GNN Fusion of Voronoi Spatial Graphs and City–Year Temporal Graphs for Climate Analysis

Alex Romanova
Independent Researcher
McLean, VA, USA
sparkling.dataocean@gmail.com

Abstract

We present a two-stream graph framework for climate similarity that fuses geography with long-term dynamics. A globe-spanning Voronoi network links cities whose cells share a boundary, while per-city temporal graphs encode decades of daily temperatures in 1000 cities over 40 years. We learn (i) temporal embeddings via a GNN graph-classification model on city–year graphs and (ii) spatial embeddings via a GNN link-prediction model on the Voronoi backbone, using either raw climatology vectors or the learned temporal embeddings as inputs. Treating cosine similarity as edge weights (using 1-cosine) enables graph-mining views: closeness maps highlight dense climate belts, and betweenness maps surface long-range "bridges" connecting distant regions. The fused approach uncovers patterns that simple averages miss, including nearby cities with low similarity (microclimates, urban form, or data aliasing) and far-apart cities with high similarity (shared seasonal regimes/latitude bands). We also incorporate the Delaunay triangulation - the dual of Voronoi - to provide a geometrically well-posed neighbor network that stabilizes these patterns. The method is scalable and reproducible, and the same template - spatial adjacency + temporal history + GNN fusion - extends beyond temperature to additional variables and to urban and infrastructure applications.

Keywords

graph neural networks, spatiotemporal modeling, climate analysis, Voronoi tessellation, Delaunay triangulation

1 Introduction

Understanding global climate patterns is critical to the climate–change challenge. In this study, we explore a graph-based framework that integrates geographic layout with long-term temporal behavior.

As a data source, we use climate records for 1,000 of the world's most populated cities with 40 years of daily temperatures. This dataset (Kaggle [7]) provides geolocations and multi-decade time series, allowing us to combine spatial and temporal perspectives.

Our spatial backbone is a Voronoi graph: from city coordinates, each city receives a Voronoi cell (the region closer to that city than to any other), and two cities are connected when their cells share a border—an interpretable, globally consistent notion of proximity. Alongside Voronoi, we also construct the Delaunay triangulation over the same points. Delaunay provides a complementary, dual view of neighborhood structure and enables triangle-based analyses; we use it as a robustness check to ensure results are not tied to a single choice of spatial adjacency.

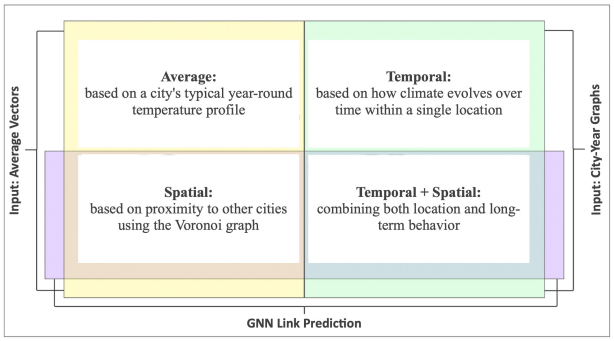

Figure 1: Node feature types for climate similarity.

For temporal behavior, each city is represented by a graph whose nodes are city–year pairs with daily-temperature profiles as features. Years are linked when their profiles exceed a cosine-similarity threshold. We add a virtual node so that each city graph forms a single connected component.

To analyze climate across space and time, we use basic vectors and pre-final vectors from Graph Neural Network (GNN) models. Figure 1 illustrates four representations used throughout the paper:

- *Average* — climatology vectors (365-day averages) per city;
- *Temporal* — embedded city graphs: pre-final vectors from a GNN graph classification model on per-city year graphs;
- *Spatial* — embedded Voronoi nodes: pre-final vectors from a GNN link-prediction model on the Voronoi graph with average vectors as inputs;
- *Spatial+Temporal* — re-embedded nodes: pre-final vectors from a GNN link-prediction model on the Voronoi graph using temporal embeddings as inputs.

We previously introduced the use of pre-final vectors from a GNN graph classification model on city temporal graphs [17] and applied linear-algebra analyses to those outputs.

In this study we contribute:

- Construction of a globe-spanning Voronoi spatial graph *and* its Delaunay triangulation as complementary spatial backbones;
- Comparisons across input climatology vectors, output city-graph embeddings, and spatial node embeddings from link prediction;
- Graph-mining analyses on induced graphs from each vector type, highlighting agreements and differences across spatial and temporal representations.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2025, Ljubljana, Slovenia
© 2025 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2025.skui.1600

2 Related Work

In 2012, two milestones reshaped AI: AlexNet's convolutional neural network set a new benchmark in large-scale image classification, far surpassing prior methods [9, 12], and Google's Knowledge Graph operationalized entity–relationship understanding at web-scale, transforming data integration, search, and management [15].

