***Report for end-semester evaluation of CE 498 course***

**Clustering Methods for the Analysis of Spatio-Temporal Distribution of Precipitation and Temperature Data in Northeast India**

**Submitted**

**By**

**Rohan Kumar Ishwar**

Under the supervision of

**Dr. Sreeja Pekkat**



**Department of Civil Engineering**

**Indian Institute of Technology Guwahati**

**November 2024**

**CERTIFICATE**

It is certified that the work contained in the project report entitled “ **Clustering Methods for the Analysis of  Spatio-Temporal Distribution of Precipitation and Temperature Data in Northeast India** ”, by **Rohan Kumar Ishwar** (210104089) has been carried out under my/our supervision and that this work has not been submitted elsewhere for the award of a degree or diploma.

Date: 27-11-2024 Signature

**Dr. Sreeja Pekkat**

Department of Civil Engineering

Indian Institute of Technology Guwahati

**ABSTRACT**

This study explores the application of advanced clustering techniques—K-Means, DBSCAN, and Hierarchical Clustering—to analyze the spatio-temporal distribution of precipitation and temperature in Northeast India, with the objective of uncovering regional climate patterns and trends. Each method is evaluated for its efficacy in managing spatial heterogeneity, temporal variability, and noise. Comparative analysis highlights DBSCAN's strength in detecting irregular and non-linear patterns, K-Means' efficiency in identifying structured clusters, and Hierarchical Clustering's capacity to reveal nested relationships. The findings underscore the potential of clustering to identify precipitation hotspots and temperature gradients, providing actionable insights for climate adaptation strategies, resource planning, and disaster mitigation. The report concludes by emphasizing the practical applicability of clustering techniques for analyzing complex spatio-temporal climatic data and outlines plans for further analysis in the next phase.

**ACKNOWLEDGEMENT**

I would like to extend my heartfelt gratitude to my supervisor, Dr. Sreeja Pekkat, Associate Professor, Department of Civil Engineering, Indian Institute of Technology Guwahati, for her invaluable guidance, insightful suggestions, and unwavering encouragement throughout the progression of this study and the preparation of this thesis.

My sincere thanks also go to Abdul Rahman (PhD Scholar) for his support, positive outlook on this work, and invaluable feedback throughout the course of this project.

I am deeply appreciative of their contributions and thank them wholeheartedly.

Date:27-11-2024 **Rohan Kumar Ishwar**

|  |  |  |  |
| --- | --- | --- | --- |
| **TABLE OF CONTENTS** | | | |
| **Page no.** | | | |
| **Certificate** | | 2 | |
| **Abstract** | | 3 | |
| **Acknowledgement** | | 4 | |
| **List of Figures** | | 7 | |
|  | |  | |
| **Chapter 1: Introduction** | |  | |
|  | * 1. General   2. Study area   3. Climate data | | 8  9  10 | |
|  | * 1. Clustering analysis   2. Motivation for Work | | 10  10 | |
| **Chapter 2**: **Literature Review** | |  | |
|  | 2.1 Introduction to Climate Data and Spatio-Temporal Analysis  2.2 Datasets for Climate Analysis  2.3 Classification of PCMs | | 11  11-12 | |
|  | 2.3.1 Organic phase change compounds  2.3.2 Inorganic phase change compounds  2.3.3 Eutectics | | 12-13  13-14  14  14-15  15-16 | |
| 2.4 Methods of PCM incorporation in concrete  2.4.1 Direct methods of PCM incorporation in concrete  2.4.2 Indirect methods of PCM incorporation in concrete | |
| **Chapter 3: DesignBuilder Software**  3.1 Evaluation of Building Energy Efficiency Using Design  Builder Software  **Chapter 4: Objectives and Scope of Study**  4.1 Research Gaps 18  4.2 Objectives 18  4.3 Scope of Study 18  **Chapter 5: Methodology 19**  **Chapter 6: Conclusion**  6.2 Future Work 20  **REFERENCES** | | 17 | |
|  |  | |  | |

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
|  | **List of Figures** |  |
| **Figure No.** | Title | **Page No.** |
| 1 | World total primary energy supply by fuel in 2010 | 8 |
| 2 | Energy consumed by building sector in U. S | 9 |
| 3 | Sensible heat and Latent heat storage for the case solid- liquid | 9 |
|  |  |  |

