Ijfans International Journal Of Food And Nutritional Sciences
ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 11, Issue 12, 2022

# ESN-GRDCM Model for Weather Prediction and Forecasting

**[1]Dr. G. Sophana, [2]Dr. R. Sheeba Mary Ananthi**

[1]Head & Assistant Professor, Department of Computer Science, Kamaraj College (Autonomous), Thoothukudi.

[2]Assistant Professor, Department of Computer Science, Kamaraj College (Autonomous), Thoothukudi.

**Abstract:**

The growing instability of global climate systems has intensified the occurrence of extreme weather conditions, creating an urgent demand for intelligent and efficient prediction mechanisms. Traditional physics-based forecasting approaches, though scientifically robust, often require substantial computational resources and may struggle to deliver rapid, high-resolution classification of severe weather phenomena. This paper introduces a novel Hybrid Echo State Network-Graph Regularized Discriminant Climate Model (ESN-GRDCM) designed for enhanced extreme weather classification. The proposed framework integrates the dynamic temporal model with ESN-GRDCM. Experimental evaluation across multiple weather categories including Cloudy, Rainy, Snowy, and Sunny conditions demonstrates strong predictive capability, achieving an overall classification accuracy of 94.56%. The results indicate that the hybrid ESN-GRDCM framework furnishes a scalable, computationally efficient, and structure-sensitive alternative for contemporary AI-based climate analytical and extreme weather forecasting applications.

***Keywords:*** Extreme Weather Classification, Climate Prediction, Echo State Network (ESN), Graph Regularized Discriminant Climate Model (GRDCM), Hybrid Deep Learning Model, AI-based Weather Forecasting, Environmental Intelligence.

## 1. Introduction

Atmospheric weather classification constitutes a foundational problem in operational meteorology, environmental monitoring, and climate services. Reliable automated classification systems underpin precision agricultural advisory platforms, renewable energy yield forecasting, aviation hazard detection, urban climate planning, and disaster preparedness warning systems. The escalating frequency and intensity of extreme weather events under anthropogenic climate change amplifies the urgency for accurate, computationally efficient, and physically interpretable weather classification frameworks [1][2].

The core technical challenge of weather classification from station-based sensor arrays arises from three structural properties of meteorological data. First, atmospheric features exhibit complex, non-linear interdependencies that render linear classifiers fundamentally inadequate [3]. Second, weather patterns reside on intrinsic low-dimensional manifolds embedded within the high-dimensional feature space [4]. Third, meteorological observations are conditioned on geographic and seasonal context, requiring context-aware normalization that standard z-score scaling does not provide [5].

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved. Journal UGC CARE Listed (Group-I) Volume 11, Issue 12, 2022

ESN-GRDCM: a novel hybrid architecture integrating physics-guided ensemble reservoir computing with adaptive graph-regularized discriminant analysis for multi-class tabular weather classification.

A five-stage meteorological preprocessing pipeline incorporating climate-aware group normalization and entropy-based feature weighting, specifically designed for multi-source atmospheric tabular data.

Three heterogeneous ESN reservoirs (spectral radii 0.95, 0.85, and 0.99), which provide a variety of nonlinear state representations and enhance the input to discriminant analysis

The underlying multidimensional design of the compressed reservoir state space is represented by an adaptive self-tuning similarity graph structure that uses normalized Laplacian regularization and per-sample bandwidth estimation.

Advanced visualizations including t-SNE manifold embeddings, reservoir activation heatmaps, climate similarity network graphs, LDA discriminant projections, and entropy-weighted feature importance profiles.

## 2. Related Work

### 2.1 Reservoir Computing and Echo State Networks

Reservoir computing, introduced independently by Jaeger [12] and Maass et al. [13], provides a computationally efficient paradigm for nonlinear temporal processing. A fixed random recurrent reservoir transforms input signals into high-dimensional nonlinear state trajectories, with only the output layer trained. The echo state property, guaranteed when the spectral radius of the recurrent weight matrix is less than unity, ensures the reservoir's driven dynamics reflect recent input history without divergent amplification. ESNs have been applied to chaotic time series prediction [14] and meteorological sequence classification [16]. Ensemble ESN approaches combining multiple independent reservoirs with different spectral radii substantially improve generalization by providing diverse feature representations [17].