These lines of work initially evolved in parallel—CNNs excelled on grid-structured data, while graph methods targeted relational structure. The emergence of graph neural networks (GNNs) in the late 2010s bridged this gap by combining deep learning with graph computation to model complex dependencies [2]. Despite the rise of large language models (LLMs) since 2022, GNNs remain essential for tasks grounded in explicitly graph-structured data.

GNNs are now standard for classification and link prediction on graph-structured data [14, 1]. At web scale, industrial recommender systems adopt scalable inductive variants such as PinSage [20], while temporal/dynamic settings leverage trajectory-predictive embeddings like JODIE [10]. Community benchmarks have further standardized evaluation for large graph learning (e.g., OGB) [5]. In geophysics, recent studies demonstrate the effectiveness of GNNs for medium-range global weather forecasting [11], global atmospheric prediction [8], and spatiotemporal hydrology and geoscience tasks such as groundwater dynamics [19] and frost-event forecasting with attention mechanisms [13], supporting the view that graph-based inductive biases are well suited to environmental systems with strong spatial and temporal structure.

Voronoi tessellations provide natural adjacency via shared cell boundaries and have a long history in climate and global modeling [6]. Recent applications use Voronoi-induced graphs for urban risk modeling and natural hazards: Gan et al. propose a Voronoi-based spatiotemporal GCN for traffic crash prediction [3], while Razavi-Termeh et al. leverage Voronoi entropy in flood susceptibility mapping [16]. Our work synthesizes these ideas by constructing a global Voronoi-based spatial graph of cities enriched with long-term temperature signals and combining it with per-city temporal graphs encoded by GNNs.

3 Methods

3.1 Graph Construction

We construct a global spatial graph by computing a planar Voronoi diagram on Web Mercator (EPSG:3857) city coordinates; two cities are adjacent if their cells touch. The Voronoi/Delaunay is used only to define adjacency (not distances/areas), yielding a simple, interpretable map of city neighborhoods worldwide.

We evaluate four alternative node-feature sets:

(1) **365-day climatology vectors** — for each city, a 365-value day-of-year climatology averaged across all available years.
(2) **Temporal vectors** — pre-final embeddings from GNN graph-classification model on each city's year-by-year graph (years linked when their daily profiles exceed a cosine-similarity threshold).
(3) **Link-prediction vectors (from averages)** — pre-final embeddings from a GNN link-prediction model on the Voronoi graph using the 365-day climatology vectors as inputs.

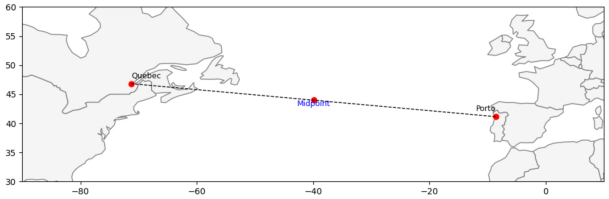

Figure 2: Voronoi edge between distant cities: Québec and Porto are neighbors because their cells meet across the Atlantic.

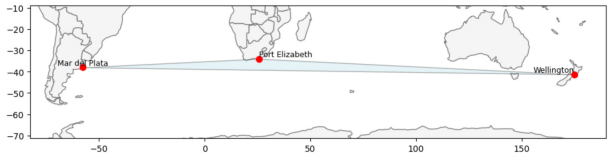

Figure 3: Largest Voronoi triangle: Wellington–Port Elizabeth–Mar del Plata illustrates long edges formed in sparsely populated regions.

(4) **Link-prediction vectors (from temporal vectors)** — the same GNN link-prediction setup, but with temporal GNN embeddings as inputs.

This design allows direct comparison of spatial, temporal, and hybrid representations within a single framework; see Figure 1.

3.2 GNN Graph Classification Model

We apply a GNN graph classification model (PyTorch Geometric) to per-city temporal graphs. Each graph has one node per year, with that year's daily-temperature profile as the node features. We add a virtual node to each graph and connect it to ensure every city graph is a single connected component. For supervision, we split cities into two equal groups by absolute latitude (closer vs. farther from the equator) and train the model to classify the graphs. We then use the pre-final vector as the city's temporal embedding for downstream analysis.

