Deep learning and machine learning models in climate change research: A comprehensive review

Madhusudhan HS[1], Ajay Kumar[2]

[1] Lincoln University College, Malaysia; Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India.

[2] Lincoln University College, Malaysia; School of Computer Science & Engineering, IILM University, Greater Noida, Delhi NCR, India.

[1] pdf.madhusudhan@lincoln.edu.my, [2]pdf.ajaykumar@lincoln.edu.my

---

Abstract: Climate modeling is confronting a transformative shift due to incorporation of deep learning (DL) and machine learning (ML) techniques. These methods, which are based on data, provide novel means to enhance the precision, efficiency, and comprehensibility of climate simulations and forecasts. Although traditional numerical models are effective, they require considerable computational resources and frequently have difficulty resolving fine-scale phenomena and capturing nonlinear dynamics. Conversely, ML and DL methods are outstanding at recognizing patterns and can represent intricate processes with the help of extensive observational and simulated datasets. This review consolidates recent advancements in the use of ML and DL within climate science, examines emerging trends and provides information on the state of AI and ML applications today, offering insights into their effectiveness, difficulties, and prospective benefits for furthering environmental sustainability and climate change research.

Keywords: Deep Learning; Machine Learning; Climate Change; Environment.

Introduction

Grasping and forecasting the Earth's climate system is among the most vital scientific challenges we face today. Extreme atmospheric events have a severe effect on societies [1], leading to hundreds of thousands of deaths annually and resulting in significant collateral effects like migrations, damage to infrastructure, transportation issues, and harm to agriculture or ecosystems. Given that the frequency and severity of extreme events have been rising over the past few decades, likely because of climate change processes. In this context, analysis and detection of these events are extremely important [2,3].

For many decades, traditional climate models have played a fundamental role in the field of climate science. These models are based mainly on physical laws that are expressed as differential equations. However, these models require considerable computation and often involve trade-offs in spatial and temporal resolution, particularly when simulating long-term climate behaviour or fine-scale regional phenomena.

The development of deep learning (DL) and machine learning (ML) in recent years has led to significant developments in climate modelling. ML approaches are well-suited for tasks like parameterisation, data assimilation, and downscaling because they have demonstrated enormous promise in extracting intricate, nonlinear connections from high-dimensional data. Subgrid-scale turbulence modelling, precipitation nowcasting, and short-term weather prediction have all benefited from the effective use of deep learning models.

In this paper, we discuss deep learning and machine learning methods applied to climate related problems. We can categorize atmospheric changes based on their physical traits and effects on human society and ecosystems. Moreover, various DL/ML techniques have been linked to issues related to atmospheric changes. For instance, feature selection/extraction problems in ML are often connected to the detection of atmospheric changes. In this context, ML algorithms can identify the most significant features that trigger such changes. By incorporating specific factors to train the ML models, we can address the attribution of atmospheric changes.

Related work

In this work, we analyse existing research that evaluates the worldwide impacts of climate change on various sectors.

a. Climate change and temperature level/CO2

One of the main issues facing the globe today is the increasing global temperature, which has an influence on every nation's economy, ecology, and healthcare system, with regional effects being more noticeable.

Authors in [4] presented deep learning-based methods like LSTM, RNN and DNN to predict temperature and also heat wave events were analysed in this work. The study was carried out on Indo-Gangetic Plain and the experiment results showed that LSTM outperforms RNN and DNN in heat wave prediction and it performs better in all GCM model datasets in terms of downscaling when compared to observed maximum temperature data.

In [9] authors addressed improve of climate model and projections using deep learning. A novel and promising method that creates more reliable and stable climate simulations and opens the door to more physically-based causal deep learning techniques is the coupling of causal discovery with deep learning. For both global 4 K sea surface temperature increase and cooling, the authors reveal that the causal drivers of subgrid-scale processes, as determined by our causal discovery technique in SPCAM, are constant across climates.

Forecasts for more frequent low-wind and high-temperature episodes in the Eastern United States are among the many GCMs whose predicted changes in wind, solar, and temperature variables are downscaled using machine learning [10]. The technique generates data with high-resolution geographical and temporal characteristics akin to historical datasets and operate 40 times quicker than conventional downscaling techniques.