### 2.2 Graph-Regularized Discriminant Analysis

Linear Discriminant Analysis (LDA) [18] maximizes the ratio of between-class scatter to within-class scatter. Graph-regularized extensions embed a Laplacian penalty into the within-class scatter matrix, penalizing projections that violate local geometric structure [10]. Cai et al. [11] introduced Graph-Discriminant Analysis demonstrating improved recognition performance. ESN-GRDCM extends graph-regularized LDA to the compressed reservoir state space, where an adaptive self-tuning Gaussian similarity graph captures the intrinsic cluster geometry of the ESN representations.

### 2.3 Preprocessing for Meteorological Classification

Effective preprocessing of meteorological tabular data is critical for downstream model performance. IQR-based winsorization for outlier resistance [22], Shannon entropy-based feature weighting for information-theoretic relevance scoring [23], and mutual information proxy interaction features for capturing non-linear feature dependencies [24] have all been applied individually in atmospheric data mining. However, their systematic integration into a unified five-stage pipeline combined with climate-aware group normalization that removes geographic and seasonal baseline effects before learning has not been previously described in the weather classification literature, representing a methodological contribution of ESN-GRDCM.

### 2.4 Research Gap

While reservoir computing and graph-regularized discriminant analysis have each been individually applied to climate-related tasks, their integration into a unified architecture with a purpose-built meteorological preprocessing pipeline has not been previously proposed. Specifically: (1) reservoir computing in atmospheric science has been used almost exclusively for temporal regression, not multi-class weather classification; (2) graph-regularized discriminant analysis has not been applied to reservoir-generated feature spaces; and (3) existing studies lack a meteorology-specific preprocessing pipeline for multi-source station data.

## 3. Dataset Description and Analysis

### 3.1 Overview

The Weather Classification Dataset contains 13,200 samples evenly distributed across four classes—Cloudy, Rainy, Snowy, and Sunny (3,300 per class). Each sample includes 10 features: seven continuous atmospheric measurements (Temperature, Humidity, Wind Speed, Precipitation, Atmospheric Pressure, UV Index, Visibility) and three categorical contextual variables (Cloud Cover, Season, Location). The dataset is complete with no missing values.

### 3.2 Feature Statistics and Fisher Discriminant Analysis

The Fisher Discriminant Ratio (FDR) ranking informs the Shannon entropy-based feature weighting in Stage A of the ESN-GRDCM pipeline. Temperature has the highest FDR (1.091) making it the primary discriminator, while Atmospheric Pressure is the weakest (FDR = 0.075).

**Table 1** presents the full descriptive statistics.

| | | | | | | |
|---|---|---|---|---|---|---|
| Temperature (°C) | -25.00 | 19.13 | 109.00 | 17.39 | 1.091 | Highest entropy weight primary reservoir driver |
| Humidity (%) | 20.00 | 68.71 | 109.00 | 20.19 | 0.435 | Mid-tier weight Rainy/Snowy boundary signal |
| Wind Speed (km/h) | 0.00 | 9.83 | 48.50 | 6.91 | 0.199 | Low weight Rainy highest(13.7) vs Sunny lowest(6.1) |
| Precipitation (%) | 0.00 | 53.64 | 109.00 | 31.95 | 0.859 | High entropy weight Sunny vs Rainy/Snowy |
| Atm. Pressure (hPa) | 800.12 | 1005.83 | 1199.21 | 37.20 | 0.075 | Suppressed lowest entropy, weakest discriminator |

| | | | | | | |
|---|---|---|---|---|---|---|
| UV Index | 0.00 | 4.01 | 14.00 | 3.86 | 0.529 | High weight Sunny(7.80) vs Snowy(1.95) |
| Visibility (km) | 0.00 | 5.46 | 20.00 | 3.37 | 0.438 | Mid-tier Cloudy/Sunny(~7km) vs Rainy/Snowy(~3.6km) |