3.3 GNN Link Prediction Model

We apply a GNN link prediction model (Deep Graph Library), using the GraphSAGE aggregator [4], to the Voronoi spatial graph of cities. Unlike the GNN graph classification model, which produces one embedding per city graph, link prediction runs on the global spatial graph and refines each city's node representation using both adjacency and input features. We evaluate two node-feature variants: (i) 365-day climatology vectors (averaged across years) and (ii) temporal embeddings from the classification model. After training, we extract pre-final node embeddings as enhanced city feature vectors for downstream analysis.

Notes and code are provided on our technical blog [18].

4 Experiments

4.1 Voronoi Graph Construction

We build the spatial graph from city coordinates with a Voronoi tessellation: each city gets a cell, and two cities are linked when their cells touch. This gives a clear, globe-spanning picture of who is naturally close, without picking an arbitrary distance cutoff. Alongside this, we also use the Delaunay triangulation

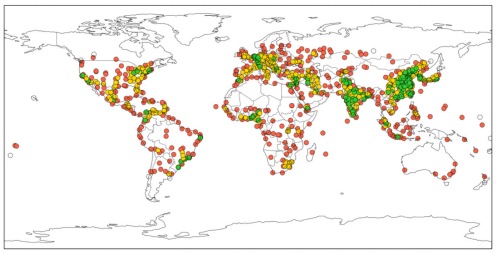

Figure 4: Voronoi area (normalized): green=low, yellow=mid, red=high.

on the same points—the dual view that connects cities exactly when their Voronoi cells meet and highlights triangle-based local structure.

Sometimes this setup links places that are far apart because there are few large cities between them. For example, Québec (Canada) and Porto (Portugal) become neighbors across the Atlantic when their cells meet (Figure 2). Larger patterns show up in the Delaunay view as well: the largest triangle—Wellington (New Zealand), Port Elizabeth (South Africa), and Mar del Plata (Argentina)—illustrates how isolated regions can still form direct connections (Figure 3).

To show spatial density, we color each city by Voronoi cell size (Figure 4). Small cells (green) mark tight clusters—for example, parts of eastern China and northern India—while large cells (red) indicate sparse areas such as interior Australia or northern Canada. Dense hubs shorten edges and raise local connectivity; sparse zones create longer links that act as bridges.

4.2 GNN Models

Across both GNNs (temporal graph classification and spatial link prediction), we use only pre-final embeddings for downstream analysis; we do not report task metrics (edge AUC/AP or classification accuracy) because our goal is weighted-path/centrality analysis on a geometric prior.

4.3 How Similar Are Distant or Nearby Cities?

This section examines climate similarity for both distant and neighboring city pairs using the four representations (*Average*, *Temporal*, *Spatial*, *Spatial+Temporal*). Tables 1 and 2 highlight highlight representative examples: one for geographically distant pairs and one for nearby pairs.

Many distant pairs show very high similarity, especially when temporal history and spatial context are both considered. For example, Wellington (New Zealand) and Mar del Plata (Argentina), though thousands of kilometers apart, score highly across all four metrics—suggesting that similar seasonal regimes and latitude can outweigh raw distance.

Nearby pairs typically agree across metrics as well. In the second table, examples such as Barranquilla–Soledad and Barcelona–Puerto La Cruz show consistently high similarity, reflecting shared local climate.

There are exceptions. New York and Brooklyn, despite being only a few kilometers apart, score low on the *Spatial* and *Spatial+Temporal* measures. This may reflect microclimates, urban effects, or dataset/aliasing issues (e.g., borough vs. city records). Such cases show that short geographic distances can mask meaningful environmental differences, underscoring the value of combining temporal and spatial modeling.

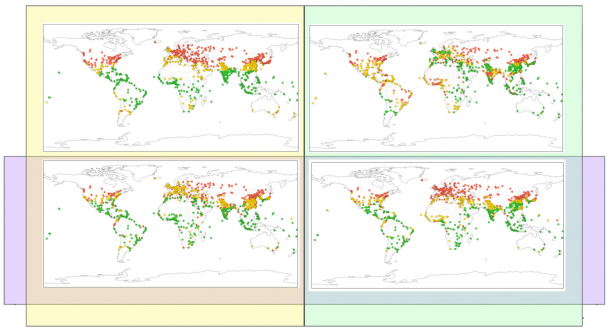

Figure 5: Closeness centrality across four vector types; red = high, yellow = mid, green = low.

