BIG DATA AND AI-BASED CLIMATE PATTERN ANALYSIS FOR SUSTAINABLE INFRASTRUCTURE PLANNING IN RAJASTHAN

Poorvi Kalkani, Dr. Priyanka Sharma, Dr. Dilendra Hiran

Research Scholar, Assistant Professor, Assistant Professor
Computer Science
Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India

Abstract: In the current era, climatic conditions of Rajasthan State are highly unstable and unpredictable. With extreme heat to extreme winters, the last few years have recorded some major changes. The conditions affected farmers, field yields, local fauna and resources. Therefore, by using Big Data and AI, the trendiest and uprising technologies of our generation can come in handy to collect a large number of datasets comprising temperatures, precipitation rates, humidity, etc., of the last few decades and analyze them. This data can be used by farmers, city planning committees, environmentalists, health officials, the public, etc. to maintain quid pro quo with ever-changing climatic conditions and to adapt to them. As the weather conditions are predetermined and forecasted, the government can take appropriate actions and provide safe spaces.

Sustainability is a major concern around the globe for balancing environmental, economic and social well-being. It is highly recommended to make decisions that are more environment prone. Rajasthan with its scarcity of vegetation, dry and infertile lands, vast landfills and current precipitation rates, is advisable to maintain and sustain the environment for future generations.

Moreover, with these findings, appropriate infrastructure and well-planned housing can be developed to withstand extreme or unpredictable conditions. This paper aims to feature the current usage of Big Data, AI or any other technology by government of Rajasthan and private companies to maintain sustainability of the state. Nevertheless, it also highlights challenges with these technologies and how to make them more efficient.

Index Terms - **Sustainable infrastructure, climate conditions, Big Data with AI, Rajasthan climate, Human-AI**

I. INTRODUCTION

Climate change is a major concern in the entire world currently. With speedy transformation to urbanization, the impact is significant and challenging. State like Rajasthan, having arid, semi-arid, and precipitation-prone areas, is highly unstable to predict the everchanging climatic conditions.

Arid is derived from Latin word “aridus” meaning dry or parched. With desert-filled lands and extreme temperatures in recent years, it is highly recommended to predict these conditions beforehand and be ready. Not only this, but the precipitation rates are also higher due to more greenery in the state. The local flora and fauna are affected by humid climate, which once was dry enough. Although higher precipitation means more groundwater and fertile lands, the change is affecting housing and building structures.

The infrastructure of rural areas of Rajasthan is not adaptable to intensity of humidity in monsoon regions. With wide courtyard and stone/mud walls, the traditional housing is more suitable for dry and cold climates. With western-region getting greener and people settling in Thar dessert, the climate has been changed with flash flooding in many parts, for which the traditional infrastructure is not adapted to handle.

With rapid growth in Artificial Intelligence and Machine Learning, complex datasets of Big Data (Honti and Abonyi, 2019) can give real-time solutions to agriculture, health, infrastructure and many other challenges (Bertot et al., 2014). AI and ML are major aspects for recognizing patterns and small changes in the climate with speed and precision. These algorithms can scale their work for bigger dataset and monitor minuscule changes in the environment, which some of the old technologies lack. (Sebestyen V et al. 2021). The predictive models use satellite imagery, IoT Sensors, meteorological data for creating and processing datasets (Reichstein et al., 2019).

The purpose of this paper is to address how technological advancements are helping government and common folk and how much more efficient and effective they can be. It explores current technologies, their methodology, limitations, implications and policies and which of them can be used to make sustainable infrastructure development better in Rajasthan State.

Most of the existing studies focus on national-scale climate prediction models and generalized forecasting frameworks. There is very limited research specifically for Rajasthan's semi-arid ecosystem. With AI-driven climate prediction this paper aims to connect the results directly with sustainable infrastructure planning.

II. METHODOLOGY APPROACH

The following points are the current technologies that are being used for research purposes for temperature/climate patterns in various parts of world including Delhi, Haryana, Rajasthan etc.