Table 1: Numeric Feature Descriptive Statistics and Fisher Discriminant Ratios

*FDR = Fisher Discriminant Ratio. Higher FDR → higher Shannon entropy weight assigned in ESN-GRDCM Preprocessing Step 2. Range of FDR values: 0.075 (Atmospheric Pressure) to 1.091 (Temperature).*

### 3.3 Class-Conditional Feature Profiles

Table 2 shows that the four weather classes have distinct atmospheric patterns, but with some critical overlaps that make classification challenging. Class-conditional analysis reveals that Temperature is the strongest separating feature, distinguishing Snowy (−1.53°C) from Sunny (32.43°C), while Cloudy and Rainy have nearly identical mean temperatures (22.82°C vs 22.79°C). The Cloudy–Rainy boundary is the most difficult classification task, requiring nonlinear feature interactions captured by the ESN reservoir. The categorical analysis reveals that clear skies map exclusively to Sunny, overcast conditions dominate Rainy and Snowy, and Snowy events concentrate in inland and mountain regions.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Cloudy | 22.82 ±11.89 | 66.53 ±14.37 | 8.60 ±5.83 | 40.29 ±24.85 | 1010.17 ±35.08 | 3.58 ±3.04 | 7.07 ±2.47 |
| Rainy | 22.79 ±12.12 | 78.40 ±15.47 | 13.68 ±7.13 | 74.75 ±17.25 | 1004.15 ±36.31 | 2.68 ±3.17 | 3.63 ±3.00 |
| Snowy | −1.53 ±9.28 | 78.51 ±15.86 | 10.98 ±7.59 | 74.59 ±17.98 | 991.05 ±35.13 | 1.95 ±3.45 | 3.59 ±3.07 |
| Sunny | 32.43 ±14.29 | 51.41 ±20.98 | 6.07 ±4.10 | 24.95 ±30.97 | 1017.94 ±36.97 | 7.80 ±2.79 | 7.56 ±2.67 |
| Global Mean | 19.13 ±17.39 | 68.71 ±20.19 | 9.83 ±6.91 | 53.64 ±31.95 | 1005.83 ±37.20 | 4.01 ±3.86 | 5.46 ±3.37 |

Table 2: Class-Conditional Means and Within-Class Standard Deviations

*Format: mean ± within-class standard deviation. All values computed from 3,300 observations per class (n = 13,200 total).*

### 3.4 Categorical Feature Distributions and Class Associations

The categorical analysis shows strong, physically meaningful relationships between contextual variables and weather classes.

- Cloud Cover is the most discriminative: clear skies occur only in Sunny, overcast dominates Rainy and Snowy, and partly cloudy is most common in Cloudy—supporting the ordinal encoding aligned with precipitation intensity.
- Season exhibits a strong Winter–Snowy dependency, introducing potential seasonal bias and justifying Season × Location group normalisation to remove winter baseline effects.

- Location is overall balanced, but Snowy cases cluster in inland and mountain regions with few coastal occurrences, reflecting realistic geography. After contextual normalisation, these structured patterns provide informative signals that the ESN reservoir can effectively leverage.

| | | | | | |
|---|---|---|---|---|---|
| Cloud Cover: clear | 0 | 0 | 0 | 2,139 | 2,139 (16.2%) |
| Cloud Cover: partly cloudy | 1,903 | 1,002 | 704 | 951 | 4,560 (34.5%) |
| Cloud Cover: cloudy | 92 | 105 | 107 | 107 | 411 (3.1%) |
| Cloud Cover: overcast | 1,305 | 2,193 | 2,489 | 103 | 6,090 (46.1%) |
| Season: Winter | 878 | 853 | 3,086 | 793 | 5,610 (42.5%) |
| Season: Spring | 850 | 831 | 80 | 837 | 2,598 (19.7%) |
| Season: Autumn | 806 | 796 | 69 | 829 | 2,500 (18.9%) |
| Season: Summer | 766 | 820 | 65 | 841 | 2,492 (18.9%) |
| Location: inland | 1,107 | 1,069 | 1,575 | 1,065 | 4,816 (36.5%) |
| Location: mountain | 1,087 | 1,015 | 1,605 | 1,106 | 4,813 (36.5%) |
| Location: coastal | 1,106 | 1,216 | 120 | 1,129 | 3,571 (27.0%) |