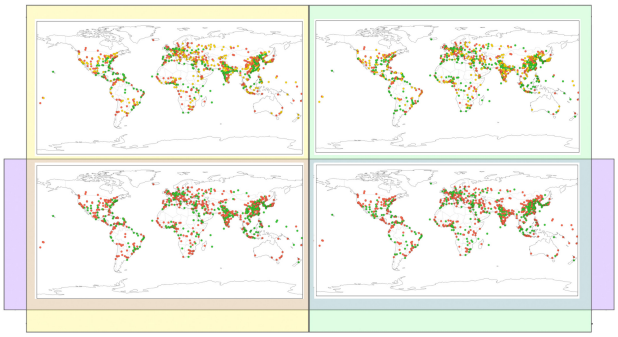

Figure 6: Betweenness centrality across four vector types; red = high, yellow = mid, green = low.

4.4 Centrality and Betweenness Patterns Across Vector Types

Throughout, *climate similarity* means *cosine similarity* between the indicated vectors; for path-based metrics we use edge weights $w = 1 - \text{cosine}$. Each set of maps uses the same spatial backbone: edges come from the Voronoi graph, where two cities are adjacent if their cells share a border. What changes across panels is the edge weight, derived from cosine similarity computed from one of four representations (*Average*, *Temporal*, *Spatial*, *Spatial+Temporal*), with vectors normalized prior to cosine. The topology stays fixed; the weights—and therefore any shortest-path–based measures—change with the chosen vectors. Smaller weights mean higher climate similarity.

In the closeness centrality maps (Figure 5), cities with high closeness are, on average, at short weighted distance from many others—i.e., they are similar to many cities. Dense climate regions such as Europe and East Asia typically stand out. Differences between panels reveal how each representation defines "similar," shifting which cities appear most central.

In the betweenness maps (Figure 6), different weightings emphasize different connectors: high-betweenness cities lie on many shortest routes. The *Spatial+Temporal* view surfaces more long-range intermediaries than *Average* (notably in Africa, South America, and the Pacific). We also observe slight polarization in *Spatial* and *Spatial+Temporal*; the reason for this requires further research.

Our centrality and betweenness maps are only a starting point, with extended graph experiments expected to uncover additional structures and recurring pathways.

Table 1: Climate similarity between distant city pairs

City 1	City 2	Distance (km)	Average	Temporal	Spatial	Spatial+Temporal
Wellington, NZ	Mar del Plata, AR	25870.97	0.9922	1.0000	1.0000	1.0000
Port Elizabeth, ZA	Wellington, NZ	16639.04	0.9982	0.9963	0.9999	1.0000
Melbourne, AU	Port Elizabeth, ZA	13299.30	0.9916	0.9958	0.9872	0.9993
Reykjavik, IS	Krasnoyarsk, RU	12911.14	0.7375	0.7482	0.9861	0.9338
Nuku'alofa, TO	Concepcion, CL	11549.31	0.9838	0.9882	0.9995	0.9997

Table 2: Climate similarity between nearby city pairs

City 1	City 2	Distance (km)	Average	Temporal	Spatial	Spatial+Temporal
Jerusalem, IL	Al Quds, PS	2.27	1.0000	1.0000	0.9998	1.0000
Barranquilla, CO	Soledad, CO	5.63	1.0000	0.9585	0.9999	0.9999
Barcelona, VE	Puerto La Cruz, VE	6.32	1.0000	1.0000	1.0000	1.0000
Khartoum, SD	Omdurman, SD	6.88	1.0000	0.8749	0.9590	0.9988
New York, US	Brooklyn, US	7.05	1.0000	0.5220	0.0857	0.0878

5 Conclusion

In conclusion, the novelty of this work is the explicit fusion of a Voronoi spatial graph with temporal GNN embeddings to reveal climate "neighborhoods" that traditional, single-view methods tend to miss. By running a GNN graph-classification model on per-city year graphs and a GNN link-prediction model on the global Voronoi backbone, we combine geography with long-term dynamics. We compare simple average-by-day climatology vectors against pre-final vectors from both GNN models and then use these vectors for downstream analysis.

This fusion surfaces informative outliers: nearby cities with low cosine similarity—consistent with microclimates, urban form, or data aliasing—and distant city pairs with high similarity, suggesting long-distance climate links. Using these vectors as edge weights enables graph-mining views: closeness maps highlight dense climate belts, while betweenness maps elevate long-range "bridges." Adding the Delaunay triangulation—the dual of the Voronoi diagram—provides a geometrically well-posed neighbor network that stabilizes these patterns.