2.1 Big Data for Analysis

The data driven from sensors, satellites and weather stations is being used by Hadoop System and Apache Spark to make predictions and climate modelling. The dataset relies on IoT sensors, soil moisture, land surface temperature (LST), satellite datasets (MODIS, CHIRPS) and precipitation, humidity etc. It follows the major 3Vs of big data (Variety, Volume, Velocity) (Laney, 2001) with climatic condition being varied, voluminous and generating speedily.

2.2 Google Earth Engine

By using GEE (a cloud-based platform), the predictions for drought-like conditions in remote to urbanized areas is easily done. GEE uses high-resolution satellite imagery (Gorelick et al., 2017) and climate data to provide comprehensive dataset for analyzing weather patterns. (Nathawat, R., Singh, S.K., Sajan, B. et al.) The platform is being used by various researchers and environmentalists to make sustainable decisions in time. However, the platform is lacking with still developing regions, as they are unable to provide the infrastructure to compute high scale data.

2.3 Machine Learning Models

Supervised learning algorithms are being used widely to speculate patterns among heterogeneous datasets. Support Vector Machines (SVM) are generally used for rainfall prediction, while XGBoost or Gradient Boosting is being used for detecting any slight anomaly like small heatwave, precipitation etc. to provide information beforehand. (Rolnick et al., 2019).

2.4 Deep Learning/Neural Network Models

Convolutional Neural Networks (CNN) are being used to extract temporal features from datasets. Deep Learning models are also being used for improving and predicting fluctuating sea level or ENSO (El Nino–Southern Oscillation) which is˜ far better than traditional methods (Ham et al. 2019). Long Short-Term Memory (LSTM) networks are also being used for temporal predictions. And a CNN and LSTM hybrid model (Li, B., and Qian, Y. 2024) can also predict monsoon intensity and duration.

2.5 Explainable AI (XAI)

A major limitation of deep learning systems is interpreting ability. With black-box AI, humans cannot comprehend how decisions are being made. Therefore, explainable AI techniques such as SHAP (SHapley Additive exPlanations) can improve transparency and policy trust (Samek et al., 2017). This is particularly important when AI is being used as an influencer for infrastructure investments.

III. DATASET COLLECTION AND PREPROCESSING

This study uses historical climate data collected from publicly available meteorological and remote sensing repositories, including the National Oceanic and Atmospheric Administration (NOAA), NASA POWER Climate Data Services, and the Indian Meteorological Department (IMD). The dataset comprises climatic observations from various districts of Rajasthan covering the period from 2000 to 2024.

The collected dataset includes parameters like average temperature, rainfall, relative humidity, wind speed, and seasonal variations. These parameters are highly impactful on sustainable infrastructure, climate change, water resource management and development of urban lifestyle. The table below presents the primary variables used for study:

Table 1 Dataset Description

Feature	Description
Temperature	35°C
Rainfall	300mm – 390mm
Humidity	38%
Wind Speed	9 km/h
District	Jodhpur
Year	2000-2024

IV. PROPOSED AI-BASED CLIMATE PREDICTION FRAMEWORK

For this study, a proposed framework is developed that integrates Big Data technologies and Artificial Intelligence techniques for climate forecasting and sustainable infrastructure planning within Rajasthan. The framework combines heterogeneous data sources including satellite imagery, weather stations, Internet of Things (IoT) sensors, and historical meteorological databases.

Initially, climate data are collected and stored within a distributed data environment capable of handling large-scale temporal and spatial datasets. The collected information undergoes preprocessing, normalization, and feature engineering to improve data quality and eliminate inconsistencies.

Subsequently, ML models are trained to identify hidden climatic patterns and forecast future environmental conditions. Three prediction models are considered in this study: Linear Regression, Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) neural networks.

And lastly, the predicted climate indicators are further analyzed for planning a sustainable infrastructure that will be related to renewable energy, water resource management and for extreme heat and cold.