Authors in [11] proposed a CNN architecture to forecast daily precipitation and mean, minimum, and maximum temperatures using ERA5 data. This work aims in building a multi-model ensemble of downscaled temperature and precipitation estimates across the Iberian Peninsula using CNN. The results could provide a new method for supplying local climate change data for adaption plans.

Ocean conditions, particularly sea level and temperature, are changing due to climate change. These modifications have an impact on marine productivity and monsoon precipitation in the Bay of Bengal, which are vital to the Indian economy [13]. To get future climate estimates, a novel data-driven deep learning model is presented. The estimates for 2024–2100 are then corrected by the trained model, which was trained using historical data from 1950–2014.

Human life is seriously threatened by human-induced global warming, which is mostly caused by an increase in atmospheric CO2. Although the majority of research focuses on estimating CO2 emissions on a yearly basis, which is essential for establishing long-term emission reduction goals, accurate daily CO2 emission prediction is just as important for establishing short-term goals. The performance of 14 models in forecasting daily CO2 emissions data from January 1, 2022, to September 30, 2023, across the top four polluting areas (China, India, the United States, and the EU27&UK) is examined in this paper [19].

b. Climate change and Heat Level/Solar Radiation

To simulate and optimize heat transfer in smart city infrastructure and make it both thermally comfortable and energy efficient, the authors suggested a framework based on deep learning, transfer learning, and multi-objective optimization [7]. The paper's framework includes a hybrid CNN-LSTM-based building thermal dynamics prediction model that was refined using data from five major Indian cities after being applied to a large dataset (12.56 metric tons) of Indian structures with a variety of features.

Authors in [4] addresses evaluation of heat wave events through deep learning models for indo-gangetic plain. In this study, LSTM outperformed the observation (12–28 events) in heat wave prediction throughout an area with a comparable temporal range (12–36 events) and geographical occurrence.

In [12], a novel stochastic weather generator (SWG) based on statistical machine learning (SML) is used to assess power flow. The suggested SML model allows the creation and assessment of simulated hourly meteorological data all year long by combining generative adversarial networks

(GANs), probability theory, and information theory. Seasonal patterns, diurnal fluctuations, and weather uncertainty are just a few of the weather variation aspects that the GAN model represents. In stochastic weather modeling, the suggested deep learning model shows notable advantages over shallow learning approaches.

c. Climate change and groundwater recharge/landslide/soil moisture

The most dependable supply of freshwater for use in homes, businesses, and agriculture is groundwater. However, effective groundwater management has been hampered by human interventions in the water cycle.

Author's goal in [5] is to understand how rainfall will primarily replenish groundwater in the future as the climate changes. Precipitation, soil type, land slope, temperature, potential evapotranspiration, aridity index (ArIn), land use and land cover (LULC), and precipitation are all predictors of groundwater recharge.

Mapping landslide susceptibility is essential for both sustainable land-use planning and disaster management. The purpose of this study was to identify landslide-prone areas in Wayanad, Kerala, India. Elevation, slope, aspect, curvature, stream power index, topographic wetness index, land use and land cover, rainfall, flow accumulation, geology, and geomorphology were among the many huge geospatial datasets utilized in the study. The use of several machine learning and deep learning algorithms to identify the areas vulnerable to landslides comes next [16].

The ability of the top soil layer to retain moisture is known as surface soil moisture (SSM). It is a crucial component of the surface water budget. To lessen the consequences of precipitation shortfalls and identify the most effective strategies for managing natural ecosystems in the face of climate change, soil moisture monitoring is essential. The proposed work [20] used MERRA-2 to gather daily SSW data with a spatial resolution of 0.5° × 0.625° for the Tel River Basin in Odisha, India, between 2001 and 2020. This study investigates the dependability of three deep learning (DL) models to predict SSW time series (SSWTS) one step ahead of time.

d. Climate change and Rainfall/Air pollution

Among Earth systems modeling, rainfall forecasting is of paramount importance. Given its critical role in determining India's agricultural production, it is imperative that the country's monthly rainfall be accurately predicted.

