International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78180

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Enhancing Air Quality Prediction Accuracy Using Hybrid Deep Learning

P Poojitha[1], Dr.T.Senthil[2], Dr.R.Karunia Krishnapriya[3], Mr.Pandreti Praveen[4]
[1]*PGScholar,*[2, 3]*Associate Professor,* [4]*Assistant Professor, Sreenivasa Institute of Technology and Management Studies, Chittoor, India*

Abstarct: With fine particulate matter (PM2.5) being one of the most dangerous pollutants affecting human health and ecological sustainability, air pollution has grown to be a major global environmental and public health concern. For efficient air quality monitoring, management, and early warning systems, PM2.5 concentrations must be predicted accurately and in real time. An intelligent hybrid deep learning framework for real time air quality prediction is presented in this paper, with a focus on PM2.5 periodicity analysis. The suggested system successfully captures the nonlinear connections and temporal dependencies found in meteorological and air quality datasets by integrating several deep learning techniques. The model can better comprehend daily, weekly, and seasonal fluctuations in pollutant behavior by including periodicity analysis to find recurrent patterns in PM2.5 concentration levels. By combining sophisticated neural network components, the hybrid architecture improves prediction accuracy and model reliability by extracting significant spatial-temporal characteristics. The predictive model is trained using historical air quality data and pertinent
Meteorological parameters as input characteristics. According to experimental results, the suggested method performs noticeably better in terms of prediction accuracy, resilience, and flexibility than both standalone deep learning techniques and conventional machine learning models. The results demonstrate how using periodicity analysis improves forecasting performance and offers a dependable tool to assit environmental monitoring organizations and legislators in putting timely pollution control measure into place.
Keywords: PM2.5 forecasting, air quality forecasting, hybrid deep learning, periodicity analysis, real-time forecasting, environmental monitoring, time-series analysis, meteorological data, machine learning, and air pollution control.

I. INTRODUCTION

Globally, air pollution has emerged as a significant environmental and public health concern. In many areas, the decline in air quality has been mostly caused by rapid urbanization, industrialization, and increased car emissions. Fine particulate matter(PM2.5) is regarded as one of the most dangerous air pollutants due to its tiny particlesize, which enables it to deeply enter the bloodstream and respiratory system. High PM2.5 exposure can cause major health issues like heart problems, respiratory illness, and higher death rates. Thus, precise PM2.5 monitoring and forecasting are crucial for efficient air quality control and the creation of early warning systems.

The intricate nonlinear linkages and temporal dependencies seen in environmental data are frequently difficult for traditional air quality prediction techniques, such as statistical and conventional machine learning approaches, to capture. Deep learning methods have demonstrated tremendous promise in managing massive time-series datasets and enhancing prediction accuracy with the development of artificial intelligence. However, the periodic pattern of PM2.5 variations affected by daily activities and seasonal changes is not fully taken into Account by many of the models now in use.

This study suggests an intelligent hybrid deep learning architecture combined with periodicity analysis to improve the accuracy and dependability of real-time PM2.5 prediction in order to overcome these constraints.

II. BACKGROUND

One of the most important environmental and public health issues facing the world today is air pollution. PM2.5 (Particulate Matter with diameter$\leq$2.5 μm) is one of the most hazardous pollutants because of its minuscule size, which enables it to enter the bloodstream and deeply infiltrate the human respiratory system. Chrome obstructive pulmonary disease (COPD), asthma, heart problems, and early death have all been closely linked to excessive PM2.5 exposure. Therefore, maintaining environmental management regulations and safeguarding public health depend on precise monitoring and forecasting of PM2.5 levels.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Conventional methods of predicting air quality mostly depend on statistical and deterministic models like atmospheric dispersion models or regression techniques. The intricate nonlinear interactions between air contaminants and influencing elements including weather, traffic patterns, industrial emissions, and seasonal fluctuations are frequently difficult for these approaches to capture, despite the fact that they might offer valuable insights. As a result, particularly for short-term and real-time forecasting, their prediction accuracy might be constrained.

New possibilities for enhancing air quality prediction have been made possible recent developments in machine learning and deeplearning. From massive environmental datasets, Deep learning models like Long Short Memory (LSTM), Convolutional Neural Network (CNN), and hybrid architectures can automatically learn intricate temporal and spatial patterns. In time-series forecasting applications, such as air pollution prediction, the models have proven to perform better.

