INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 28/03/2025

Accepted: 10/06/2025

Published: 13/07/2025

Citation: Dhwani N, Pranay P, Haragopal VV, Manjusha T (2025) Year-wise Longitudinal Comparative Analysis of Temperature Prediction Models for Hyderabad (2017–2024). Indian Journal of Science and Technology 18(27): 2179-2184. https://doi.org/10.17485/IJST/v18i27.559

* **Corresponding author**.

pettempranay@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Year-wise Longitudinal Comparative Analysis of Temperature Prediction Models for Hyderabad (2017–2024)

N Dhwani[1,2], P Pranay[1]*, V V Haragopal[3], T Manjusha[2]

1 Chaitanya (Deemed to be University), Warangal, Telangana, India
2 Little Flower Degree College, Uppal, Hyderabad, Telangana, India
3 Department of Mathematics, BITS, Hyderabad, Telangana, India

Abstract

Objective: The study aims to compare and analyze temperature forecasting models for Hyderabad from 2017 to 2024 in the context of modeling long-term accuracy and the influence of critical meteorological conditions on temperature. **Methods:** Past meteorological data, including temperature, humidity, solar radiation, wind speed, and precipitation, were analyzed. Five models for predicting were applied: Multiple Linear Regression, Decision Tree Regression, Random Forest Regression, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) Neural Networks. A chi-squared test determined the connections between temperature and qualitative meteorological features. The performance of models was evaluated based on R^2, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). **Findings:** Maximum and minimum temperatures had strong and persistent correlations with mean temperature throughout all years, whereas the effect of humidity decreased significantly in 2024. Between models, Multiple Linear Regression presented the best accuracy (R^2 = 0.989, RMSE = 0.4246) in 2024, closely followed by Random Forest. Decision Tree and ANN models had mid-level performance, whereas LSTM had poor performance with negative R^2 values for some years. These findings emphasize the strength of linear models and the weakness of deep learning for this data. **Novelty:** Through its comparative analysis spanning several years, this study presents revealing insights into the resilience of regression models, the drawbacks of deep learning, and the need for adaptive forecasting techniques in urban climate modeling.

Keywords: Temperature Prediction; Weather forecasting; Meteorological variables; Machine learning models; Hyderabad Climate

1 Introduction

Extreme weather events like storms, heat waves, and freezing conditions disrupt energy systems by damaging infrastructure and causing blackouts. Understanding these impacts is essential for developing resilient energy networks through technological upgrades and adaptive strategies [1]. Despite significant progress in numerical weather models, computing power, and real-time data, accurate weather forecasting remains difficult due to the inherently chaotic and sensitive nature of the atmosphere. Predictability limits still constrain long-range and high-impact event forecasting [2].

In rapidly growing cities like Hyderabad, rising urban temperatures driven by unplanned expansion and air pollution significantly impact public health, infrastructure, and environmental quality. Understanding Urban Heat Island dynamics is key to developing sustainable urban strategies and improving overall livability [3]. Forecasting meteorological variables such as temperature, humidity, solar radiation, and wind speed requires models that can capture their complex and non-linear relationships. Traditional methods like Multiple Linear Regression often lack accuracy, making way for more effective approaches that significantly improve prediction performance [4].

Predicting ozone levels is vital for managing air quality and minimizing health and environmental risks. Due to the complex interactions of pollutants, accurate forecasting supports timely actions and effective mitigation strategies [5]. Random Forest Regression (RFR) has demonstrated superior accuracy over traditional regression and SVM models in forecasting climatic parameters like solar radiation and wind speed. By effectively capturing complex relationships without relying on detailed atmospheric physics, RFR offers a reliable and cost-efficient alternative for weather prediction [6].

Recent advancements in deep learning have significantly enhanced temperature forecasting by utilizing Long Short-Term Memory (LSTM) networks. These models are highly effective at capturing temporal patterns and local dynamics, enabling accurate, high-resolution predictions at the suburban scale [7]. Deep learning models require large volumes of labeled data to perform effectively, but many real-world applications, including meteorology, suffer from data scarcity. This limitation makes it challenging to train reliable models, often leading to issues like overfitting and reduced generalization [8].