**Chapter 1**

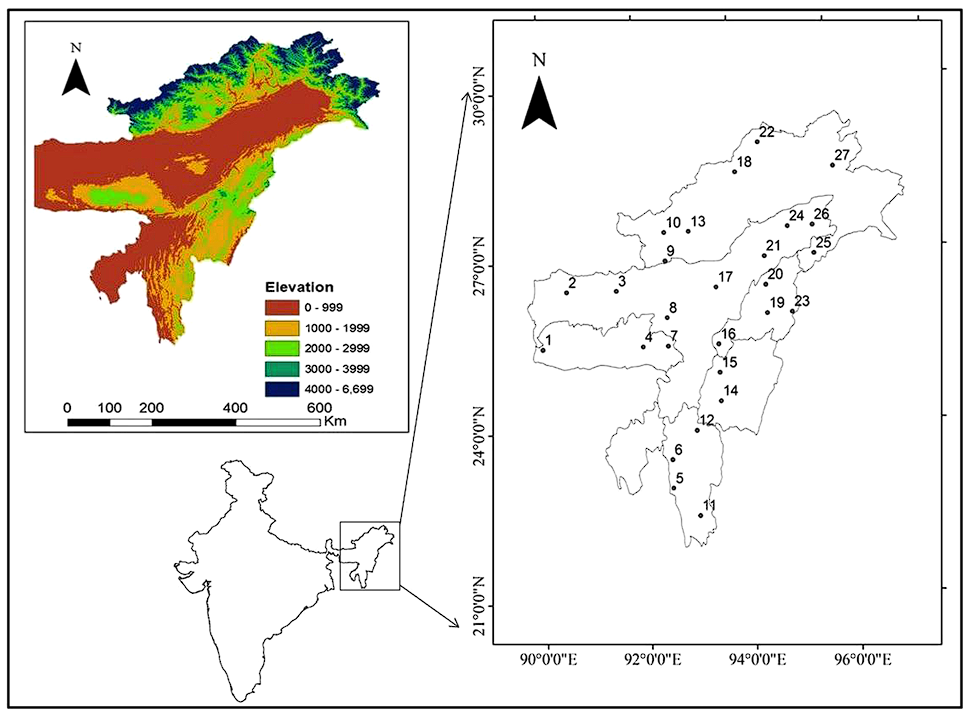
**INTRODUCTION**

* 1. **General**

Climate variability and change have become critical global concerns due to their far-reaching impacts on ecosystems, economies, and human livelihoods. The Northeast region of India, renowned for its complex topography and diverse climatic conditions, is particularly vulnerable to climate-induced challenges. This region has witnessed a significant rise in extreme weather events, including heavy precipitation, floods, and temperature fluctuations, which have resulted in considerable socio-economic disruptions. According to recent reports, India experienced a 26% increase in extreme rainfall events between 2001 and 2020, with Northeast India contributing a substantial share. Such events exacerbate issues related to agriculture, water resource management, and disaster preparedness, underscoring the urgency of understanding regional climatic patterns.

**1.2 Study area**

Northeast region of India(as shown in fig.1) which consists of the eight states Assam, Meghalaya, Arunachal Pradesh, Tripura, Nagaland, Mizoram, Manipur and Sikkim is selected for the study. This region extends from latitude 22.4°N to 28.7°N and longitude 88.2°E to 96.5°E. Elevation varies from 28 m above mean sea level to 7000 m above mean sea level. Region is very diverse in nature, and plains mainly comprise Brahmaputra and Barak valleys. Onset of monsoon occurs from middle of May and continues till October. On average, the NE region receives about 2450 mm of rainfall. The Cherrapunjee–Mawsynram range receives rainfall as high as 11,500 mm, about 60% area under forest with Arunachal Pradesh hav ing about 80% of its area under different kinds of forest, while Assam has the minimum percentage of forest area (30%). Large altitude differences and varying physical features are the main reason of diverse climate from near tropical to temperate. The annual rainfall in the region is received mainly from the southwest monsoon, annually (Das et al. 2009; Dash et al. 2012; Goyal 2014). Northeast India is one of the most vulnerable regions toward climate change, and the largest threat to its biodiversity and hydrological system is from changing climate (Ravindranath et al. 2011). For studying the extreme precipitation indices and its trend, 27 stations from 0.5° × 0.5° gridded precipitation and temperature data from India Meteorological Department (IMD) were selected for the period of 1965–2014.



**Fig. 1** Study area and location of selected station used to make gridded dataset

**1.3 Climate Data**

The increasing availability of high-resolution climatic datasets, driven by advancements in satellite observations and ground-based measurements, has revolutionized the field of climate science. These datasets offer unprecedented opportunities to analyse complex climate dynamics at finer spatial and temporal scales. However, the inherent complexity and vast volume of such datasets present significant analytical challenges. Extracting actionable insights requires methodologies capable of capturing subtle spatial variability and dynamic temporal interactions, which are often overlooked by traditional statistical approaches. The gridded data helped us analyse the trends using Mann–Kendall nonparametric test (Mann 1945; Kendall 1948) was used for calculating the trends of precipitation indices.

**1.4 Clustering Analysis**

Clustering analysis is an unsupervised learning technology, which can discover the structure and other information contained in the data itself without knowing the correct results in advance (Amin et al., 2020; Kar et al., 2015; Zhou et al., 2020). There are three common clustering methods: K-means clustering, hierarchical clustering and Density-based Spatial Clustering of Application with Noise (DBSACN) (Long et al., 2020; XIONG et al., 2020; Zheng et al., 2020). This approach facilitates the segmentation of datasets into meaningful clusters, which can reveal hidden relationships and regional characteristics that are critical for understanding climate variability and change.

By leveraging clustering methods such as K-Means, DBSCAN, and Hierarchical Clustering, it is possible to delineate climate zones, identify hotspots of extreme weather, and track seasonal transitions with greater precision. These insights are invaluable for addressing pressing challenges in climate-sensitive regions, such as optimizing agricultural practices, managing water resources, and improving disaster preparedness. The integration of clustering with high-resolution data analytics represents a significant step toward informed decision-making, enabling policymakers and researchers to better understand and mitigate the impacts of climate variability on ecosystems and human systems.

**1.5 Motivation for work**

The region's intricate topography and diverse climatic systems necessitate advanced analytical approaches to decipher spatio-temporal climate dynamics. With the advent of high-resolution datasets from satellite observations and ground-based systems, traditional statistical techniques fail to fully exploit the multidimensional complexity of the data. By deploying state-of-the-art clustering algorithms, this study seeks to uncover latent patterns, delineate precipitation anomalies, and identify temperature gradients, thereby enabling precise climate modeling and enhanced regional adaptive strategies for disaster risk reduction and sustainable resource management.