While this study centers on climate and temperature, the dual Voronoi–Delaunay framework with GNN fusion is broadly applicable. The same geometric scaffold can analyze urban connectivity and infrastructure networks, surface social or economic linkages in dense regions, and support practical tasks like traffic management and siting of schools, parks, or grocery stores. It offers a stable way to reason about spatial relationships beyond climate. The approach is also a starting point for continued work: enrich node features, adopt spherical/geodesic tessellations, learn the graph via contrastive or metric objectives, and explore dynamic temporal GNNs with attribution, counterfactuals, and uncertainty.

References

[1] Jakub Adamczyk. 2022. Application of graph neural networks and graph descriptors for graph classification. *arXiv preprint arXiv:2211.03666*. doi:10.48550/arXiv.2211.03666.

[2] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. 2021. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*. doi:10.48550/arXiv.2104.13478.

[3] Junjie Gan, Qing Yang, Dong Zhang, Li Li, Xinyu Qu, and Bin Ran. 2024. A novel voronoi-based spatio-temporal graph convolutional network for traffic crash prediction considering geographical spatial distributions. *IEEE Transactions on Intelligent Transportation Systems*. doi:10.1109/TITS.2024.3452275.

[4] William L. Hamilton, Rex Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems (NeurIPS)*. https://arxiv.org/abs/1706.02216.

[5] Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. 2020. Open graph benchmark: datasets for machine learning on graphs. In *Advances in Neural Information Processing Systems (NeurIPS)*. https://arxiv.org/abs/2005.00687.

[6] Lili Ju, Todd Ringler, and Max Gunzburger. 2011. Voronoi tessellations and their application to climate and global modeling. In *Numerical Techniques for Global Atmospheric Models*. Lecture Notes in Computational Science and Engineering. doi:10.1007/978-3-642-11640-7_10.

[7] Kaggle Dataset. 2020. Temperature history of 1000 cities 1980 to 2020. https://www.kaggle.com/datasets/sudalairajkumar/daily-temperature-of-major-cities. (2020).

[8] Ryan Keisler. 2022. Forecasting global weather with graph neural networks. *arXiv preprint arXiv:2202.07575*. doi:10.48550/arXiv.2202.07575.

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*. doi:10.1145/3065386.

[10] Srijan Kumar, Xikun Zhang, and Jure Leskovec. 2019. Predicting dynamic embedding trajectory in temporal interaction networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. doi:10.1145/3292500.3330895.

[11] Rosalia Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, et al. 2023. Learning skillful medium-range global weather forecasting. *Science*. doi:10.1126/science.adi2336.

[12] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature*. doi:10.1038/nature14539.

[13] Hugo Lira, Luis Martí, and Nayat Sanchez-Pi. 2022. A graph neural network with spatio-temporal attention for multi-source time series data: an application to frost forecast. doi:10.3390/s22041486.

[14] Xia Liu, Jie Chen, and Qingsong Wen. 2023. A survey on graph classification and link prediction based on gnn. *arXiv preprint arXiv:2307.00865*. doi:10.48550/arXiv.2307.00865.

[15] Natasha F. Noy, Yuval Gao, Anshu Jain, Anant Narayanan, Alan Patterson, and Jamie Taylor. 2019. Industry-scale knowledge graphs: lessons and challenges. *acmqueue*. doi:10.1145/3329781.3332266.

[16] S. Vahideh Razavi-Termeh, Amir Sadeghi, Faisal Ali, Rana Abdul Naqvi, et al. 2024. Cutting-edge strategies for absence data identification in natural hazards: leveraging voronoi-entropy in flood susceptibility mapping with advanced ai techniques. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2024.132337.

[17] Alex Romanova. 2024. Utilizing pre-final vectors from GNN graph classification for enhanced climate analysis. In *Proceedings of the 21st Workshop on Mining and Learning with Graphs (MLG 2024)*. Co-located with ECML PKDD 2024.

[18] sparklingdataocean.com. [n. d.] Temporal–spatial gnn fusion for climate analytics. http://sparklingdataocean.com/2025/06/25/voronoiGNN/.

[19] Marco L. Taccari, Hua Wang, James Nuttall, Xue Chen, and Peter K. Jimack. 2024. Spatial-temporal graph neural networks for groundwater data. doi:10.1038/s41598-024-75385-2.

[20] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L. Hamilton, and Jure Leskovec. 2018. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. doi:10.1145/3219819.3219890.