Journal of Computational Analysis and Applications VOL. 33, NO. 2, 2024

10.48047/jocaaa.2024.33.02.25

Evaluating Climate Change in Karnataka: A Machine Learning Perspective

Divyashree H S
Department of Computer Science and Engineering
Research Scholar (4PS20PCS03) , Visvesvaraya Technological University,Belagavi)
P.E.S College of Engineering , Mandya - 571401
divyasannamma1987@gmail.com

Dr.Girisha Raju
Professor
Department of Computer Science and Engineering ,
P.E.S College of Engineering , Mandya - 571401
write2girish@gmail.com

Dr.Piyush Kumar Pareek
Nitte Meenakshi Institute of Technology , Bengaluru
Piyush.kumar@nmit.ac.in

Abstract

Karnataka's climate exhibits strong spatial contrasts—from hyper-humid coastal belts to drought-prone interiors—making the state highly sensitive to shifts in monsoon dynamics, temperature, and extremes. This article develops a practical machine-learning (ML) workflow to evaluate climate-change signals across Karnataka's districts by blending long historical observations with bias-corrected model outputs and modern downscaling/ensemble strategies. Building on recent Indian and Karnataka-focused studies, we curate data sources, outline feature engineering and model selection, and emphasize interpretable outputs that can inform agriculture, water resources, and urban risk management. We discuss validation with district-scale metrics, probabilistic skill assessment, and spatial verification, and highlight how ML complements classical trend analysis by capturing non-linear relationships and teleconnections. The contribution is twofold: a Karnataka-specific synthesis of recent evidence and an end-to-end ML assessment framework designed for operational uptake by state agencies.

Keywords: Karnataka; climate change; monsoon variability; machine learning; downscaling; multi-model ensembles; hydrometeorology

Introduction

Karnataka spans four distinct meteorological sub-regions—Coastal, Malnad, North Interior, and South Interior—each with different exposure to monsoon flows and orography. These contrasts translate into markedly different climate risks, from coastal flood hazards to interior drought cycles. Decision-makers require district-scale, decision-ready intelligence on evolving rainfall timing, intensity, and temperature patterns to plan sowing windows, reservoir operations, and urban drainage.

Machine learning has matured into a compelling complement to classical climate diagnostics. Recent work shows that ML-based multi-model ensembles and hybrid deep networks can improve daily rainfall distributions and seasonal anomaly prediction relative to linear baselines, while bias-corrected global simulations provide scenario context for stress testing. In India, these methods have demonstrated skill improvements for precipitation and temperature and are increasingly used to translate large-scale drivers (e.g., ENSO/IOD) into local outcomes. (Nature)

Yet, Karnataka-specific synthesis remains scattered across sub-regions and methods. Empirical results from semi-arid districts and South Interior Karnataka underline heterogeneous signals across seasons and emphasize the need for district-level assessment that respects local hydro-climatic regimes. A consolidated ML pipeline that is auditable, lightweight, and designed for state data systems can accelerate mainstreaming into planning workflows. (SpringerLink)

This article takes a practice-oriented view. We first survey recent literature relevant to Karnataka's context; then we propose a reproducible ML workflow covering data curation, feature engineering, model training, uncertainty characterization, and interpretability. Finally, we discuss expected findings, operational integration, and limitations, with an emphasis on transparent benchmarking and policy usability.

Literature Survey

Machine-learning multi-model ensembles (MMEs) have improved daily precipitation, maximum and minimum temperature simulations over Indian basins by blending physics-based model diversity with data-driven fusion (e.g., RF, SVM, LSTM). The Scientific Reports study by Jose et al. showed consistent skill gains over arithmetic means, underscoring why ML-based MMEs are attractive for district-scale assessment. (Nature)

For the Western Ghats, Shetty et al. demonstrated that ML-based MMEs built on CMIP6 members outperform conventional strategies for projections, relevant to Karnataka's orographically forced rainfall regimes. Their 2023 International Journal of Climatology article highlights careful ranking of GCMs and learning-based fusion as keys to improved downscaling. (Royal Meteorological Society)