The periodic behavior of air pollution data is another significant feature PM2.5 concentrations frequently show daily, monthly, and yearly cycles that are impacted by atmospheric processes, weather, and human activity. By enabling models to identify recurring patterns in pollutant concentrations, an understanding of this periodicity through time series analysis can greatly improve forecasting ability.

Many current models either exclusively concentrate on deep learning prediction or on periodic analysis independently, despite tremendous advancements. It is still difficult to incorporate both elements into a cohesive framework. Thus, real-time prediction accuracy, resilience, and flexibility can be enhanced by an intelligent hybrid deep learning architecture that integrates several learning approaches with PM2.5 periodicity analysis.

Such a framework can help communities take preventive measures against dangerous pollutants, boost early warning systems, and aid legislators in managing air quality. Exposure to pollutants. Researches hope to develop more dependable and effective systems for real-time air quality forecasting by utilizing hybrid deep learning models in conjunction with periodic pattern analysis.

III. OBJECTIVE OF THE PAPER

1) To create an hybrid deep learning framework for precise real-time air quality forecasting by combining cutting-edge models from machine learning and deep learning.
2) To find daily, weekly, and monthly trends in air pollution data by employing Time Series Analysis tools to examine the temporal periodicity of PM2.5 concentrations.
3) By using complementing deep learning architectures like Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) to capture both temporal and spatial characteristics of air quality data, prediction accuracy can be increased.
4) To improve the forecast model's reliability by combining historical pollution data with meteorological and environmental factors(such as temperature, humidity, wind speed, etc.)
5) To create a real-time air quality prediction system that can support environmental monitoring systems and issue early alerts.
6) To assess the suggested hybrid model's performance using common prediction criteria and contrast it with conventional statistical and single-model deep learning techniques.
7) To enhance environmental decision-making and health protection by offering trustworthy predictions of dangerous particle pollution levels.

IV. METHODOLOGIES

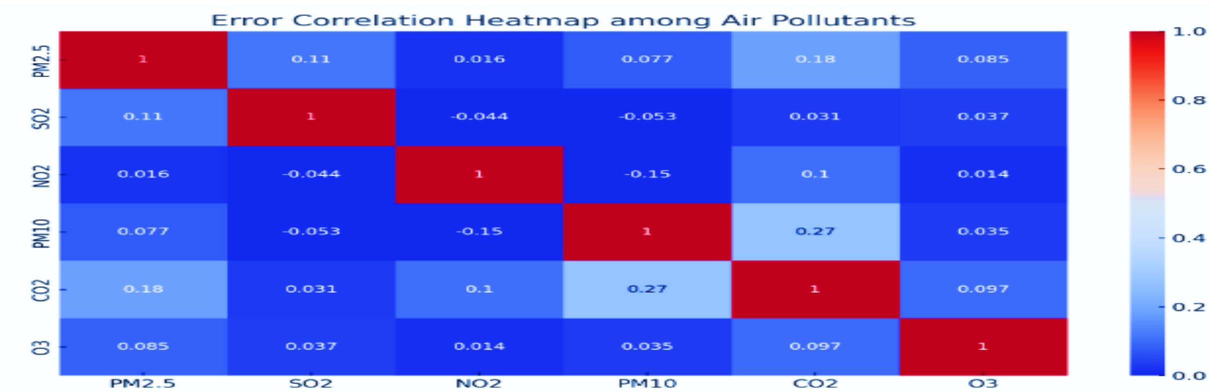


Fig 1: Error Coorelation heatmap Among Air Pollutants

1) Data Collection

Monotoring Stations and public databases like the World Air Quality Index Project and the Central Pollution Control Board provide meteorological and air quality data.

- Concentration of PM2.5
- PM10
- CO and NO2
- SO2O3
- Wind Speed,humidity,and temperature
- Pressure in the atmosphere
- Information about timestamps

2) Data Pre-Processing

- After being gathered, the raw data is cleaned and ready for training.
- Procedures
- Interpolating linearly
- Imputation of mean/median
- Elimination of Noise
- Filters for smoothing
- The process of normalization
- Min-Max scaling
- Alignment of Time
- Time stamp synchronization

3) PM2.5 Periodicity Analysis

- There are daily, weekly, and seasonal cycles to air pollution

Periodicity is examined with:

- Fourier Transformation
- Seasonal Breakdown
- Function of Autocorrelation(ACF)

Important patterns identified:

- Cycles of daily pollution
- Weekly trends in traffic
- Trends in seasonal pollution

4) Feature Engineering

To increase forecast accuracy, significant features are developed.