**Chapter 2**

**LITERATURE REVIEW**

* 1. **Introduction to Climate Data and Spatio-Temporal Analysis**

The study of precipitation and temperature distributions requires high-resolution spatio-temporal datasets to capture complex climatic variations. Advances in remote sensing technologies and ground-based observation networks have significantly improved the granularity of climate datasets, providing robust inputs for analytical frameworks. However, deriving actionable insights from these datasets poses challenges due to their sheer volume and dimensionality. This necessitates the use of machine learning approaches like clustering to extract patterns and trends from the data.

* 1. **Datasets for Climate Analysis**
     1. **Sources and Characteristics**

Climate studies often rely on datasets from sources such as:

* **Satellite Observations**: Datasets like TRMM (Tropical Rainfall Measuring Mission) and GPM (Global Precipitation Measurement) offer high temporal and spatial resolution for precipitation data.
* **Ground-Based Stations**: Meteorological station networks provide localized and accurate measurements of precipitation and temperature.
* **Reanalysis Data**: Products like ERA5 and MERRA2 integrate observational data with model outputs for comprehensive climatic datasets.

These datasets are characterized by high dimensionality, temporal continuity, and spatial variability, requiring advanced computational techniques for analysis. The complexity of these characteristics—interdependencies across time, space, and variables—necessitates advanced computational techniques capable of handling the intricacies of spatial-temporal interactions.

**2.2.2 Challenges in Dataset Utilization**

* **Heterogeneity**: Integration of data from multiple sources introduces inconsistencies.
* **Missing Values**: Incomplete observations often hinder robust analysis.
* **Scale Variability**: Differences in spatial and temporal resolutions affect data harmonization.
  + 1. **Data Preprocessing**

Outline pre-processing steps, such as handling missing values, standardizing data, or aggregating temporal resolutions (e.g., daily to monthly). By systematically addressing the preprocessing steps, the dataset becomes well-prepared for clustering, ensuring robust and interpretable results for spatio-temporal climate analysis.

* 1. **Clustering Techniques**

Traditional statistical approaches often fall short in addressing such multidimensional interdependencies, necessitating the use of machine learning methods like clustering. Clustering techniques are pivotal in identifying climate zones, precipitation hotspots, and temperature gradients. The following subsections review the clustering methods applied in this study:

**2.3.1 K-Means Clustering**

* **Overview**: K-Means is a partition-based clustering method that minimizes intra-cluster variance. It is widely used for structured datasets due to its computational efficiency.
* **Applications in Climate Studies**: Studies have demonstrated the utility of K-Means in classifying precipitation zones and temperature regimes. For example, K-Means has been employed to cluster rainfall patterns across monsoonal regions.
* **Limitations**: The method assumes spherical cluster shapes and requires the pre-specification of the number of clusters (k), making it less effective for irregularly shaped climate zones.

**2.3.2 DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**

* **Overview**: DBSCAN identifies clusters based on density, making it effective for datasets with noise and irregular cluster shapes.
* **Applications in Climate Studies**: DBSCAN excels in detecting precipitation anomalies and extreme events. Its ability to classify noise points is particularly valuable for identifying outliers in spatio-temporal climate data.
* **Limitations**: DBSCAN's performance is sensitive to parameters like epsilon (eps) and minimum samples, which require careful tuning based on data characteristics.

**2.3.3 Hierarchical Clustering**

* **Overview**: Hierarchical clustering builds a dendrogram to represent nested clusters, offering a flexible approach to cluster analysis without predefined cluster numbers.
* **Applications in Climate Studies**: Hierarchical clustering has been used to analyse nested climate patterns, such as sub-regional rainfall behaviours and temperature gradients.
* **Limitations:** The method is computationally intensive and sensitive to distance metrics and linkage methods, which can affect the resulting cluster structure**.**

**2.3.4 Comparative Insights from Literature**

The literature underscores the strengths and applications of various clustering methods in spatio-temporal climate analysis:

* **K-Means:** Effective for structured, low-noise datasets with well-defined clusters.
* **DBSCAN:** Robust for identifying irregular clusters and handling noise, particularly useful in detecting extreme climatic events.
* **Hierarchical Clustering:** Suitable for capturing nested relationships and multi-scale patterns, often applied in exploratory climate research.

The selection of a clustering method is influenced by dataset characteristics, such as noise levels, spatial variability, and temporal dynamics, as well as the specific analytical goals. A combined approach leveraging these methods provides a holistic understanding of spatio-temporal climatic patterns, enhancing the accuracy and depth of insights.

**Chapter 3**

**Methodology**

**3.1. Data Preprocessing**

**3.1.1 Handling Missing Values**

The preprocessing of climatic datasets (as shown in fig.2) involved addressing data(daily mean precipitation in mm) gaps caused by sensor malfunctions, poor connectivity, or data loss during transmission. Missing data were treated using:

* **Temporal Interpolation:** Suitable for filling gaps in continuous time-series data to preserve temporal trends.
* **Grid-wise Mean Imputation:** Addressed spatial gaps by averaging precipitation or temperature values across nearby grid points, minimizing spatial distortion.

The combination of these methods maintained the dataset's integrity while ensuring reliable input for downstream clustering analysis.