In [6] authors proposed a rainfall modeling technique based on a transformer-based deep learning architecture. This method's ability to parallelize sequential incoming data via an attention technique is its main differentiator. Larger datasets may be processed and trained more quickly because to this feature. The transformer-based architecture's predictive performance was evaluated using monthly rainfall data from India over a 41-year period, from 1980 to 2021.

The environment and public health are seriously threatened by air pollution. Implementing successful mitigation solutions requires an understanding of and ability to anticipate PM2.5 concentrations.

In [14] authors study PM2.5 in Jaipur city using five machine learning techniques and air pollutants and meteorological indicators from 2019–2023. The results demonstrated that every air contaminant was above WHO (2021) standards. With NO2, SO2, and NH3, a rising trend in PM2.5 was seen. According to the sensitivity study, SO2 and O3 were more sensitive to variations in PM2.5 concentrations.

The study in [18] uses deep learning (DL) models to anticipate the hourly AQI in Azamgarh, Uttar Pradesh, India. A total of 8760 data points were collected between July 2022 and June 2023 to quantify hourly particulate matter (PM2.5, PM10), gaseous concentrations (NO2, SO2), and meteorological parameters (temperature, relative humidity, wind direction, wind speed, and UV radiation). With higher AQI readings in the winter than in the summer, the projected annual mean hourly AQI was 123, suggesting moderate pollution. To determine the statistical significance of changes in PM2.5, PM10, SO2, and NO2, we employed a MANOVA.

e. Climate change and thunderstorms

Reliable thunderstorm predictions are necessary because thunderstorms are a serious threat to both society and the economy.

In [21] a thunderstorm forecasting method is implemented using CNN. CNNs to take use of the spatial features included in the meteorological data. Using satellite data, the convection prediction learning task is structured as a binary-classification issue. According to the study, CNN-based designs enhance point-prediction model performance, with the highest results coming from fully-convolutional neural-network architectures.

Authors in [22] proposed feedforward neural network model to forecast thunderstorm. Lightning observations and convection-resolving ensemble projections over central Europe are used to train the model. SALAMA correctly calibrates the inference of thunderstorm occurrence probability using just a set of pixel-wise input characteristics linked to thunderstorm development that are taken from NWP data.

The prediction model was developed using the Random Forest method and tested and trained using historical weather data [23]. With a mean absolute error of 2.34 storms and an index of agreement of 0.926, the model showed strong predictive ability, accounting for about 72.6% of the variation in thunderstorm occurrences. The significance of atmospheric moisture content, namely TCW and TCRW, in forecasting thunderstorm frequency was demonstrated by key studies.

Using a meteorological dataset, the DL technique has been used to categorize thunderstorms and non-thunderstorms [24]. Incidence of thunderstorms and non-thunderstorms has been classified using daily observational hourly data sets from 2016 to 2021. Three DL approaches are employed, and when they are contrasted with traditional ML procedures, it is evident that DL approaches performed better than ML. With an accuracy of 92.01%, the CNN method outperforms all other DL and ML techniques.

Table 1 depicts various machine learning and deep learning methods used for climate modeling under various parameters.

Table 1. Study of various machine learning and deep learning techniques for climate modeling

Ref	Methodology Used	Dataset	Parameters/Metrics					
			Temperature Prediction/ warming/cooling	Heat Wave Prediction / Heat Transfer/ solar radiation	Groundwater Recharge/ Landslide/ Soli Moisture	Forecast Rainfall/ Air pollution/ Flood/ CO2 / Thunderstorm	Energy consumption prediction	Rice Crop Prediction / wheat yield Prediction
4	LSTM, DNN, RNN	India Meteorological Department (IMD)	✓	✓				
5	XGBoost	CGWB website			✓			
6	Transformer-based deep learning architecture	World Bank's Climate Change Knowledge Portal (CCKP) website				✓		
7	CNN-LSTM model	Indian Meteorological Department Station, BEEP and EESL		✓			✓	