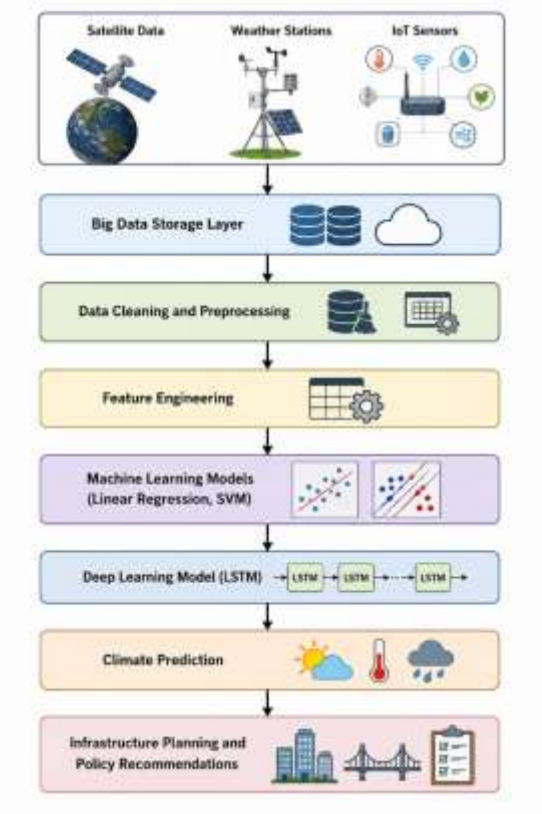

Figure 1 AI Based Climate Prediction Framework

The proposed framework enables proactive decision-making and assists policymakers in transitioning from reactive infrastructure management to predictive climate-resilient planning.

V. APPLICATIONS FOR SUSTAINABLE INFRASTRUCTURE PLANNING IN RAJASTHAN

Sustainable infrastructure planning in Rajasthan is in high demand for climate-responsive strategies currently due to the vulnerability of the state to extreme temperatures, water scarcity, and occasional droughts. Therefore, below are a few policies and technologies that can be implemented to face challenges like water-resource management, urban expansion, and renewable energy systems. These technologies can revolutionize infrastructure and help in real-time, sustainable decision making.

5.1 Urban Heat and Temperature Stress Mapping

Rapidly urbanizing cities like Jaipur, Udaipur and Jodhpur have contributed to an observable rise in urban temperatures, a phenomenon aggravated by heat island effects and reduced green cover (Nathawat et al., 2025; Rangarajan et al., 2025). Empirical remote sensing data demonstrates that average land surface temperature (LST) in Jaipur has increased by nearly 1.8 °C over two decades, related to a 42% decline in vegetative cover (Nathawat et al., 2025).

Table 2 Urban vs Rural Temperature

Area	Temperature (Celsius)
Urban	42°C
Rural	38°C

GIS-based heat mapping can be combined with AI-enabled pattern analysis for the identification of high-temperature areas and mitigate the challenge by urban greening, reflective surfaces, green corridoes and adaptive building codes while enhancing ecosystem.

For example, geospatial frameworks that integrate LST, normalized difference vegetation index (NDVI), slope, and land use data through weighted overlay analysis provide spatial suitability indices for Urban Green Infrastructure (UGI) development within semi-arid cities like Jaipur (Nathawat et al., 2025).

Table 3 ANNUAL AVERAGE TEMPERATURE TREND IN RAJASTHAN (2000–2024)

Year	Avg. Min Temp (Celsius)	Avg. Max Temp (Celsius)	Avg. Mean Temp (Celsius)
2000	16.3°C	32.4°C	24.3°C
2005	17.9°C	34.1°C	26.0°C
2010	18.4°C	34.1°C	26.2°C
2015	17.8°C	34.9°C	26.3°C
2020	17.5°C	28.2°C	22.8°C
2024	18.1°C	24.9°C	21.5°C

5.2 Water Resource and Drought Risk Infrastructure

Rajasthan's frequent drought cycles are significantly affecting water availability, agricultural productivity, and rural infrastructure in many remote places. The recent studies are using cloud-based data for drought indices known as Standardized Drought Composite Index (SDCI). It is generated from satellite monitoring of precipitation and temperature parameters, which are processed via Google Earth Engine (GEE) that reveals drought severity variations across the state (Nathawat et al., 2025). These analyses are taken in consideration for infrastructure planning by highlighting areas where water storage, irrigation systems, and watershed structures must be developed to accommodate extreme climate.