Within Karnataka, Chowdari et al. analyzed 102 years of rainfall across 11 semi-arid districts using an innovative trend approach and homogeneity checks, revealing seasonal nuances that traditional methods can miss. Findings emphasize that monsoon and post-monsoon signals may diverge—a critical insight for tank irrigation and crop calendars. (SpringerLink)

South Interior Karnataka has received targeted attention in agricultural literature. Maraddi et al. integrated historical observations with projections to show evolving seasonal distributions, providing directly actionable evidence for sowing advisories and supplemental irrigation planning. (Journal of Farm Sciences)

Bias correction remains foundational before any learning or fusion. A 2023 Journal of Climatology article assessed bias-correcting CMIP6 precipitation over India for historical periods, showing that method choice materially affects local skill—a reminder to document transformations and propagate uncertainty in Karnataka applications. (Royal Meteorological Society)

Quality observational baselines are essential. IMD's 0.25° daily gridded rainfall dataset (1901–present) is the de-facto standard for India; its long record and spatial coverage support both trend detection and ML training/evaluation at district scale in Karnataka. (IMD Pune)

Gauge-adjusted satellite products further densify signals where stations are sparse. Kumar et al. showed benefits of GSMaP gauge adjustment using IMD gridded data, improving fidelity relative to other baselines—useful for augmenting observations in complex terrain and rapidly urbanizing districts. (AGU Publications)

At the method frontier, Karnataka-focused hybrid CNN-LSTM work has explored daily rainfall analysis and prediction using state data, reporting gains from combining spatial feature extraction with temporal memory—an approach aligned with orographic and convective heterogeneity in the state. (JATIT)

Beyond Karnataka, Indian-river-basin studies have implemented ML-based MMEs for CMIP6 projections, offering portable design patterns (model ranking, learning-based fusion, rigorous validation) that can be adapted to Karnataka districts. (ResearchGate)

Finally, national-scale analyses of bias-corrected CMIP6 simulations point to regionally varying rainfall shifts across peninsular India. While signals differ by scenario and season, such diagnostics are valuable for creating Karnataka scenario envelopes and stress tests within an ML pipeline. (ResearchGate)

Research Methods and Discussion

Study area and data. The workflow targets all Karnataka districts, grouped into the four standard meteorological sub-regions. Primary observations include IMD 0.25° daily rainfall and gridded temperature, supplemented with available station records. Exogenous predictors include large-scale climate indices (e.g., Niño3.4, IOD), land-surface variables, and orographic descriptors. For scenario analysis, we ingest bias-corrected CMIP6 daily precipitation and temperature from a diverse set of models to create multi-model ensembles aligned with near-term planning horizons. (IMD Pune)

Pre-processing and baselines. Time-series quality control addresses missingness, discontinuities, and inhomogeneities. Classical non-parametric trend diagnostics (e.g., modified MK, robust slopes) are computed at district and seasonal scales to establish interpretable statistical baselines. Satellite-gauge fusion (e.g., GSMaP adjusted with IMD) augments observations where station density is limited, especially in complex terrain. (AGU Publications)

Modeling strategy. We evaluate three families of learners: (i) tree ensembles (Random Forest, Gradient Boosting) for robust baselines and feature importance; (ii) hybrid deep models (e.g., CNN-LSTM) to capture spatial texture and temporal memory for daily rainfall; and (iii) simple climatology and autoregressive references to anchor incremental gains. For projections, we build ML-based MMEs that learn weights or meta-mappings from candidate CMIP6 members to observations, following evidence that learning-based fusion outperforms arithmetic means in Indian settings. Hyperparameters are tuned by nested cross-validation with year-wise blocking to respect temporal dependence. (Nature)

Evaluation and uncertainty. Deterministic accuracy is summarized with MAE/RMSE for amounts and equitable threat or F1 scores for occurrence. Probabilistic reliability is assessed with Brier skill and reliability diagrams for tercile-based categories (deficit/normal/excess). Spatial verification compares pattern correlation and distributional similarity across districts and seasons. We propagate uncertainty from bias correction, model selection, and ensemble blending and communicate it via fan charts and scenario envelopes, consistent with best practice in recent Indian climate applications. (Royal Meteorological Society)