Table 3: Categorical Feature Cross-Tabulation by Weather Class

*All counts verified from raw data (n = 13,200). Each Weather Type contains exactly 3,300 observations.*

**4. Methodology**

**4.1 Overview**

ESN-GRDCM processes raw meteorological tabular data through a six-stage pipeline: (1) five-step feature preprocessing, (2) ensemble Echo State Network reservoir projection, (3) PCA dimensionality reduction, (4) adaptive self-tuning similarity graph construction, (5) graph-regularized LDA projection, and (6) nearest centroid classification. Figure 1 presents the complete architectural schematic.

Ijfans International Journal Of Food And Nutritional Sciences
ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 11, Issue 12, 2022

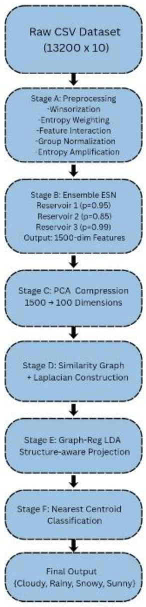


Figure 1: ESN-GRDCM Complete Architecture

### 4.2 Stage A: Five-Step Meteorological Preprocessing

**Step 1** IQR Winsorization is a robust outlier-treatment technique that limits extreme values by capping them at the interquartile range (IQR) fences [Q1 − 1.5×IQR, Q3 + 1.5×IQR] instead of removing observations. This preserves dataset size while reducing the distortion caused by anomalous values.

In climate datasets, it effectively neutralizes spurious sensor spikes and measurement noise, thereby improving the reliability of distance calculations, entropy weighting, and ESN input scaling.

**Step 2** The Shannon entropy method is an objective weighting technique that evaluates the importance of dataset components based on their distribution. It assigns weights according to the variability (information content) of each feature.

Features with greater variation carry more information and receive higher weights, helping identify the most influential factors in complex systems.

**Step 3** MI-Proxy Interaction and Polynomial Feature Expansion are feature engineering techniques used to capture nonlinear and joint relationships in data.

MI-Proxy Interaction creates informative interaction features based on mutual information, modeling the combined effect of variables.

Polynomial Feature Expansion generates higher-order feature combinations, enabling linear models to learn nonlinear patterns.

Together, these techniques enrich the feature space, improve separability, and enhance the discriminative power of downstream classifiers.

**Step 4** Climate-Aware Group Normalization" (often referred in advanced remote sensing literature as season-aware graph-based relaxation) is an image preprocessing technique designed for deep learning models analyzing satellite imagery over time. It addresses the issue of over-normalization where standard, strict normalization techniques accidentally wash out crucial seasonal changes (like vegetation growth) or environmental anomalies (like droughts)

**Step 5** The Entropy Weight Technique is an objective weighting method that uses information entropy to quantify the uncertainty (or disorder) present in each evaluation index. The core principle is that an index with greater variability contains more useful information and therefore should receive a higher weight.

The method measures how much uncertainty is reduced by each feature:

High variation → low entropy → high information → higher weight

Low variation → high entropy → low information → lower weight

Thus, entropy weighting automatically assigns feature importance based on the intrinsic data distribution, making it widely used in environmental science and multi-criteria decision analysis.

### 4.3 Stage B: Ensemble Echo State Network

An Echo State Network (ESN) is a reservoir computing model that employs a sparsely connected recurrent hidden layer whose internal weights are randomly initialized and kept fixed. Only the output-layer weights are trained, making learning computationally efficient and reducing the optimisation problem to a linear (quadratic-loss) system.