The CHyPP approach combines physics-based models with machine learning to improve atmospheric forecasts. It delivers more accurate predictions than traditional or standalone ML models, especially for surface temperature and humidity. A hybrid model combining machine learning with a global circulation model enhances the simulation of complex atmospheric processes. It improves forecasts of phenomena like stratospheric warming and ENSO by integrating additional climate variables [9] [10].

1.1 Objectives

- To examine historical weather records of Hyderabad for the years 2017 to 2024 for major meteorological parameters like temperature, humidity, wind speed, precipitation, and solar radiation.
- To analyze the statistical relationship between temperature and other meteorological variables using correlation analysis.
- To design and compare some predictive models (Multiple Linear Regression, Decision Tree, Random Forest, ANN, LSTM) for forecasting temperature.

2 Methodology

Hyderabad, the Telangana state capital, is located in south-central India on the Deccan Plateau. Hyderabad has a semi-arid climate, with hot summers, relatively cool winters, and seasonal monsoons. Urbanization in recent decades has rendered it highly prone to urban heat island effects, with altering land use and low vegetation cover influencing local meteorological conditions. In this research, Hyderabad has been targeted because of its potential as an expanding metropolitan city under climate variability.

The data collected from the India Meteorological Department (IMD), Begumpet, Hyderabad, Telangana, and the work examines temperature forecasting for Hyderabad City spanning from January 1, 2017, to December 31, 2024, covering eight complete years. The dataset includes daily records of key atmospheric variables such as maximum and minimum temperature, dew point, relative humidity, precipitation, wind speed and direction, wind gust, and solar radiation. This period was selected to capture both recent trends and inter-annual variability, including extreme weather patterns (https://docs.google.com/spreadsheets/d/1Yk86reo4DFqACmNWx5mGY0aMnXT1gdBX/edit?usp=drive_link&ouid=112858300260286252116&rtpof=true&sd=true).

Preprocessing was performed to maintain uniformity, and a chi-squared test was performed to validate associations between temperature and categorical meteorological data. These features were utilized in the following models: OLS Regression: Applied all chosen features under linear assumptions. Coefficients of features indicate the strength and direction of their impact on temperature.

Decision Tree and Random Forest: Both models can address linear and nonlinear relationships. Importance scores for features were derived to identify predictive strengths. Artificial Neural Network (ANN): Inputs were normalized and provided to a multilayer feedforward network. Hidden layers facilitate the modeling of complex, non-linear interactions among features and temperature. Long Short-Term Memory (LSTM): Due to its prowess in predicting temporal sequence, LSTM utilized lagged values (past 7 days of features) as input to adequately capture temporal dependencies. Five models were implemented for temperature prediction: Multiple Linear Regression, Decision Tree Regression, Random Forest Regression, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) Networks. The performance of the models was assessed using the R^2 Score, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The research provides a comparative review of the prediction accuracy over time, determining the best modeling methodology for temperature forecasting in Hyderabad.

3 Results and Discussion

3.1 Descriptive Statistics

From the descriptive statistics, the Comparative Summary of Meteorological Data for Hyderabad (2017 to 2024) reveals the following trends:

- **Temperature Trends:** As shown in **Figures 1, 2, and 3**, the average maximum temperature decreased marginally from **33.06°C (2017)** to **30.87°C (2024)**, indicating a cooling trend. This relates to the findings of[11], who reported links between climate drivers and temperature variability.
- **Humidity & Dew Point:** Humidity levels increased steadily from **64.15% (2017)** to **70.35% (2024)**, indicating rising atmospheric moisture. Correspondingly, dew points rose minimally, highlighting more humid conditions in later years. These results align with[12], who identified dew point temperature, precipitable water, and lifted index as significant indicators for forecasting precipitation and severe weather events over Hyderabad.
- **Precipitation:** The average precipitation increased slightly from **2.87 mm (2017)** to **2.99 mm (2024)**, with extreme rainfalls reaching **91 mm** in certain years. This trend aligns with[13], who examined annual and extreme rainfall trends in Hyderabad from 1990 to 2020, highlighting spatial-temporal variability driven by urbanization and land use changes.
- **Wind & Solar Radiation:** Wind speeds showed a slight reduction over the years from **13.43 km/h (2017)** to **13.91 km/h (2024)**, while wind gusts averaged around **32 km/h**. Solar radiation decreased from **222.15 W/m^2 (2017)** to **177.82 W/m^2 (2024)**, potentially due to higher cloud cover or pollution. Consistent with[14], who examined long-term changes in solar radiation and sunshine duration across India from 1985 to 2019, identifying regional disparities and a declining trend in solar energy potential.

3.2 Pearson correlation coefficients between temperature and key meteorological variables in Hyderabad (2017 - 2024)

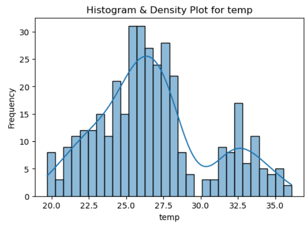
Correlation analysis between 2017 and 2024 indicates that temperature always highly correlates with maximum and minimum temperatures, verifying their direct relationship. Solar radiation highly correlates positively, supporting its heating function. Humidity highly negatively correlates with temperature, particularly in 2019 and 2020, indicating that higher humidity typically comes with cooler weather. Dew point correlations change from year to year, while wind variables such as speed and gust have weak or no correlations. Wind direction possesses a weak relationship in a few years, while rainfall has little to no correlation. These patterns indicate that temperature is predominantly controlled by TempMax, TempMin, solar radiation, and humidity.

3.3 Predictive analysis of temperature

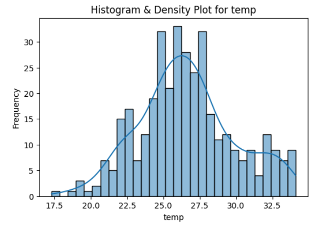
With RMSE values of low magnitude and high R^2 values (over 0.98), the Multiple Linear Regression (MLR) model performs best throughout and always outshines the rest with excellent predictive accuracy in all years[15]. This indicates that predictors and temperature bear a consistent linear relationship. Also effective are Random Forest (RF) and Decision Tree (DT) models; specifically, RF is effective in nonlinear trends,[16] as evidenced by its R^2 values above 0.96 in most years with comparably low RMSE. Compared to DT and RF, MLR does well with a lower $\mathbf{R^2}$ and higher RMSE.

The Artificial Neural Network (ANN) model demonstrates average performance, with mixed outcomes. R^2 scores vary from 0.62 to 0.96, while RMSE scores are higher, especially in 2023, reflecting some volatility. The Long Short-Term Memory (LSTM) model fares poorly in all the years, with low or negative R^2 values and high RMSE, which indicates that it might not be appropriate for this dataset or has been compromised due to overfitting, underfitting, or inadequate sequence data preparation.

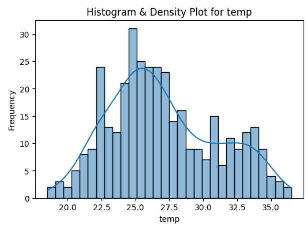
Year 2024 emerges as the statistically strongest model, boasting the highest R^2 (0.989) and extremely low RMSE (0.4246), with the strongest predictive capability among all years. Its regression equation is more finely tuned to both temperature extremes and humidity, and thus, the strongest year for predictive modeling within the study. This implies that 2024 data could have been influenced by more stable environmental trends or higher-quality data.