8	CatBoost, LightGBM, Orthogonal Matching Pursuit	ERA5, MODIS, NASA's TERRA and AQUA satellites, and a cropland mask (CROPGRIDS)						Rice
9	NN	SPCAM data	✓					
10	Open-source generative ML	NSRDB, WTK, Sup3rCC, HRRR	✓					
11	CNN	ERA5 , Iberia01	✓					
12	deep learning and statistical learning	MERRA2		✓ (solar radiation)				
13	Deep learning (CNRM, UNET)	ORAS5, CNRM-CM6 suite	Sea surface temp and sea level					
14	SVR, LR, ANN, RF, KNN, CNN, GRU	Central Pollution Control Board (CPCB)				Air pollution (PM2.5)		
15	SMLR, ANN, SVR, DNN	punjab.data.gov.in						wheat yield
16	RF	https://earthexplorer.			landslide susceptible zones			

		usgs.gov/, https://code.earthengine.google.com/						
17	RF and Bayesian	Google Earth Engine (GEE), Sentinel-1 data				flood susceptibility mapping		
18	Feed forward NN	NA				AQI		
19	ARIMA, SVM, RF, LSTM. CNN-RNN	https://carbonmonitor.org				CO2 emissions		
20	GRU, LSTM, RNN	NA			Soil moisture monitoring			
21	CNN	European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS)				Thunderstorm		
22	Feedforward NN	ICON-D2-EPS				Thunderstorm		

23	RF	ERA5, ECMWF				Thunderstorm		
24	CNN, RNN, LSTM	IMD Ranchi's data				Thunderstorm		

Conclusions

This work has significantly improved our theoretical understanding of the various ways that climate change affects significant areas including temperature, water resources, heat waves, rainfall, air pollution, thunderstorms, etc. Every one of these sectors is important, and our study has shed light on the intricate relationships that exist between each sector and climate change. It makes significant theoretical advances that advance a thorough comprehension. It is observed that machine learning and deep learning models were used to forecast or predict future instances using the real time datasets.

References

1. Horton RM, Mankin JS, Lesk C, Coffel E, Raymond C (2016) A review of recent advances in research on extreme heat events. Curr Clim Change Rep 2(4):242–259
2. Herring SC, Hoerling MP, Kossin JP, Peterson TC, Stott PA (2015) Explaining extreme events of 2014 from a climate perspective. Bull Am Meteor Soc 96(12):1–172
3. Easterling DR, Kunkel KE, Wehner MF, Sun L (2016) Detection and attribution of climate extremes in the observed record. Weather Clim Extremes 11:17–27
4. Chaturvedi, Manisha, Rajesh Kumar Mall, Saumya Singh, Pawan K. Chaubey, and Ankur Pandey. "Statistical downscaling of maximum temperature under CMIP6 global climate models and evaluation of heat wave events using deep learning methods for Indo-Gangetic Plain." *International Journal of Climatology* 44, no. 3 (2024): 953-972.
5. Banerjee, D., Ganguly, S., & Kushwaha, S. (2024). Forecasting future groundwater recharge from rainfall under different climate change scenarios using comparative analysis of deep learning and ensemble learning techniques. *Water Resources Management*, *38*(11), 4019-4037.
6. Nayak, G. H., Alam, W., Singh, K. N., Avinash, G., Ray, M., & Kumar, R. R. (2024). Modelling monthly rainfall of India through transformer-based deep learning architecture. *Modeling Earth Systems and Environment*, *10*(3), 3119-3136.