The AI-augmented water-based infrastructure helps in planning canal networks, replenishing groundwater preservers, hydrological modeling etc. The Rajasthan Watershed Development Portal further demonstrates how geospatial and participatory data integration can support community-oriented watershed interventions and rural road infrastructure aligned with water resource stability (Rajasthan Watershed Portal, 2024).

Table 4 ANNUAL RAINFALL VARIABILITY IN RAJASTHAN (2000–2024)

Year	Rainfall (mm)
2000	420
2005	510
2010	380
2015	600
2020	450
2024	530

5.3 Transportation

AI-based climate predictions can also fundamentally improve the planning of transportation and energy systems by anticipating weather beforehand. The state faces rapid degradation for road-maintenance materials, pavement stress and overall route planning in summers. Therefore, predictive modeling of temperature helps engineers beforehand for selecting heat-resilient materials and adjust maintenance schedules for arterial roads. Similarly, the data derived from IoT sensors and predictive analytics can inform officials about flood-risk for vulnerable bridges and transit corridors of the state.

5.4 Energy Infrastructure Resilience

In the energy sector, integrated AI analytics are already being used for demand forecasting, renewable integration, and grid optimization in Rajasthan. Regulatory initiatives supporting digital infrastructure and AI-enabled smart grid analytics aim to improve renewable energy utilization, operational performance, and fault detection across the grid (Government of Rajasthan, 2025). These predictive capabilities are crucial as the state transitions to higher shares of solar and wind power, which are inherently variable and climate dependent.

5.5 Policy and Implementation Frameworks

The effective application of AI and Big Data in infrastructure planning requires a strong framework. Recent policy initiatives in Rajasthan emphasizes on AI-powered data hubs and geospatial analytics to support decision-making across sectors including agriculture, disaster management, and infrastructure risk assessment (Government of Rajasthan, 2026). Strategic investments in climate data systems, drone policies for environmental monitoring, and Centres of Excellence for climate research are expected to catalyze evidence-based planning and cross-sectoral collaboration.

However, challenges remain in ensuring data interoperability, integration of localized climate models with national frameworks, and capacity building among planning agencies. Strengthening policy mechanisms that mandate climate risk assessments and AI-based analytics in infrastructure project approval processes can help institutionalize resilient planning practices.

VI. TECHNOLOGICAL AND STRUCTURAL GAPS

With incorporating multiple AI/ML models, Big Data Frameworks and above technological advancements, Rajasthan is still facing challenges. While the advancements are working well with urbanized and modern mindsight, it is lacking in remote and rural areas. Not only the technology, but people and infrastructure are not equipped to house and understand the technology. The following section highlights what chal lenges are being faced and what policies/solutions are being introduced to solve them.

6.1 Data Quality and Efficiency

The foremost challenge is insufficient and inconsistent data for the dataset and analyzing. There are limited sources for getting older datasets, as digitization was scarce then. The current situation is somewhat concerning, as there are limited workstations for getting geospatial data in rural regions, which affects the quality of data. The state is currently not equipped to house this level of energy and resources.

6.2 Model Limitation

The most common challenge of AI is finding out how decisions are made. The black box decision-making process leads to chaos and inefficiency for us to know how machines are working. Moreover, some limited models use hybridization of Physical and AI-based, but the count is still limited. Old school folk are still lacking in using updated AI-based models and are not technologically advanced, therefore, they are relying on physical models. Reichstein et al. (2019) emphasized the need for integrating process-based climate models with machine learning to improve the reliability of the models and their performance. Lastly, the model overfits the datasets that leads to storage issues and redundancy. Small regions can work fine if the data is limited. Therefore, models need to vary from region to region, making it more efficient.

6.3 Infrastructure Barriers

The high-performance computational workstations and AI modelling architectures like CNN, LSTM, and various frame works require GPU-Enabled clusters, high-performance and large-scale computing. These resources are constrained due to budgetary and institutional limitations. As noted by Reichstein et al. (2019), these models require parallelized data pipelines and large memory architecture, which are absent in rural and developing regions. Furthermore, over-reliance on cloud-based computing increases the operational cost as the infrastructure is not fully equipped to store data locally.