A black screen with white numbers

Description automatically generated

**Fig. 2** Sample Dataset Representation: Extracted Precipitation Metrics

**3.1.2 Temporal Aggregation**

To simplify the analysis of climatic patterns, daily data were aggregated into:

* **Monthly Means:** Highlighted seasonal variability, particularly useful for detecting monsoonal behaviour in the dataset.
* **Seasonal Aggregation:** Divided data into three climatic seasons (winter, monsoon, and summer), allowing a high-level analysis of intra-annual variability. For easy handling of data we divided each season having 4 months as shown in fig. 3.
  + - Winter- November to February
    - Summer- March to June
    - Monsoon- July to October

**3.1.3 Standardization**

Climatic variables displayed large variability due to differences in precipitation intensity and temperature ranges across regions. Z-score normalization was applied to:

* Reduce biases caused by extreme values.
* Improve clustering performance by ensuring equal weighting of variables across all grid points.

This preprocessing laid the foundation for accurate clustering and reliable trend analysis. Finally, we had daily, monthly and seasonal data with corresponding latitude and longitude after this step.

**3.2. Trend Analysis**

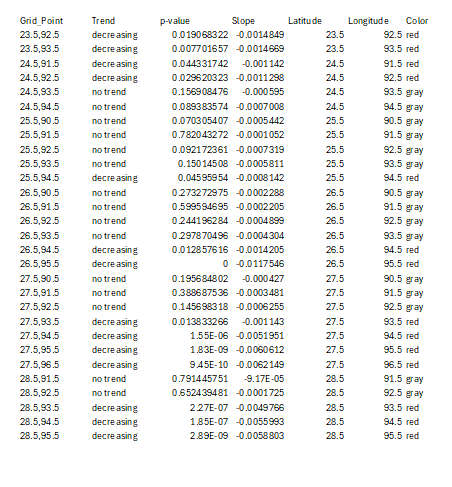
**3.2.1 Mann-Kendall Test and Sen’s Slope Estimator**

**Significant Trends:** The Mann-Kendall test identified:

* **Positive Precipitation Trends:** Concentrated in northern high-altitude regions, potentially linked to orographic effects.
* **Negative Trends:** Notable in the southern plains, suggesting reduced monsoonal activity.
* The **p-value** indicates the probability that the observed trend is due to random chance. A **small p-value** suggests a statistically significant trend (generally people take <0.05).

**Quantified Slopes:** Sen’s Slope Estimator calculated trends in mm/year. The results:

* The **Sen’s Slope** is the **median** of all individual slopes. It highlights localized hotspots of climatic variability.
* Provided quantitative metrics for comparative spatial analysis.