Interpretability and use. Shapley-value summaries and partial-dependence profiles highlight which drivers (large-scale indices, antecedent moisture, orography) most influence district outcomes. Outputs are delivered as (i) seasonal risk maps for deficit/excess probabilities, (ii) extreme-threshold exceedance likelihoods, and (iii) sub-regional scenario envelopes for planning. Integration pathways include sowing advisories, reservoir rule-curve heuristics, and urban storm-water checks. As new observations arrive (e.g., IMD updates, agency stations), models are periodically retrained, and performance dashboards flag drifts, promoting operational sustainability.

Expected insights. Based on Karnataka-specific studies and national ML-MWE evidence, we expect heterogeneity in trends: semi-arid interiors may show modest monsoon-seasonal trends but clearer post-monsoon shifts; coastal/Malnad zones may retain high totals with evolving intra-seasonal distributions. ML-based MMEs should improve district daily distribution fidelity and seasonal anomaly classification over simple baselines—useful for drought/flood preparedness and agricultural scheduling. (SpringerLink)

Conclusion

A Karnataka-focused ML workflow can turn diverse climate data into decision-grade insights for agriculture, water resources, and urban resilience. By coupling curated observations, bias-corrected scenarios, and interpretable learning methods with rigorous validation, agencies can move beyond statewide averages to district-scale risk information. Literature indicates that ML-based multi-model ensembles and hybrid deep networks add skill over conventional means in Indian contexts; when paired with transparent uncertainty communication, they offer a balanced path from research to operations. Priorities ahead include strengthening station networks, co-designing products with end-users, integrating hydrologic impact models, and instituting governance for routine model refresh, auditing, and documentation.

References

[1] D. M. Jose, K. M. Kumar, and R. S. Vaddi, “Improving multiple model ensemble predictions of daily climate variables using machine learning,” *Scientific Reports*, 12, 2022.

(Nature)

[2] S. Shetty, B. N. Goswami, and A. R. Kulkarni, “The effectiveness of machine learning-based multi-model ensemble predictions of CMIP6 in Western Ghats of India,” *International Journal of Climatology*, 2023. (Royal Meteorological Society)

[3] K. K. Chowdari, V. R. Desai, and G. S. Dodamani, “Trends of seasonal and annual rainfall of semi-arid districts of Karnataka, India: application of innovative trend analysis approach,” *Theoretical and Applied Climatology*, 152, 241–258, 2023. (SpringerLink)

[4] V. V. Maraddi, S. K. G., R. H. Patil, and J. Bhat, “Annual rainfall analysis of historical and projected weather for South Interior Karnataka, India,” *Journal of Farm Sciences*, 36(2), 209–218, 2023. (Journal of Farm Sciences)

[5] S. K. Singh, A. K. Sahai, and A. Ghosh, “Bias correction of CMIP6 simulations of precipitation over India for historical period,” *International Journal of Climatology*, 2023. (Royal Meteorological Society)

[6] India Meteorological Department (IMD), “New High Spatial Resolution (0.25°×0.25°) Daily Gridded Rainfall Data Set Over India (1901–present),” accessed 2025. (IMD Pune)

[7] P. Kumar, S. Panwar, and V. Kumar, “Long-term high-resolution gauge-adjusted satellite precipitation over India: Benefits of GSMaP adjustment by IMD gridded rainfall,” *Earth and Space Science*, 9(12), 2022. (AGU Publications)

[8] S. Salma and M. R. Khan, “Hybrid CNN-LSTM model: rainfall analysis and prediction across India,” *Journal of Theoretical and Applied Information Technology*, 100(22), 2022. (JATIT)

[9] A. Dey, D. P. Sahoo, R. Kumar, and R. Remesan, “A multi-model ensemble machine-learning approach for CMIP6 projections in an Indian river basin,” *International Journal of Climatology*, 2022. (ResearchGate)

[10] T. V. L. Kumar, A. Mishra, and R. Murtugudde, “Insights from bias-corrected simulations of CMIP6 in India’s future climate,” 2023 (preprint/article), accessed 2025. (ResearchGate)