus W.A. Jr., Dutta S., Maitra R., Newman J., Weber E., Jirak I.: A machine learning approach to improve the usability of severe thunderstorm wind reports. Bulletin of the American Meteorological Society. 105(3):E623-E638. doi: 10.1175/BAMS-D-22-0268.1 (2024)
[16] Ortland S.M., Pavolonis M.J.: An object-based evaluation of output from a deep learning model for thunderstorm nowcasting. Artificial Intelligence for the Earth Systems. doi: 10.1175/AIES-D-24-0071.1 (2025)
[17] Lee H., Kim J., Kim E.K., Kim S.: A novel convective storm location prediction model based on machine learning methods. Atmosphere. 12(3):343. doi: 10.3390/atmos12030343.(2021)
[18] Jergensen G.E., McGovern A., Lagerquist R., Smith T.: Classifying convective storms using machine learning. Weather and Forecasting. 35(2):537-559. doi: 10.1175/WAF-D-19-0170.1 (2020)
[19] Arora K., Ray K., Ram S., Mehajan R.: The role of instability indices in forecasting thunderstorm and non-thunderstorm days across six cities in India. Climate. 11(1):14. doi: 10.3390/cli11010014 (2023)
[20] Shield S.A., Houston A.L.: Diagnosing supercell environments: a machine learning approach. Weather and Forecasting. 2022;37(5):771-785. doi: 10.1175/WAF-D-21-0098.1 (2022)
[21] Sobash R.A., Ahijevych D.A.: Evaluating machine learning–based probabilistic lightning forecasts using the HRRR: a comparison to three forecast baselines. Weather and Forecasting. 2025;40(5):759-774. doi: 10.1175/WAF-D-24-0140.1 (2025)
[22] Ukkonen P,, Mäkelä A.: Evaluation of machine learning classifiers for predicting deep convection. Journal of Advances in Modeling Earth Systems. 11(6):1784-1802. doi: 10.1029/2018MS001561 (2019)
[23] Kyei-Manuh S., Osei M.A., Aryee J.N.A., Quansah E., Obuobie E., Amekudzi L.K.: Lightning and thunderstorms, the source of severe weather over Ghana, West Africa. PLOS Climate. 4(9):e0000703. doi: 10.1371/journal.pclm.0000703 (2025)
[24] Kamangir H., Collins W., Tissot P., King S.A.: A deep-learning model to predict thunderstorms within 400 km² South Texas domains. Meteorological Applications. 27:e1905. doi: 10.1002/met.1905 (2020)
[25] Leinonen J., Hamann U., Germann U., Mecikalski J.R.: Nowcasting thunderstorm hazards using machine learning: the impact of data sources on performance. Natural Hazards and Earth System Sciences. 2022;22:577-597. doi: 10.5194/nhess-22-577-2022 (2022)

[26] Ortland S.M., Pavolonis M.J., Cintineo J.L.: The development and initial capabilities of ThunderCast, a deep learning model for thunderstorm nowcasting in the United States. Artificial Intelligence for the Earth Systems. doi: 10.1175/AIES-D-23-0044.1 (2023)
[27] Amjad K., Malik M.H., Ghous H., Hussain A., Ismail M.: Thunderstorms prediction using satellite images. Int J Inf Syst Comput Technol. 2(1):1-24. doi: 10.58325/ijisct.002.01.0044 (2023)
[28] Sekyere A.O.O.: Modeling thunderstorms using machine learning for WA. Preprint. doi: 10.20944/preprints202410.0433.v1 (2024)
[29] Azad M.A.K., Islam A.R.M.T., Rahman M.S., Ayen K.: Development of novel hybrid machine learning models for monthly thunderstorm frequency prediction over Bangladesh. Nat Hazards. 107:1109-1133. doi: 10.1007/s11069-021-04722-9 (2021)
[30] Yunqing L., Chen M., Si J., et al.: TG-Net: a physically interpretable deep learning forecasting model for thunderstorm gusts. J Meteorol Res.39(2):1-22. doi: 10.1007/s13351-025-4080-y (2025)
[31] Manan R., Bharmal S.R., Hitawla F.A.A., Gupta P.: Machine learning for weather prediction and forecasting for a local weather station using IoT. International Research Journal of Engineering and Technology7(2):419-423 (2020)
[32] Ajinaa A., Jaya Christiyan J.C.K.G., Bhat D.N., Saxena K.: Prediction of weather forecasting using artificial neural networks. (2023)
[33] Cavaiola M., Cassola F., Sacchetti D., Ferrari F., Mazzino A.: Hybrid AI-enhanced lightning flash prediction in the medium-range forecast horizon. Nat Commun.15:1188. doi: 10.1038/s41467-024-44697-2 (2024)
[34] Shin K., Kim K., Lee G.W.: Supervised learning-based prediction of lightning probability in the warm season. Remote Sens.16(19):3621. doi: 10.3390/rs16193621 (2024)
[35] Cintineo JL, Pavolonis MJ, Sieglaff JM. ProbSevere LightningCast: a deep-learning model for satellite-based lightning nowcasting. Wea Forecasting. 37(7):1239-1257. doi: 10.1175/WAF-D-22-0019.1 (2022)
[36] Jardines A., Eivazi H., Zea E., García-Heras J., Simarro J., Otero E., Soler M., Vinuesa R.: Thunderstorm prediction during pre-tactical air-traffic-flow management using convolutional neural networks. Expert Syst Appl.241:122466. doi: 10.1016/j.eswa.2023.122466 (2024)
[37] Wang J., Zhang Y., Tan Y., Chen Z., Zheng D., Zhang Y., Fan Y.: Fast and fine location of total lightning from low frequency signals based on deep-learning encoding features. Remote Sens.13(11):2212. doi: 10.3390/rs13112212 (2021)
[38] Flora M.L., Potvin C.K., Skinner P.S., Handler S., McGovern A.: Using machine learning to generate storm-scale probabilistic guidance of severe weather hazards in the Warn-on-Forecast System. Mon Weather Rev. 149(5):1535-1557. doi: 10.1175/MWR-D-20-0194.1 (2021)
[39] Zhou W., Wang W., Wang X.: Research on lightning prediction based on GCN-LSTM model integrating spatiotemporal features. Atmosphere.16(4):447. doi: 10.3390/atmos16040447 (2025)
[40] Chakrabarty H., Bhattacharya S.: Application of K-nearest neighbor technique to predict severe thunderstorms. Int J Comput Appl. 110(10):1-4. doi: 10.5120/19349-7677 (2015)