A grid of red and black dots

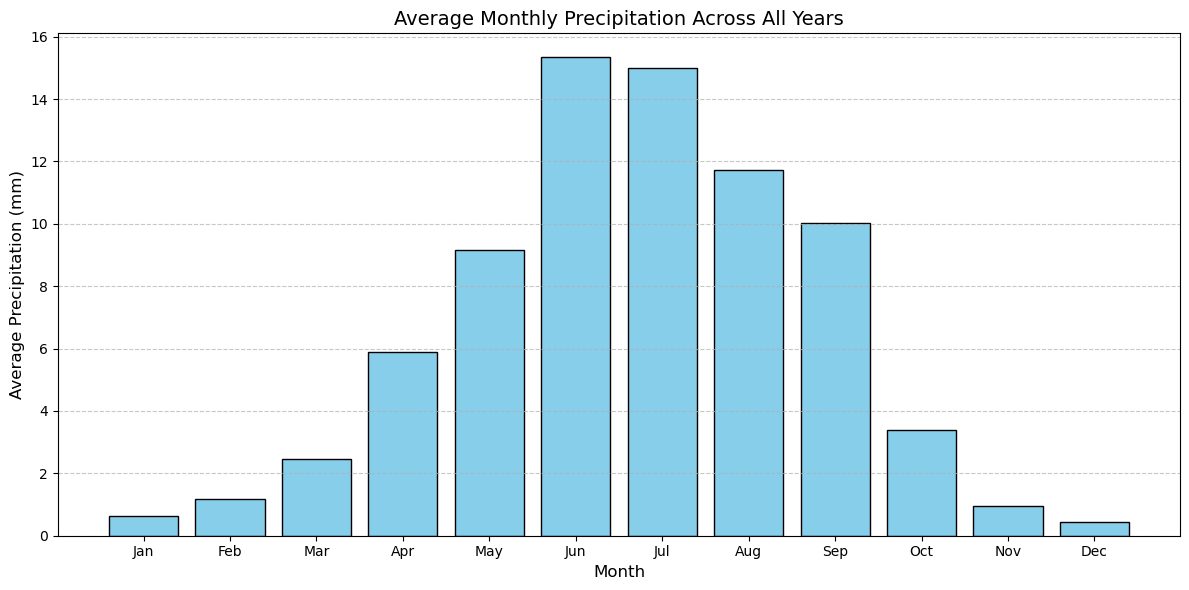
Description automatically generated

**Fig. 3** Data after trend analysis

**3.3 Temporal Analysis**

**3.3.1 Monthly Aggregation Analysis**

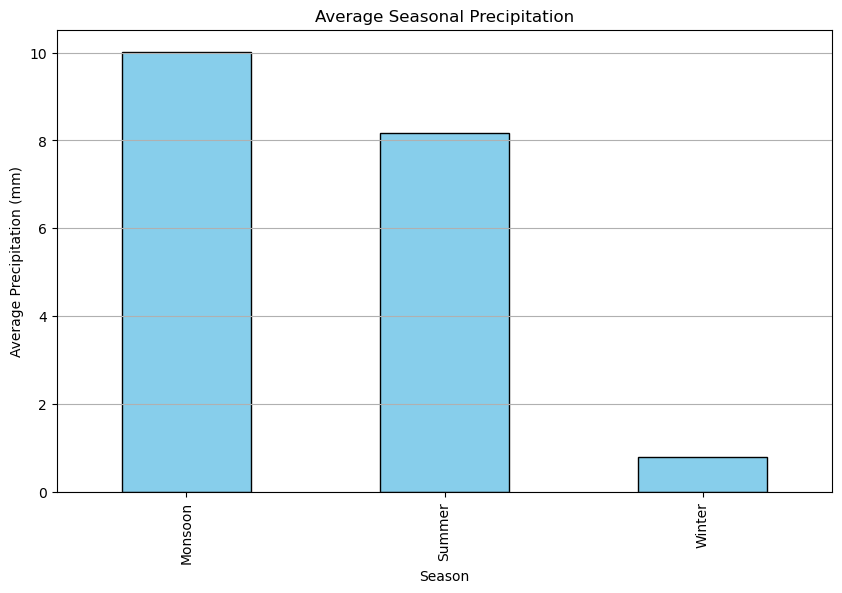
* **Precipitation Peaks:** Observed during the monsoon season (June–September), consistent with the region's dependence on monsoonal rainfall that can be seen in fig. 4.
* **Dry Spells:** Winter months (December–February) showed negligible precipitation, aligning with historical drought patterns.



**Fig. 4** Average monthly precipitation across all years

**3.3.2 Seasonal Aggregation Analysis**

* **Monsoonal Dominance**: Seasonal analysis highlighted the dominance of monsoon rainfall (June–September) across the region, contributing over 70% of the annual precipitation. This is clearly represented in the seasonal precipitation plot in Fig. 5, where the spatial intensity of rainfall aligns with Northeast India's dependence on monsoonal patterns.
* **Post-Monsoon Transition**: The post-monsoon period (October–November) exhibited moderate precipitation levels, reflecting the gradual withdrawal of the monsoon and localized rainfall events.
* **Winter Dryness**: Minimal precipitation was observed during winter months (December–February), indicating prolonged dry spells typical of the region's seasonal drought history. This is corroborated by the temporal patterns visualized in Fig. 4.
* **Pre-Monsoon Activity**: Early pre-monsoon months (March–May) showed rising precipitation levels, likely driven by convectional activity and localized thunderstorms.

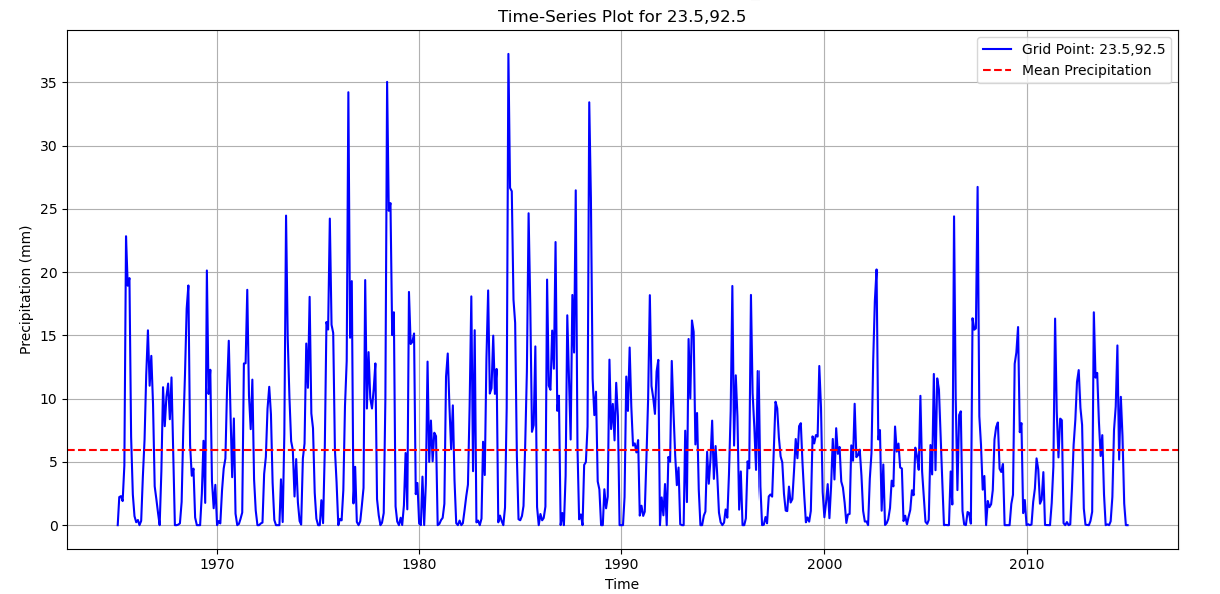


**Fig. 5** Average seasonal precipitation

**3.3.3 Time Series Analysis at Grid Level**

The temporal distribution for a representative grid location (e.g., 23.5, 92.5) is shown in Fig. 6, illustrating clear seasonal cycles:

* **Peak Activity**: Noticeable peaks during monsoonal months with significant rainfall accumulation.
* **Periodic Variations**: Clear periodic oscillations reflecting seasonal climatic variability.
* **Low Noise**: Consistent patterns validate the preprocessing steps in removing outliers and ensuring data integrity.