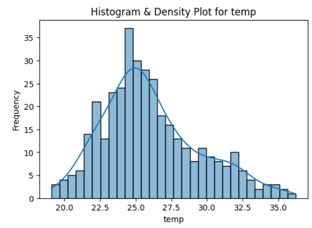
Histogram & Density plot for temperature in the Year 2017



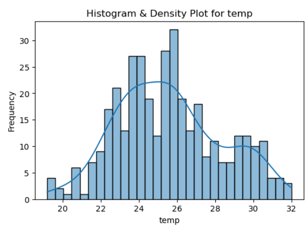
Histogram & Density plot for temperature in the Year 2018



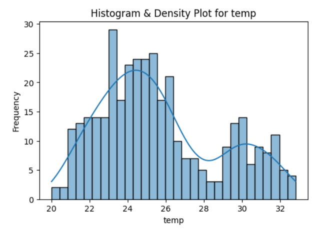
Histogram & Density plot for temperature in the Year 2019



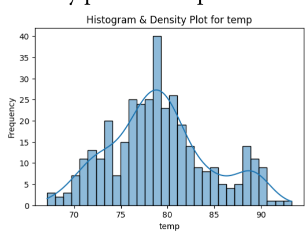
Histogram & Density plot for temperature in the Year 2020



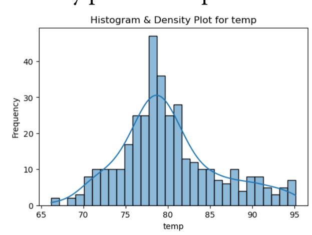
Histogram & Density plot for temperature in the Year 2021



Histogram & Density plot for temperature in the Year 2022



Histogram & Density plot for temperature in the Year 2023



Histogram & Density plot for temperature in the Year 2024

Table 1. Year-wise correlation of temperature with meteorological parameters (2017 to 2024)

Variable Pair	2017	2018	2019	2020	2021	2022	2023	2024
Temp ∼ TempMax	0.909	0.905	0.935	0.927	0.8653	0.914	0.88	0.916
Temp ∼ TempMin	0.7964	0.808	0.844	0.8353	0.7602	0.813	0.829	0.8414
Temp ∼ Dew Point	0.105	0.3977	0.0316	0.1959	0.2614	0.2764	0.253	0.0588
Temp ∼ Humidity	-0.454	-0.393	-0.6894	-0.6201	-0.3841	-0.421	-0.405	-0.1199
Temp ∼ Solarradiation	0.4922	0.4832	0.573	0.6976	0.3679	0.535	0.484	0.539
Temp ∼ Wind Speed	0.1917	0.1384	0.2673	0.1059	0.0624	0.0208	0.2374	0.0726
Temp ∼ Wind gust	0.12006	0.0692	0.1409	0.0026	0.0789	0.0244	0.253	0.0552

Table 2. Predictive analysis of temperature for the years (2017-2024)

Model	**2017**		**2018**		**2019**		**2020**		**2021**		**2022**		**2023**		**2024**	
	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**	R^2	**RMSE**
MLR	0.985	0.5754	0.981	0.4553	0.9863	0.4455	0.9918	0.2921	0.987	0.3439	0.9819	0.4246	0.9801	0.7239	0.989	0.4246
DT	0.9475	0.8322	0.9347	0.8435	0.9568	0.7914	0.9424	0.7733	0.9376	0.6439	0.9374	0.7896	0.9537	1.1048	0.9374	0.7896
RF	0.9729	0.5977	0.964	0.6267	0.9778	0.5674	0.9786	0.4710	0.9616	0.505	0.9609	0.6239	0.9718	0.8623	0.9609	0.6239
ANN	0.9523	0.7933	0.9179	0.9463	0.9633	0.7297	0.9635	0.6159	0.8658	0.9446	0.9278	0.8477	0.6239	3.1496	0.9329	0.8174
LSTM	-1.4817	1.7071	0.2049	2.0127	0.2919	1.2194	0.0307	1.5961	-0.162	1.6687	-1.3939	1.6766	0.3629	2.4887	-1.535	1.7253