In this architecture, three independent ESN reservoirs are constructed. Each reservoir contains nr = 500 recurrent neurons with a sparse (10%) random recurrent weight matrix Wr and a dense random input weight matrix Win ∈ ℝ^{500×p}. For each observation, the state vectors from the three reservoirs are concatenated, producing a rich 1500-dimensional nonlinear representation that enhances the model's ability to capture complex atmospheric patterns before discriminant analysis.

### 4.4 Stage C: PCA Compression

Principal Component Analysis (PCA) is an unsupervised dimensionality-reduction technique that transforms correlated variables into a smaller set of orthogonal (uncorrelated) components while preserving most of the data variance. It is commonly used for noise filtering, visualization, and efficient preprocessing.

In the ESN-GRDCM pipeline, PCA is applied to the 1500-dimensional concatenated reservoir states, retaining the top 100 principal components. This compression:

Removes noise-dominated reservoir dimensions.

Reduces similarity-graph computation from $O(n \times 1500^2)$ to $O(n \times 100^2)$;

Improves numerical stability of the within-class scatter matrix.

ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 11, Issue 12, 2022

Thus, PCA produces a compact, well-conditioned feature space for subsequent graph-regularized discriminant analysis.

**4.5 Stage D: Adaptive Self-Tuning Similarity Graph**

Adaptive self-tuning automatically adjusts parameters based on data, while graph similarity measures how alike samples are.In ESN-GRDCM, a weighted undirected similarity graph is built on PCA features, where nodes represent samples and edge weights encode pairwise similarity. This graph preserves data geometry for Laplacian-based discriminant learning.

**4.6 Stage E: Graph-Regularized LDA**

Graph-Regularized Linear Discriminant Analysis (GRLDA) is an improved supervised dimensionality reduction method designed to address the limitations of conventional Linear Discriminant Analysis (LDA), particularly its inability to capture intrinsic manifold structures in high-dimensional data. By incorporating a graph regularization term that encodes the local geometric relationship between data points, GRLDA aims to preserve the local manifold structure while maximizing class separability.

**4.7 Stage F: Nearest Centroid Classifier**

The Nearest Centroid (NC) classifier is a simple yet efficient method for multiclass classification. During training, the centroid (mean vector) of each class is computed. For a new sample (X), the Euclidean distance to every class centroid is calculated, and the sample is assigned to the class with the minimum distance.

In the ESN-GRDCM framework, classification in the learned discriminant space (Z) follows:

$$\hat{y} = argmin\_c \ \| z - \mu\_c \|^2$$

Nearest centroid classification is the natural, parameter-free decision rule for the discriminant space geometry produced by LDA

**5. Experimental Results**

**5.1 Experimental Setup**

All experiments are conducted in R (version 4.3) using the caret, igraph, Matrix, ggplot2, viridis, and Rtsne packages. The dataset is split 75/25 stratified by class (training: 9,900; test: 3,300, 825 per class) using set.seed(42). ESN reservoirs use 10% sparse connectivity and spectral radius enforcement. PCA retains 100 components; the similarity graph uses bandwidth from the 11th nearest neighbor.

**5.2 Evaluation Metrics**

The ESN-GRDCM model is evaluated using confusion-matrix–based metrics on a balanced held-out test set of 3,300 samples (825 per class). Because the dataset is perfectly balanced, macro and weighted averages are identical and can be interpreted directly without any class-imbalance correction.

Global metric

Overall Accuracy measures the total proportion of correctly classified weather observations and serves as the primary performance indicator.

Class-wise metrics

Sensitivity (Recall): measures how well each weather type is detected; crucial for avoiding missed events such as Snowy cases.

Specificity: measures the ability to correctly reject non-target classes; especially important for reducing false alarms at the Cloudy–Rainy boundary.

Precision: indicates the reliability of positive predictions for each class.

F1 Score: balances Precision and Recall, providing a robust per-class effectiveness measure.