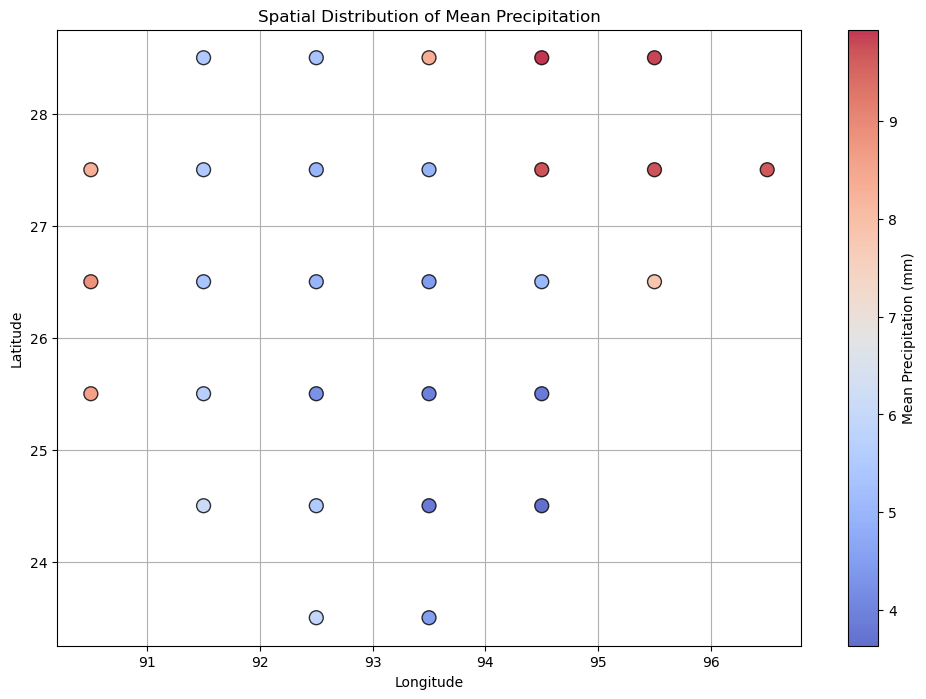


**Fig. 6** Sample time series plot of latitude:23.5, longitude:92.5

**3.4 Spatial Distribution**

Spatial analyses revealed:

* **High Precipitation Zones:** Concentrated in northeastern highlands, such as Cherrapunji.
* **Dry Zones:** Identified in lower valleys, reflecting reduced monsoonal dependency.



**3.5. Clustering Analysis**

**3.5.1 K-Means Clustering**

**Performance Analysis:**

* K-Means effectively segmented structured regions, identifying six distinct clusters based on precipitation and temperature distributions.
* To determine the optimal number of clusters (k), the elbow method and silhouette method was employed as represented in Fig. 7, which evaluates the cohesion and separation of clusters to identify the most suitable value for k, ensuring well-defined and meaningful clusters.
* We used optimal\_K =6 based on silhouette method result.

A graph with a line

Description automatically generated

A graph with a number of clusters

Description automatically generated

**Fig. 7** Methods for finding optimal K

**Cluster Insights:**

* + Regions with homogeneous precipitation patterns, such as the Brahmaputra Valley, formed uniform clusters.
  + High-elevation areas, like Meghalaya, emerged as unique clusters due to distinct rainfall patterns.

**Advantages:**

* Efficient for structured datasets with low noise.
* Computational simplicity allowed scalability for large datasets.

**Limitation:**

* Failed to capture irregular or non-linear patterns, as evidenced by the merging of heterogeneous climatic zones in some regions.

**Clusters:**

We obtained 6 clusters and namely cluster-3 had the highest number of grid-point in it as shown in Fig. 8, it also visually shows the representation of clusters obtained from K-mean clustering algorithm on precipitation data.

A black screen with white text

Description automatically generated

A graph of a graph

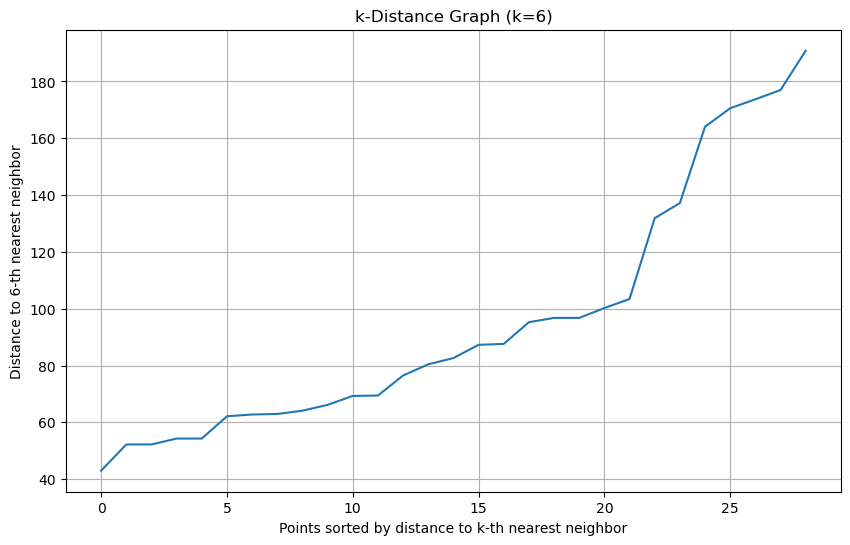
Description automatically generated with medium confidence

**Fig. 8** K-means clustering

**3.5.2 DBSCAN Clustering**

**Robustness to Noise:**

* DBSCAN proved effective in isolating anomalous grid points as noise, particularly those affected by extreme climatic events like flash floods.
* **Optimal Parameters:**
  + Eps=105.0 and Min\_samples=7, determined using a k-distance graph (Fig. 9).