Table 3. Year-wise Summary of Regression Coefficients for Temperature Prediction Models in Hyderabad (2017–2024)

Year	**Intercept** (β_0)	**Temp. Max** (β_1)	**Temp. Min** (β_2)	**Solar Radiation** (β_3)	**Humidity** (β_4)	R^2	**RMSE**
2017	3.52	0.436	0.489	-0.0003	-0.0221	0.985	0.5754
2018	1.89	0.5235	0.4253	-0.0013	-0.014	0.981	0.4553
2019	4.42	0.4622	0.4573	-0.0021	-0.0289	0.986	0.4455
2020	4.13	0.4425	0.4672	-0.0003	-0.0252	0.987	0.3439
2021	3.52	0.436	0.489	-0.0003	-0.0221	0.981	0.4246
2022	3.12	0.4312	0.4798	0.0011	-0.0161	0.982	0.4246
2023	5.87	0.4556	0.4812	0.0022	-0.0212	0.98	0.7239
2024	7.62	0.4503	0.4996	-0.0027	-0.0452	0.989	0.4246

4 Conclusions

The overall assessment of Hyderabad's weather data, spanning from 2017 to 2024, shows relative climatic steadiness, but with increasing trends of environmental change that may be initiated by urbanization. Both maximum and minimum temperatures consistently exhibit strong correlations with the mean temperature, as they play a significant role in shaping broad trends within temperature. Conversely, the contributions of humidity and solar radiation have fluctuated: humidity's contribution is seen to decrease in the latter years, indicative of lower atmospheric moisture possibly due to declining green cover or increasing urban surfaces, while solar radiation had higher correlations in the middle years (2019–2021) and leveled off thereafter, possibly suggesting clearer skies and lower monsoonal fluctuation.

The remaining parameters, like wind speed, gust, and rainfall, indicated weak correlation with temperature and thus no significant contribution to daily temperature variability. MLR and Random Forest predictive models yielded high R^2 and low RMSE, validating linear relationships and the appropriateness of such models in comparison to higher-order models like ANN and LSTM, which had inconsistent performance.

Notably, the rising intercepts in regression equations—most dramatically so in 2024—point towards an increasing base temperature, possibly the beginning of an urban heat island phenomenon. Overall, the results demonstrate that although Hyderabad's climate has been fairly consistent, such incremental but vital warming trends are observed, and hence, the pressing need for sustainable urban planning and climate-resilient environmental management.

5 Acknowledgments

The author would like to thank the India Meteorological Department, Begumpet, Hyderabad, Telangana.

- Author Contributions Statement
  - Dhwani. N: Basic statistical analysis, writing and editing the documentation
  - P. Pranay: Research Guidance
  - V.V. Haragopal: Conceptualization & Methodology
  - T. Manjusha: Analysis
- Conflict of statement: The authors declare no conflicts of interest.
- Ethical Compliance statement: This study does not involve human participants.
- Funding Statement: This research does not support any funding agency.
- Data Availability Statement: The dataset was taken from the India Meteorological Department (IMD), Hyderabad.