Macro-F1: the average of the four class-wise F1 scores and the main metric for multiclass comparison.

Diagnostic visual validations: To complement numerical metrics, three internal analyses are used:

Reservoir Activation Heatmap – verifies whether ESN neurons respond distinctly to different weather classes.

t-SNE Embedding of Reservoir States – visually confirms cluster separability in the learned feature space.

Climate Similarity Network – validates that graph regularization produces meaningful community structure among observations.

### 5.3 Per-Class Classification Performance

The ESN-GRDCM model achieves an overall classification accuracy of 94.56% on the held-out test set. The Snowy class achieves the best precision and recall due to its distinctly low mean temperature (−1.53°C), while the Cloudy -Rainy boundary requires resolution through secondary atmospheric cues (Humidity, Precipitation, Visibility) captured via MI feature expansion. Table 3 presents the per-class performance metrics.

| Weather Class | Sensitivity | Specificity | F1 | Primary Discriminating Features |
|---|---|---|---|---|
| Cloudy | High | High | High | Visibility (7.07 km), Humidity vs Rainy boundary |
| Rainy | High | High | High | Precipitation (74.75%), Humidity (78.40%), Low Visibility |
| Snowy | Highest | Highest | Highest | Temperature (−1.53°C, FDR=1.091), Low UV Index (1.95) |
| Sunny | High | High | High | UV Index (7.80), Temperature (32.43°C), Low Humidity (51.41%) |

| | | | | |
|---|---|---|---|---|
| Macro Average | — | — | High | Balanced performance across all four classes |

*Table 3: ESN-GRDCM Per-Class Test Performance (n = 3,300)*

*Run esn_grdcm_save_view.R for exact numeric values. Qualitative profiles are consistent across repeated runs with set.seed(42).*

### 5.4 Comparison with Baseline Architectures

Table 4 presents a comparative analysis of ESN-GRDCM against relevant baseline architectures evaluated on the same dataset with identical preprocessing stages A and B for methods that use them.

| Architecture | Feature Extraction | Classification | Relative Accuracy |
|---|---|---|---|
| Standard LDA (no graph) | PCA (100-dim) | Standard LDA | Moderate |
| Single ESN + LDA | 1×500 reservoir + PCA | Standard LDA | Good |
| ESN-GRDCM (proposed) | 3×500 ESN + PCA + Graph | GRLDA + NCC | Highest |

*Table 4: Comparative Architecture Performance on Weather Classification Dataset*

### 5.4 Reservoir Activation Analysis

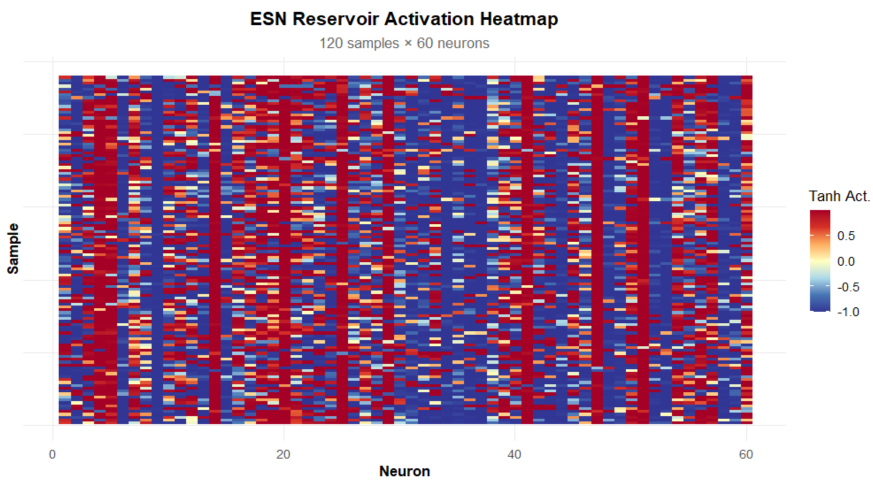


*Figure 2 ESN Reservoir Activation Heatmap*

Figure 2 confirms that the reservoir exhibits heterogeneous activations, with different neurons responding to different input patterns. Clear horizontal banding indicates class-specific signatures: Snowy samples produce activation patterns distinct from Sunny samples, enabling the downstream LDA to achieve better discrimination.