6.4 Lack of Interdisciplinary Coordination

The collaboration between AI model developers, Urban Planners, Civil Engineers, Climate Scientists, and Policy makers is an important factor while planning climate-resilient infrastructure and technology. The lack of communication and the absence of interdisciplinary platforms result in chaos, leading to an imbalance between technology and physical domains. Rasp et al. (2020) highlight the importance of integrating machine learning with physical climate modelling to Physics-Informed Neural Networks (PINNs) integrate physical laws such as Navier–Stokes equations, thermodynamic constraints, and hydrological balance equations into deep learning architectures. Unlike purely data-driven models, PINNs ensure physically consistent predictions.

VII. RESULT AND ANALYSIS

In the above-mentioned datasets, there is a consistent rise in temperature across Rajasthan over the past two decades, accompanied by irregular rainfall patterns. Machine learning models such as Support Vector Machines (SVM) and LSTM demonstrated improved prediction accuracy for short term climate forecasting.

The temperature trend analysis (Table 2) shows an increase of approximately 2.8°C over 24 years, indicating a significant warming. Rainfall variability (Table 3) highlights the unpredictable precipitation patterns over two decades, which directly affects the agriculture and water resource planning in the state.

Moreover, urban heat island analysis (Table 1) indicates that urban regions experience up to 4°C higher temperatures compared to rural surroundings, emphasizing the need for more green trees and less pollution.

These findings have showcased the significant changes in Rajasthan in 24 years which are resulting in abrupt heatwaves and unpredictable rainfall. The need to integrate AI-driven analytics into this type of climate is necessity and can be beneficial for infrastructure planning.

VIII. FUTURE RESEARCH

With such technological advancements and current barriers, Rajasthan's climate conditions can be further integrated with a few technologies that have not yet been used.

8.1 Physics-Informed Neural Networks (PINNs) for Climate Simulation

Physics-Informed Neural Networks (PINNs) integrate physical laws such as Navier–Stokes equations, thermodynamic constraints, and hydrological balance equations into deep learning architectures. Unlike purely data-driven models, PINNs ensures physically consistent predictions.

With the hybrid climate modeling framework, the combination of AI with numerical weather prediction (NWP), the models have demonstrated improved forecasting accuracy and computational efficiency (S. Rasp et al., 2020). For Rajasthan, PINNs can enhance:

- Monsoon variability forecasting
- Drought cycle modeling
- Heatwave intensity prediction

These models can reduce the overfitting risks and can also help in interpreting the data correct data from atmospheric sciences

8.2 Real-Time Climate Dashboards Integrating IoT and AI Analytics

The future infrastructure systems can leverage Internet of Things (IoT) sensor networks that can be integrated with AI driven analytics (B. Sirmacek , R. Vinuesa, 2022) These real time dashboards can get following data:

- Temperature and humidity sensor data
- Soil moisture readings
- Reservoir levels
- Air quality metrics

These dashboards can support early-warning systems for tornadoes, storms and floods. And can also be used for water allocation strategies and monitoring the infrastructure stress. Sun and Scanlon (2019) highlight that real-time big data analytics significantly improves environmental management outcomes. By implementing the AI-powered decision dash boards at municipal and district levels, Rajasthan state can shift infrastructure governance from reactive to predictive frameworks.

8.3 Federated Learning for Secure Climate Data Sharing

Federated learning (FL) enables decentralized model training across multiple institutions/workstations without centralizing raw datasets. This approach is valuable in sharing climatic contexts where different agencies may be reluctant to share raw meteorological or infrastructural datasets. Overall, FL can help in maintaining a collaborative approach for climatic ecosystem. FL architecture enhance:

- Data privacy preservation
- Cross-institutional collaboration
- Distributed model optimization

8.4 Region-Specific Climate Adaptation Models for Arid Ecosystems

Rajasthan has a unique climatic condition with less precipitation in certain parts and humidity in some. Moreover, with deserts and high evaporation rates, the weather is always unpredictable in contrast with the rest of the world. Therefore, adapting AI-based models is a necessity with semi-arid regions. The future can focus on:

- Microclimate heat mapping
- Desertification progression modeling

- Groundwater depletion forecasting

With XAI and localized dataset, it can develop an efficient way to make new policies.