References

1) Gonçalves AC, Costoya X, Nieto R, Liberato ML. Extreme weather events on energy systems: a comprehensive review on impacts, mitigation, and adaptation measures. *Sustainable Energy Research*. 2024;11(1):4. https://doi.org/10.1186/s40807-023-00097-6.
2) Brotzge JA, Berchoff D, Carlis DL, Carr FH, Carr RH, Gerth JJ, et al. Challenges and opportunities in numerical weather prediction. *Bulletin of the American Meteorological Society*. 2023;104(3):E698–E705. https://doi.org/10.1175/BAMS-D-22-0172.1.
3) Mathew A, Aljohani TH, Shekar PR, Arunab KS, Sharma AK, Ahmed MFM, et al. Spatiotemporal dynamics of urban heat island effect and air pollution in Bengaluru and Hyderabad: implications for sustainable urban development. *Discover Sustainability*. 2025;6(1):134. https://doi.org/10.1007/s43621-025-00860-3.
4) Segovia JA, Toaquiza JF, Llanos JR, Rivas DR. Meteorological variables forecasting system using machine learning and open-source software. *Electronics*. 2023;12(4):1007. https://doi.org/10.3390/electronics12041007.
5) Zhou Z, Qiu C, Zhang Y. A comparative analysis of linear regression, neural networks, and random forest regression for predicting air ozone employing soft sensor models. *Scientific Reports*. 2023;13(1):22420. https://doi.org/10.1038/s41598-023-49899-0.
6) Meenal RM, Michael PA, Pamela D, Rajasekaran E. Weather prediction using random forest machine learning model. *Indonesian Journal of Electrical Engineering and Computer Science*. 2021;22(2):1208–1215. https://doi.org/10.11591/ijeecs.v22.i2.pp1208-1215.
7) Yu M, Xu F, Hu W, Sun J, Cervone G. Using Long Short-Term Memory (LSTM) and Internet of Things (IoT) for localized surface temperature forecasting in an urban environment. *IEEE Access*. 2021;9:137406–137418. https://doi.org/10.1109/ACCESS.2021.3116809.
8) Alzubaidi L, Bai J, Al-Sabaawi A, Santamaría J, Albahri AS, Al-Dabbagh BSN, et al. A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *Journal of Big Data*. 2023;10(1):46. https://doi.org/10.1186/s40537-023-00727-2.
9) Arcomano T, Szunyogh I, Wikner A, Pathak J, Hunt BR, Ott E. A hybrid approach to atmospheric modeling that combines machine learning with a physics-based numerical model. *Journal of Advances in Modeling Earth Systems*. 2022;14(3):e2021MS002712. https://doi.org/10.1029/2021MS002712.
10) Arcomano T, Szunyogh I, Wikner A, Hunt BR, Ott E. A hybrid atmospheric model incorporating machine learning can capture dynamical processes not captured by its physics-based component. *Geophysical Research Letters*. 2023;50(8):e2022GL102649. https://doi.org/10.1029/2022GL102649.
11) Kumar R, Kuttippurath J, Gopikrishnan GS, Kumar P, Varikoden H. Enhanced surface temperature over India during 1980–2020 and future projections: causal links of the drivers and trends. *npj Climate and Atmospheric Science*. 2023;6(1):164. https://doi.org/10.1038/s41612-023-00494-0.
12) Taiyab M, Rahiman MK, Basha HA. Impact of Surface Dew Point Temperature, Precipitable Water and the Lifted Index on the Precipitation over Hyderabad (India). *Current World Environment*. 2024;19(1):207. https://doi.org/10.12944/CWE.19.1.19.
13) Hari D, Kasa RR. Assessment of Spatiotemporal Trends in Rainfall and Rainfall Extremes in the Hyderabad City, India. *Grassroots Journal of Natural Resources*. 2024;7(2):216–235. https://doi.org/10.33002/nr2581.6853.070211.
14) Kumar BS, Phukan R, Boragapu R, Nalage CB, Tathe AD, Hosalikar KS. Understanding the climatology and long-term trends in solar radiation using ground-based in-situ observations in India. *Mausam*. 2024;75(2):349–372. https://doi.org/10.54302/mausam.v75i2.6238.
15) Tahir A, Akhter M, Uddin Z, Sarim M. Neural Network and Regression Methods for Estimation of the Average Daily Temperature of Hyderabad for the Years 2018-2020. *International Journal of Economic and Environmental Geology*. 2021;12(2):87–91. https://doi.org/10.46660/ijeeg.v12i2.107.
16) Sandunil K, Bennour Z, Mahmud HB, Giwelli A. Effects of tuning decision trees in random forest regression on predicting porosity of a hydrocarbon reservoir. A case study: Volve oil field, North Sea. *Energy Advances*. 2024;3(9):2335–2347. https://doi.org/10.1039/D4YA00313F.