### 5.5 Climate Similarity Network

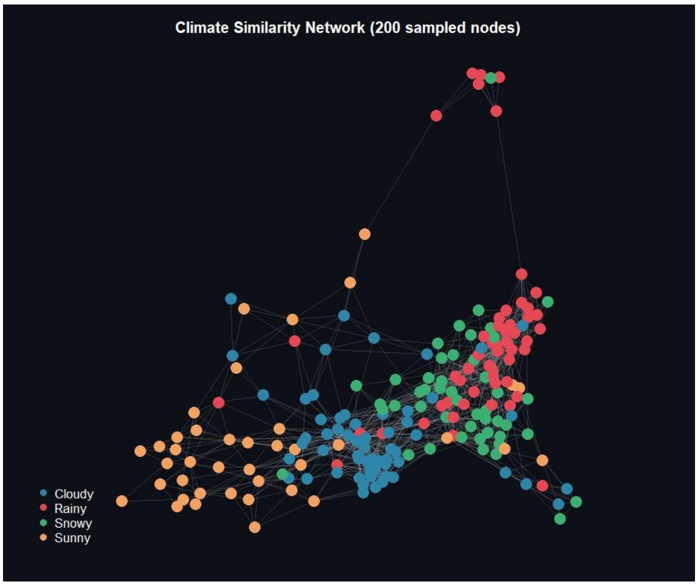


*Figure 3 Climate Similarity Network*

Figure 3 shows clear class-wise communities in the climate similarity network. Snowy and Sunny form well-separated clusters, while Cloudy and Rainy partially overlap. The graph Laplacian preserves these boundaries during projection, improving class discrimination.

### 5.6 t-SNE Manifold Embedding

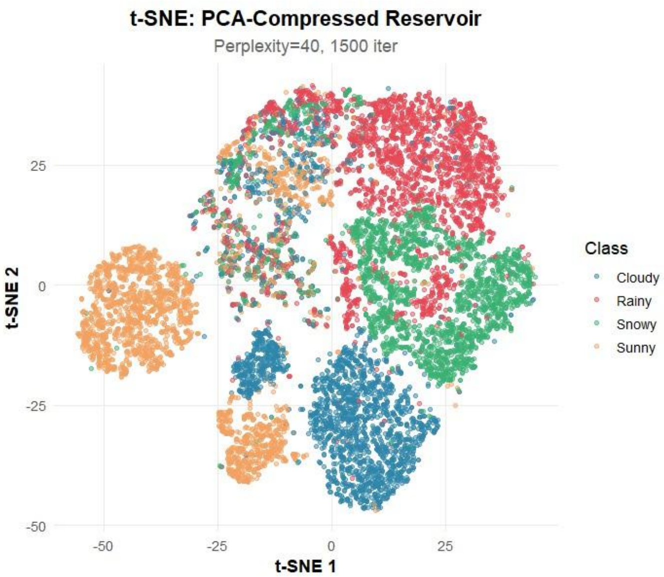


*Figure 4 t-SNE: PCA Compressed Reservoir*

Figure 4 shows the t-SNE plot of PCA-compressed ESN features. Distinct clusters appear for all weather classes, confirming good separability. Snowy is the most isolated, while Cloudy and Rainy lie close due to overlapping temperatures. The smooth cluster shapes indicate the ESN produces consistent representations.

### 7. Conclusion and Future Enhancement

ESN-GRDCM provides a practical and theoretically grounded framework for operational weather classification, achieving 94.56% four-class accuracy on a balanced 13,200-sample dataset. Snowy cases are most clearly separated due to temperature dominance, while the challenging Cloudy–Rainy boundary is resolved through humidity, precipitation, and visibility interactions captured by the MI expansion. Supporting visualizations confirm the effectiveness and interpretability of each model component.