**Fig. 9** Plot to find optimal parameters of DBSCAN

**Key Findings:**

* **Irregular Clusters:** Highlighted unique patterns in highly variable regions, such as areas impacted by shifting monsoon patterns.
* **Anomalies as Noise:** Noise points captured rare events, aiding in disaster risk mapping.

**Advantages:**

* Adaptable to irregular spatial and temporal patterns.
* Resilient to outliers.

**Limitations:**

* Parameter sensitivity required extensive tuning, particularly in densely sampled regions.

**Clusters:**

* We obtained 3 clusters and namely cluster-0 had the highest number of grid-point in it as shown in Fig. 10, it also visually shows the representation of clusters obtained from DBSCAN clustering algorithm on precipitation data.

A black screen with white text

Description automatically generated

A grid of dots and lines

Description automatically generated

**Fig. 10** DBSCAN clustering

.

**3.5.3 Hierarchical Clustering**

**Uncovering Nested Relationships:**

* Hierarchical clustering provided insights into nested climatic behaviors, with clusters revealing multi-scale patterns.

**Cluster Characteristics:**

* Larger regions were hierarchically subdivided, offering granular insights into sub-regional climate dynamics.
* Dendrogram analysis was performed using Ward's method (Fig. 11) revealed clear parent-child relationships among climatic zones.

A diagram of a diagram

Description automatically generated with medium confidence

**Fig. 11** Dendrogram analysis

**Advantages:**

* Ideal for exploratory analyses requiring a hierarchical understanding of climatic systems.
* Captured nested relationships that other methods missed.

**Limitations:**

* Computational demands were high, especially for larger datasets.

**Clusters:**

* We obtained 5 clusters and namely cluster-5 had the highest number of grid-point in it as shown in Fig. 12, it also visually shows the representation of clusters obtained from Hierarchical clustering algorithm on precipitation data.

A black screen with white text

Description automatically generated

A grid of dots with numbers

Description automatically generated

**Fig. 12** Hierarchical Clustering

**Chapter 4**

**Objectives and Scope of Study**

**4.1 Research gaps**

* There is an absence of research in macro encapsulating inexpensive inorganic salt for usage in mortar board.
* No studies are presented on reduction of thermal bridging in building using PCM incorporated concrete.

**4.2 Research objectives**

In order to create more energy-efficient buildings, it is necessary to design, develop, and evaluate the performance of thermal storage modules made with PCM that may be embedded in walls, roofs and other structural building components.

* To examine the influence of variables such as w/c ratio of cement mortar and dosage of PCM on mechanical and thermal properties of PCM-incorporated cement mortar board.

**4.3 Scope of the work**

The scope of the study is limited to the following with respect to raw materials used, manufacturing method adopted, and availability of tests: -

In preliminary study, OPC-43 grade cement is used as binding material conforming to IS269 and zone -2 fine aggregate as per IS-383 is used to prepare mortar mix. Superplasticizer (Fosroc Auramix 350) is used to get workability of 80%. Inorganic PCM Calcium chloride hexahydrate (CaCl2.6H2O) is used for the study which has a melting point of 29℃.

**Chapter 5**

**METHODOLOGY**

**5.1 Preparation of specimen**

Before the casting of PCM incorporated mortar board, control mix is to be casted to check the basic characteristics and also to compare the properties with the PCM incorporated mix.

|  |  |
| --- | --- |
| **Specimen** | **Test conducted** |
| 50x50x50 mm cube | Compressive test |
| 40x40x160 mm prism | Flexure test |
| 400x400x20mm slab | Model testing |
| 300x300x20 mm slab | Flexure test |

**Hardened Properties Test Methods**

**Compressive Strength:** To study compressive strength six cubes (for each design density and for each parameter) of size 50 mm cube specimens are casted. The specimens are then subjected to water curing for 25 days and 3 days air curing. The compressive strength tests are carried out in accordance with ASTM C 796 after 28 days.

**Flexural strength:** Flexural strength is indirect tensile strength of mortar in bending. Flexural strength of mortar will be performed as per ASTM C348. For each sample 3 specimen of size 40mm x 40mm x 160mm will be cast and their average test result will be reported.

**Small scale experimental model:** A 20 mm thick mortar board with different percentages of PCM and another without PCM will be caste. The surface area of mortar board will be 400 mm x 400 mm. The thermocouples will be connected with temperature data logger. After 28 days of curing, the board will be place on a horizontal surface above the ground level. Then heater with various temperature

will be placed above the board and it will be kept for 4 hours (sun light is maximum 4 hours). Then the temperature at different depth of mortar board and phase lag will be found out from temperature data logger. Further from this study effect of thickness on temperature control will be found out. In fresh stage heat of hydration will be found out and in hardened state thermal effectiveness will be found out.