7. Shafiq, M., Bhavani, N. P. G., Ramesh, J. V. N., Veeresha, R. K., Talasila, V., & Alfurhood, B. S. (2024). Thermal modeling and Machine learning for optimizing heat transfer in smart city infrastructure balancing energy efficiency and Climate Impact. *Thermal Science and Engineering Progress*, *54*, 102868.
8. De Clercq, D., & Mahdi, A. (2024). Feasibility of machine learning-based rice yield prediction in India at the district level using climate reanalysis and remote sensing data. *Agricultural Systems*, *220*, 104099.
9. Iglesias-Suarez, F., Gentine, P., Solino-Fernandez, B., Beucler, T., Pritchard, M., Runge, J., & Eyring, V. (2024). Causally-informed deep learning to improve climate models and projections. *Journal of Geophysical Research: Atmospheres*, *129*(4), e2023JD039202.
10. Buster, G., Benton, B. N., Glaws, A., & King, R. N. (2024). High-resolution meteorology with climate change impacts from global climate model data using generative machine learning. *Nature Energy*, *9*(7), 894-906.
11. Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V. A., & Bushenkova, A. (2024). High-resolution downscaling of CMIP6 Earth system and global climate models using deep learning for Iberia. *Geoscientific Model Development*, *17*(1), 229-259.
12. Fu, X., Zhang, C., Xu, Y., Zhang, Y., & Sun, H. (2024). Statistical machine learning for power flow analysis considering the influence of weather factors on photovoltaic power generation. *IEEE Transactions on Neural Networks and Learning Systems*.
13. Pasula, A., & Subramani, D. N. (2025). Data Driven Deep Learning for Correcting Global Climate Model Projections of SST and DSL in the Bay of Bengal. *arXiv preprint arXiv:2504.20620*.
14. Singh, Saurabh, and Gourav Suthar. "Machine learning and deep learning approaches for PM2. 5 prediction: a study on urban air quality in Jaipur, India." *Earth Science Informatics* 18, no. 1 (2025): 97.
15. Aravind, K. S., Ananta Vashisth, P. Krishnan, Monika Kundu, Shiv Prasad, M. C. Meena, Achal Lama, Pankaj Das, and Bappa Das. "Development of multistage crop yield estimation model using machine learning and deep learning techniques." *International Journal of Biometeorology* 69, no. 2 (2025): 499-515.
16. Lokesh, P., C. Madhesh, Aneesh Mathew, and Padala Raja Shekar. "Machine learning and deep learning-based landslide susceptibility mapping using geospatial techniques in Wayanad, Kerala state, India." *HydroResearch* 8 (2025): 113-126.
17. Abijith, Devanantham, Subbarayan Saravanan, K. S. S. Parthasarathy, Nagireddy Masthan Reddy, Janardhanam Niraimathi, Ahmed Ali Bindajam, Javed Mallick, Maged Muteb Alharbi, and Hazem Ghassan Abdo. "Assessing the impact of climate and land use change on flood vulnerability: a machine learning approach in coastal region of Tamil Nadu, India." *Geoscience Letters* 12, no. 1 (2025): 1.
18. Ansari, Asif, and Abdur Rahman Quaff. "Data-driven analysis and predictive modelling of hourly Air Quality Index (AQI) using deep learning techniques: a case study of Azamgarh, India." *Theoretical and Applied Climatology* 156, no. 1 (2025): 74.

19. Ajala, Adewole Adetoro, Oluwatosin Lawrence Adeoye, Olawale Moshood Salami, and Ayoola Yusuf Jimoh. "An examination of daily CO2 emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models." *Environmental Science and Pollution Research* (2025): 1-26.
20. Sankalp, Sovan, Bibhuti Bhusan Sahoo, Sushindra Kumar Gupta, Mani Bhushan, Rajib Kumar Majhi, and Santosh DT. "A Deep Learning Approach to Predict Surface Soil Wetness and Its Uncertainty Analysis Over the Tel River Basin, India." *CLEAN–Soil, Air, Water* 53, no. 1 (2025): e70003.
21. Jardines, Aniel, Hamidreza Eivazi, Elias Zea, Javier García-Heras, Juan Simarro, Evelyn Otero, Manuel Soler, and Ricardo Vinuesa. "Thunderstorm prediction during pre-tactical air-traffic-flow management using convolutional neural networks." *Expert systems with applications* 241 (2024): 122466.
22. Vahid Yousefnia, Kianusch, Tobias Bölle, Isabella Zöbisch, and Thomas Gerz. "A machine-learning approach to thunderstorm forecasting through post-processing of simulation data." *Quarterly Journal of the Royal Meteorological Society* 150, no. 763 (2024): 3495-3510.
23. Sekyere, Andrew Osei Owusu. "Modeling Thunderstorms Using Machine Learning for WA." (2024).
24. Bala, Kanchan, Dilip Kumar Choubey, Shruti Gedam, and Md Aknan. "Multi-deep learning models based analysis for classification of thunderstorms on meteorological dataset over Ranchi region." *Engineering Research Express* (2025).