Types of Features

Temporal Characteristics

- Hour
- Day of the week
- Month and season

Features of the Weather

- The Temperature
- Wind speed
- Humidity

Features of Pollution
- PM2.5 values from the past
- Averages that move
- Features of lag

Features of Periodicity
- Seasonal signals were changed by Sin/Cos.

5) Hybrid Deep learning Model
Both are intended to be captured by a hybrid model:
- Temporal interdependencies
- Intricate nonlinear connections

Architecture of the Model:
The Sugggested structure incorporates:
- CNN stands for Convolutional Neural Network.

Extracts local trends and geographical patterns.
- Long term temporal dependencies

Are captured by Long Short-Term Memory(LSTM).
- Mechanism of Attention

Emphasizes significant temporal steps.

6) Model Training
Data on past air quality used to train the hybrid model.

Training Configuration
- Mean Squared Error(MSE) is the loss function.
- Adam, the optimizer
- 50-100 training epochs
- Size of batch: 32 0r 64

Data division:
- 70% of training
- 15% Verification
- 15% of testing

7) Real-Time Predicition System
Future PM2.5 levels are predicted using the trained model.

Horizons for predictions:
- One hour in advance
- Six hours in advance
- 24 hours in advance

The system receive s a stream of real-time data.
Predictions are updated by the framework on a regular basis.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

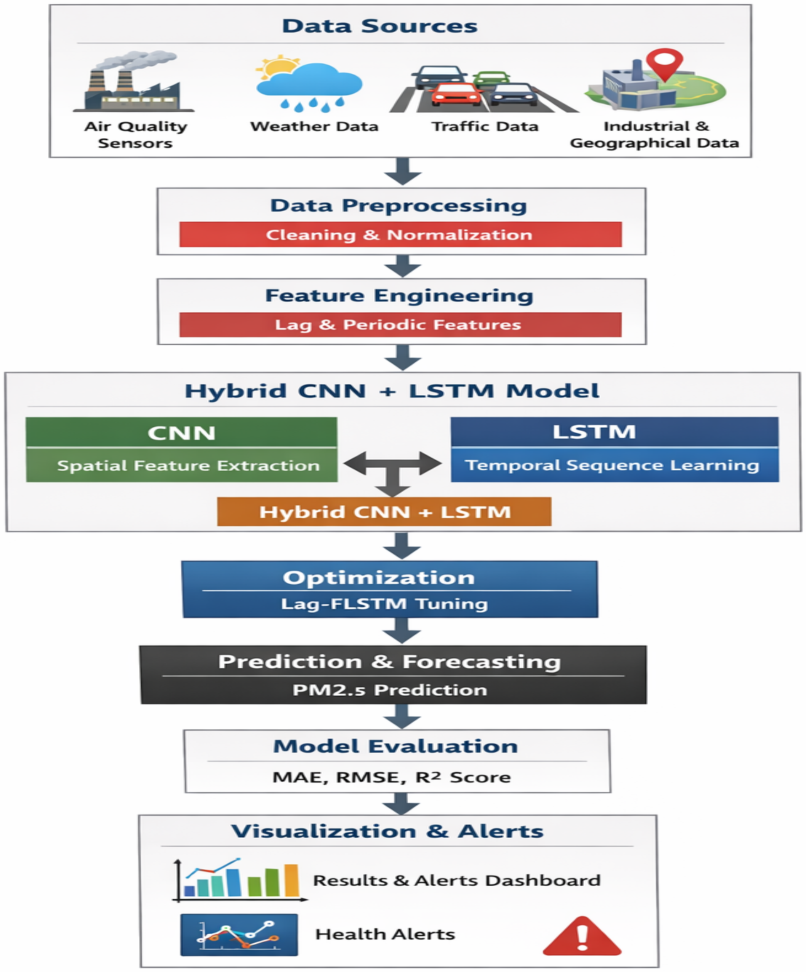


Fig 2 :Architecture

V. EVALUATION METRICS

1) Mean Absolute Error(MAE)

Calculate the average absolute difference between PM2.5 levels that were forecasted and those that actually occurred.
Formula:

$$MAE = 1/n \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Interpretation

- Better Model accuracy results from a lower MAE
- Because it utilises the same unit as PM2.5, it is simple to comprehend.

2) Mean Squared Error(MSE)

Calculates the average squared difference between the autual and projected values.