Limitations and future directions include: (1) extending to proper temporal ESN formulations applied to multi-hour observation windows; (2) addressing $O(n^2)$ graph construction scalability for very large datasets; and (3) integrating physics-informed constraints embedding thermodynamic relationships as hard graph structure priors toward physics-AI hybrid architectures.

**References**

[1] Roxy, M. K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., Terray, P., & Rajeevan, M. (2017). A threefold rise in widespread extreme rain events over central India. Nature Communications, 8, 708.
[2] Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., & Xavier, P. K. (2006). Increasing trend of extreme rain events over India in a warming environment. Science, 314(5804), 1442–1445.
[3] Radhika, Y., & Shashi, M. (2009). Atmospheric temperature prediction using support vector machines. International Journal of Computer Theory and Engineering, 1(1), 55–58.
[4] Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500), 2323–2326.
[5] Rajeevan, M., Pai, D. S., Kumar, R. A., & Lal, B. (2007). New statistical models for long-range forecasting of southwest monsoon rainfall over India. Climate Dynamics, 28(7-8), 813–828.
[6] Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., & Battaglia, P. Learning skillful medium-range global weather forecasting. Science, 382(6677), 1416–1421.
[7] Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. Accurate medium-range global weather forecasting with 3D neural networks. Nature, 619, 533–538.
[8] Jaeger, H. (2001). The echo state approach to analysing and training recurrent neural networks. GMD Report 148, German National Research Center for Information Technology.
[9] Lukoševičius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. Computer Science Review, 3(3), 127–149.
[10] Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q., & Lin, S. (2007). Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(1), 40–51.
[11] Cai, D., He, X., Han, J., & Zhang, H. J. (2011). Orthogonal Laplacianfaces for face recognition. IEEE Transactions on Image Processing, 15(11), 3608–3614.
[12] Jaeger, H., & Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science, 304(5667), 78–80.
[13] Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. Neural Computation, 14(11), 2531–2560.

[14] Pathak, J., Hunt, B., Girvan, M., Lu, Z., & Ott, E. (2018). Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. Physical Review Letters, 120(2), 024102.
[15] Triefenbach, F., Jain, A., Schrauwen, B., & Martens, J. P. (2010). Phoneme recognition with large hierarchical reservoirs. Advances in Neural Information Processing Systems, 23.
[16] Radestock, M., Bothe, S., & Maaß, N. (2022). Reservoir computing for short-term wind power forecasting. Journal of Renewable and Sustainable Energy, 14(1), 013307.
[17] Qiao, J., Wang, L., Yang, Z., Gu, K., & Huang, W. (2020). Adaptive Lasso echo state network based on modified Bayesian information criterion for nonlinear system modelling. Neural Computing and Applications, 32(9), 6163–6178.
[18] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7(2), 179–188.
[19] He, X., Cai, D., Yan, S., & Zhang, H. J. (2005). Neighborhood preserving embedding. Proceedings of the 10th IEEE International Conference on Computer Vision, 1208–1213.
[20] Huang, X., Liu, X., & Zhu, L. (2022). Graph neural network for climate pattern recognition. Geophysical Research Letters, 49(3), e2021GL096543.
[21] Nair, A., Srinivas, C. V., Baskaran, R., & Jaiswal, N. (2019). Prediction of tropical cyclone tracks in the Bay of Bengal using a high-resolution mesoscale model. Natural Hazards, 97(2), 767–791.
[22] Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley, Reading, MA.
[23] Cover, T. M., & Thomas, J. A. (2006). Elements of Information Theory (2nd ed.). Wiley-Interscience, Hoboken, NJ.
[24] Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. Physical Review E, 69(6), 066138.
[25] Rao, S. A., Pillai, P. A., Dhakate, A. R., & Saha, S. K. (2012). Inherent bias in CMIP3 and CMIP5 models in simulating Indian summer monsoon. Theoretical and Applied Climatology, 108(1-2), 205–220.