**Chapter 6**

**Conclusions**

This study demonstrates the potential of integrating Phase Change Materials (PCMs) into mortar boards to improve building energy efficiency and thermal performance. Using DesignBuilder software, the operational energy of buildings with PCM-incorporated walls was compared, highlighting significant energy savings and better thermal regulation. The findings confirm the feasibility of PCM integration as a sustainable solution for reducing energy demands in buildings. Future work will focus on experimental validation to strengthen these conclusions and refine the application of PCMs in construction**.**

**6.2 Future Work**

In the next phase of this research, similar analysis will be done to the temperature dataset and focus on integrating supplementary datasets, such as land use and atmospheric circulation models, to refine clustering precision and validate findings through empirical evaluations under real-world climatic conditions.

**References**

1. UN Environment and International Energy Agency, Towards a zero-emission, efficient, and resilient buildings and construction sector, Global Status Report 2017 (2017).

2. K. Peippo, P. Kauranen, P.D. Lund, A multicomponent PCM wall optimized for passive solar heating, Energy and Buildings 17 (1991) 259–270. [2] A.K. Athienitis, C. Liu, D. Hawes, D. Banu, D. Feldman, Investigation of the thermal performance of a passive solar test-room with wall latent heat storage, Building and Environment 32 (5) (1997) 405–410.

3. R. Kelly, Latent heat storage in Building Materials. Paper of AMEC Design. http://www.amec.com, 1997 (retrieved 30.10.08).

4. M.M. Farid, A.M. Khudhair, S.A.K. Razack, S. Al-Hallaj, A review on phase change energy storage: materials and applications, Energy Conversion and Management 45 (2004) 1597–1615.

5. S.M. Hasnain, Review on sustainable thermal energy storage technologies, Part I: heat storage materials and techniques, Energy Conversion and Management 39 (11) (1998) 1127–1138.

6. S. Jing-cang, M. Peng-sheng, A novel solid–solid phase change heat storage material with polyurethane block copolymer structure, Energy Conversion and Management 47 (2006) 3185–3191.

7. D. Feldman, M.A. Khan, D. Banu, Energy storage composite with an organic PCM, Solar Energy Materials 18 (1989) 333–341.

8. D. Feldman, M.M. Shapiro, D. Banu, C.J. Fuks, Fatty acids and their mixtures as phase-change materials for thermal energy storage, Solar Energy Materials 18 (1989) 201–216.

9. P. Kauranen, K. Peippo, P.D. Lund, An organic PCM storage system with adjustable melting temperature, Solar Energy 46 (5) (1991) 275–278.

A.F. Rudd, Phase-change material wallboard for distributed thermal storage in buildings, ASHRAE Transactions: Research 99 2 (1993) 339–346.

10. B. Zalba, J.M. Marín, L.F. Cabeza, H. Mehling, Free-cooling of buildings with phase changing materials, International Journal of Refrigeration 27 (2004) 839–849.

11. L.V. Shilei, Z. Neng, F. Guohui, Impact of phase change wall room on indoor thermal environment in winter, Energy and Buildings 38 (2006) 18–24.

12. D. Feldman, D. Banu, DSC analysis for the evaluation of an energy storing wallboard, Thermochimica Acta 272 (1996) 243–251.

13. T.K. Stovall, J.J. Tomlinson, What are the potential benefits of including latent storage in common wallboard? Journal of Solar Energy Engineering 117 (2008) 318–325.

14. R.J. Kedl, Conventional wallboard with latent heat storage for passive solar applications, in: Energy Conversion Engineering Conference, 1990. IECEC-90. Proceedings the 25th Society, August 12–17, 1990, 1991, pp. 222–225.

15. D. Feldman, D. Banu, D. Hawes, E. Ghanbari, Obtaining an energy storing building material by direct incorporation of an organic phase change material in gypsum wallboard, Solar Energy Materials 22 (1991) 231–242.

16. J.B. Drake, A study of the optimal transition temperature of PCM Wallboard for solar energy storage. Oak Ridge National Laboratory Report ORNL/TM-10210, September 1987, 1–14.

17. Charach, C., Zarmi, Y. & Zemel, A. (1987). Simple method for assessing the thermal performance of PCM panels. Proceedings of ISES Solar World Congress, Hamburg, 1987, 1212-1216. Qtd in: PEIPPO, K., KAURANEN, P. & LUND, P.D. (1991). A multicomponent PCM wall optimized for passive solar heating. Energy and Buildings 17, 259-270.

18. D.A. Neeper, Thermal dynamics of wallboard with latent heat storage, Solar Energy 68 (5) (2000) 393–403.

19. D.P. Bentz, R. Turpin, Potential application of phase change materials in concrete technology, Cement & Concrete Composites 29 (2007) 527–532.

20. S. Weber. Curing of high strength concrete using light-weight aggregates. http://www.iwb.uni-stuttgart.de/, 1996 (retrieved 11.12.08).

21. B. Niesing, Storing heat with wax, Fraunhofer Magazine 1 (2004) 36–37.

22. C. Castellón, M. Nogués, J. Roca, M. Medrano, L.F. Cabeza, Microencapsulated phase changing materials (PCM) for building applications, in: L. Stiles (Ed.), Proceedings of the 10th International Conference on Thermal Energy Storage, Stockton, Yersey, May 31, 2006, pp. 1–9.

23. A.K. Schindler, B.F. Mccullough, The importance of concrete temperature control during concrete pavement in hot weather conditions, in: Proceedings of the Annual Meeting of the Transport Research Board, Washington, DC, 2002, pp. 1–17.