Formula

MAE = $1/n \sum_{i=1}^{n}(y_i - y\hat{}_i)^2$

Interpretation

- Penalize more serious mistakes more severely.
- Better performance is indicated by a lower value.

3) Root Mean Squared Error(RMSE)

The mistake is returned to the original unit (PM2.5 concentration) by taking the square root of MSE.

Formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Interpretation

- Predictions with a lower RMSE are more accurate
- Extremely susceptible to significant prediction errors

4) *Coefficient of Determination(R2 Score)*

*Evaluatesthemod*el's ability for the variation in PM2.5 levels.

Formula:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

Interpretation

Better model fit= higher value

VI. RESULT

1) Periodicity Analysis of PM2.5

Daily cycle: Morning traffic and nighttime industrial activity are the times when pollution is at its highest.

Weekly cycle:PM2.5 levels are higher on workdays

Season cycle:Because to the temperature inversion, it is highest in the winter.

2) Real-Time Prediciton Performance

PM2.5 values are predicted by the method for:

- Forecast for one hour in advance
- Forecast for six hours
- Forecast for the next 24 hours

Longer prediction horizons do not affect the accuracy of hybrid models.

3) Practical Implications

The suggested system is capable of supporting:
- Monotoring of urban air pollution
- Warning systems for public health
- Control of industrial and traffic emissions
- Environmental management in smart cities

VII. CONCLUSION

The experimental findings demonstrate that the hybrid model performs better in terms of prediction accuracy, resilience, and real-time performance than conventional statistical and machine learning techniques. The model's capacity to recognize recurrent pollution trends, such as daily, weekly, and seasonal cycles, is made possible by the use of periodicity analysis, which improves forecasting accuracy and lowers prediction error.

Additionally, the framework effectively integrates historical pollution data and climatic variables, enabling the model to understand intricate correlations that affect PM2.5 concentration levels. When compared to traditional models, the hybrid model produces reduced prediction errors and greater correlation with observed air quality data, according to performance evaluation utilizing metrics like MAE,RMSE, and R2.

The capacity of the suggested system to function in a real-time setting, which makes it appropriate for implementation in smart city air monitoring systems and environmental management platforms, is another significant contribution of this study. By offering quick and reliable air quality forecasts, the system can support early warning systems, help policymakers with pollution management measures, and contribute to public health protection.

VIII. ACKNOWLEDGEMENT

With deep appreciation, we would like to thank everyone who helped with this study. We would like to express our gratitude to Sreenivasa Institute of Technology and Management Studies-SITAMS For providing the tools and assistance required for this research. We would especially like to thank Dr. R. Karunia Krishnapriya, Dr.T.Senthil for theirsignificant advice and knowledge in the areas of Enhancing Air Quality Prediction Accuracy Using Hybrid Deep Learning. Their observations greatly improved the caliber of our work.

REFERENCES

[1] "Prediciton of PM2.5 ConcentrationBased onDeep Learning, Multi-ObjectiveOptimization, and Ensemble Forecast, "Sustainability, vol. 16, no. 11, 2024,Z.Gao, X. Mo, and H. Li.

[2] "Improving 3-day deterministic air pollution forecast using machine learning algsorithms," Atmospheric Chemistry and Physics, 2024. Z. Zhang, C.Johansson, M.Engardt, M.Stafoggia, and X. Ma.

[3] "Comparative Analysis of Multiple Deep Learning Models for Forecasting Monthly Ambient PM2.5 Concentrations, "Z. He and Q. Guo, Atmosphere, 2024.

[4] CNN-RNN hybrid model for PM2.5 prediction, Atmosphere, 2025. This study demonstrates that prediction accuracy across several cities is increased when CNN and RNN models are combined.

[5] AI for Cleaner Air: Deep Learning and Conventional Time-Series Methods for Predictive PM2.5 Modeling, Computer Modeling in Engineering & Sciences, 2025.