XI. CONCLUSION

By leveraging deep-learning, distributed computing, and decentralization of the dataset, Rajasthan can transform from reactive to predictive infrastructure. By making rural areas more technologically advanced, the unique climate can be better dealt with.

By addressing the technological gaps, the overall technological domain can ensure resilient climate adaptation. With collaboration on the region-specific framework, the insights given by AI models can be turned into actionable infrastructure and policies.

Lastly, Rajasthan has great potential for powering and utilising the technology, it needs a more efficient way to battle for sustainable future. Moving from traditional mindset, the state needs to be ready to adapt to the more advanced and efficient technology.

REFERENCES

[1] Sebesty´en, V., Czvetk´o, T., & Abonyi, J. (2021). The Applicability of Big Data in Climate Change Research: The importance of System of Systems Thinking. Frontiers in Environmental Science, 9. https://doi.org/10.3389/fenvs.2021.619092

[2] Nathawat, R., Singh, S.K., Sajan, B. et al. (2025) Integrating Cloud-Based Geospatial Analysis for Understanding Spatio-Temporal Drought Dynamics and Microclimate Variability in Rajasthan: Impli cations for Urban Development Planning. J Indian Soc Remote Sens. https://doi.org/10.1007/s12524-025-02139-6

[3] Li, B., & Qian, Y. (2024). Weather Prediction Using CNN-LSTM for Time Series Analysis: A Case Study on Delhi Temperature Data. ArXiv. https://arxiv.org/abs/2409.09414

[4] Nathawat, R., Gupta, S. K., Kanga, S., Singh, S. K., Chakraborty, S., Marazi, A., Sajan, B., Abouleish, M. Y., Meraj, G., Ali, T., & Kumar, P. (2025). Urban Green Infrastructure Planning in Jaipur, India: A GIS Based Suitability Model for Semi-Arid Cities. Sustainability, 17(6), 2420. https://doi.org/10.3390/su17062420

[5] Nathawat, R. et al. (2025). Integrating Cloud-Based Geospatial Analysis for Understanding Spatio-Temporal Drought Dynamics and Microclimate Variability in Rajasthan: Implications for Urban Development Planning. Journal of the Indian Society of Remote Sensing.

[6] Government of Rajasthan. (2026). Drone policy and AI data hub initiatives as announced in state budget. Times of India.

[7] Rajasthan Watershed Development Portal. (2024). Climate adaptation and watershed management strategies.

[8] Eshraghi, P., Talami, R., Dehnavi, A. N., Mirdamadi, M., & Zomorodian, Z.-S. (2024). Adopting Explainable-AI to investigate impact of urban morphology design on energy and environmental performance in dry arid climates.

[9] M. Reichstein et al., "Deep learning and process understanding for data driven Earth system science," Nature, vol. 566, pp. 195–204, 2019.

[10] S. Rasp et al., "WeatherBench: A benchmark dataset for data-driven weather forecasting," J. Adv. Model. Earth Syst., vol. 12, 2020.

[11] A. Y. Sun and B. R. Scanlon, "How can Big Data and machine learning benefit environment and water management," Environ. Res. Lett., vol. 14, 2019.

[12] Thorat, P. (2025). AI-powered predictive analytics for sustainable urban development: Addressing climate impacts of La Ni˜na and El Ni˜no. Journal of Information Systems Engineering & Management, 10(53s), 436–446.

[13] Sirmacek, B., & Vinuesa, R. (2022). Remote sensing and AI for building climate adaptation applications. Results in Engineering, 15(100524), 100524.

[14] Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... Bengio, Y. (2019). Tackling climate change with machine learning. arXiv preprint arXiv:1906.05433.

[15] Samek, W., Wiegand, T., & M¨uller, K.-R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296.

[16] https://www.ncei.noaa.gov/products/land-based-station/global-historical climatology-network-daily (GHCN-Daily)

[17] https://www.weatherlog.in/?baithak=793845263