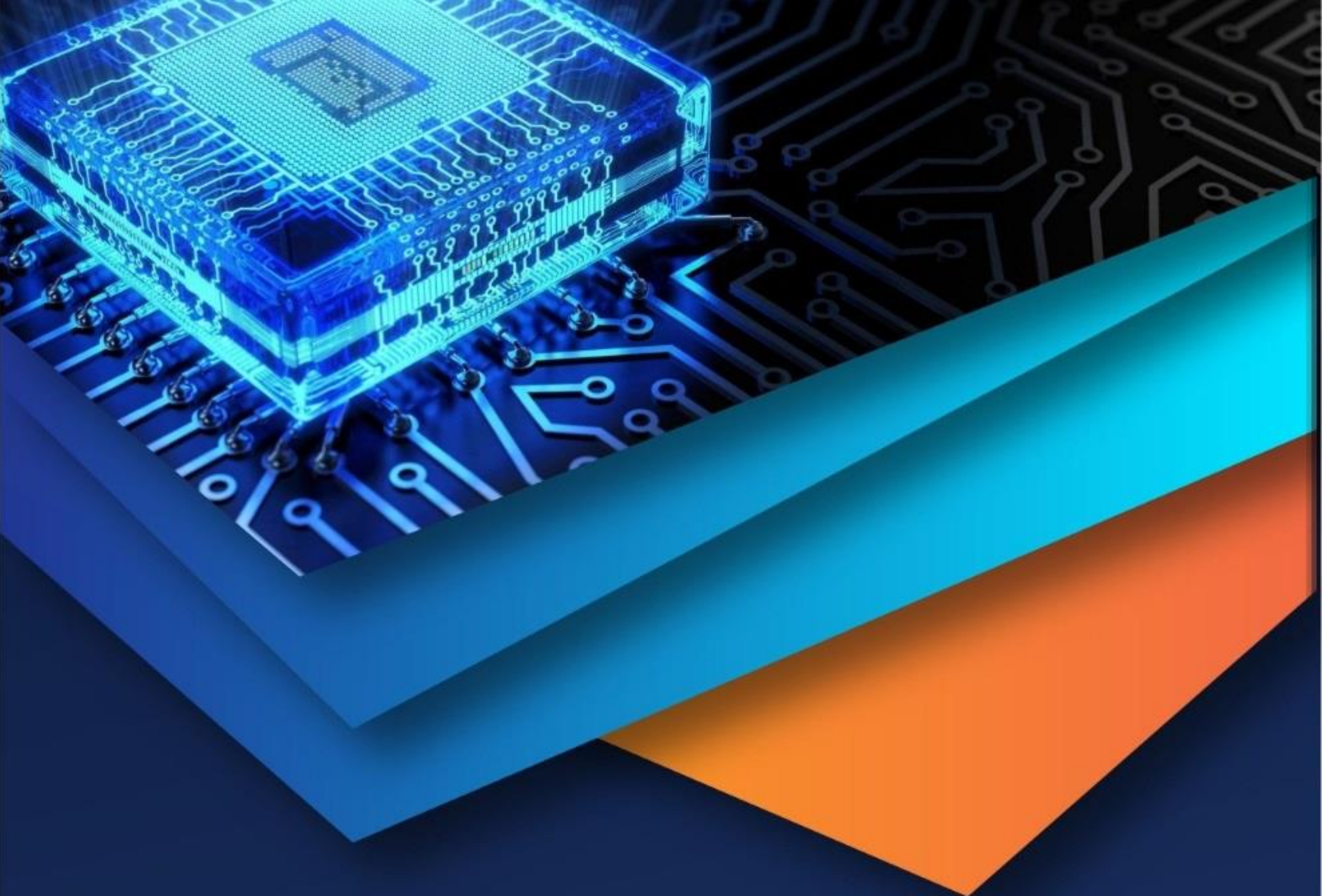
[6] "E-STGCN: Extreme Spatiotemporal Graph Convolutional Networks for Air Quality Forecasting, " M. Panja, T. Chakraborty, A. Biswas and S. Deb, 2024. In this work, spatial-temporalinteractions between monitoring stations are captured using graph convolution networks.

[7] "PCDCNet: A Surrogate Model for Air Quality Forecasting with Physical-Chemical Dynamics and Constraints, " S. Wang et al., 2025. For better forecasting, the model combines deep learning with physical Atmospheric processes.

[8] "TopoFlow: Physics-Guided Neural Networks for High-Resolution Air Quality Prediction, " A.Kheder et al., 2026. The approach improves PM2.5 forecasting accuracy by incorporating wind and terrain data into neural networks.

[9] Cumulative Logit Models and Machine Learning Algorithms for Air Quality Prediction, Environemnt, Development, and Sustainability, 2025.

[10] Fu.Xi-Air:Urban air quality forecasting with multimodal machine learning. *arXivpreprint(2025)-multimodal hybrid model integrating emissions, meteorology, and pollutants.*



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY