Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spatial interpolation alters heat-wave detection despite small temperature errors in data-sparse regions

Fabiano Fernandes Bargos

fabianobargos@usp.br

Universidade de São Paulo

Maurício Lamano Ferreira

Universidade de São Paulo

Danúbia Caporusso Bargos

Universidade de São Paulo

Research Article

Keywords: Spatial interpolation, Temperature extremes, Heat waves, Climate indicators, Interpolation uncertainty

Posted Date: August 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10629711/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Bulletin of Atmospheric Science and Technology on November 13th, 2026. See the published version at https://doi.org/10.1007/s42865-026-00155-z.

Spatial interpolation alters heat-wave detection despite small temperature errors in data-sparse regions

Fabiano Fernandes Bargos[1*], Maurício Lamano Ferreira[1] and Danúbia Caporusso Bargos[1]

[1*]School of Engineering of Lorena, University of São Paulo, Lorena, São Paulo, Brazil .

*Corresponding author(s). E-mail(s): fabianobargos@usp.br;
Contributing authors: lamano@usp.br; danubiacbargos@usp.br;

Abstract

Climate assessments frequently rely on temperature fields generated by spatial interpolation, yet the influence of interpolation methods on derived climate indicators remains insufficiently understood. This study investigates how commonly used distance-based interpolation methods affect temperature representation and heat-wave detection in the state of São Paulo, Brazil. Hourly air-temperature observations from meteorological stations (2021–2025) were evaluated using a leave-one-out cross-validation framework. The Nearest Neighbor (NN) method and multiple configurations of Inverse Distance Weighting (IDW) were compared, and interpolated temperature series were generated for the centroids of 645 municipalities to derive daily temperature statistics and heat-wave indicators. The best-performing configuration (IDW with a 250 km search radius and distance-decay exponent of 2) achieved an RMSE of 1.97 °C and an $\boldsymbol{R^2}$ of 0.877. Although IDW adequately preserved the central tendency of maximum and mean temperatures, it systematically reduced temperature variability, particularly for minimum temperatures, attenuating cold extremes and modifying temperature distributions. These differences became substantially more pronounced for threshold-based climate indicators. Compared with NN, IDW consistently identified fewer and shorter heat-wave events, with cumulative heat-wave duration differing by up to 199 days in some municipalities over the study period. The results demonstrate that interpolation methods showing satisfactory predictive accuracy may still alter the representation of climate extremes and derived heat-wave indicators. Consequently, evaluating interpolation methods solely using conventional error metrics may be insufficient for applications involving climate-risk assessment and urban climate analysis.

Keywords: Spatial interpolation, Temperature extremes, Heat waves, Climate indicators, Interpolation uncertainty

1 Introduction

Reliable climate data are fundamental to environmental risk assessment, sustainable urban planning, climate services, and evidence-based climate

policy [1]. Air temperature is a key variable in the climate system, controlling surface energy balance, atmospheric circulation, ecosystem functioning, and a wide range of socio-economic activities. Its spatial distribution underpins applications such as climate risk assessment, agricultural planning, urban and regional planning and public health management, where reliable information on mean conditions and extremes is essential. In particular, the characterization of temperature variability is central for the study of extreme events, such as heat waves, whose frequency, duration, and intensity have increased in recent decades [2, 3]. These events pose growing threats to human health, energy infrastructure, labor capacity, and economic productivity in a warming climate [4, 5].

Despite this central role, near-surface air temperature is typically observed only at discrete locations, most commonly meteorological stations that are irregularly distributed in space. Station networks often exhibit strong spatial inhomogeneities due to topographic complexity, infrastructure constraints, and historical development, resulting in gaps in spatial representativeness [6, 7]. These limitations are particularly problematic in regions with complex terrain or sparse station density, where local-scale variability may not be adequately captured by available observations. In such contexts, interpolated temperature surfaces frequently become the empirical basis for assessing climate exposure at decision-relevant administrative scales, including municipalities and regions [8]. Addressing observational gaps is therefore essential for producing actionable climate information, but it also requires explicit attention to the uncertainty introduced during spatial estimation.

To obtain spatially continuous fields, a variety of interpolation methods are routinely applied to station data. These range from deterministic approaches, such as Inverse Distance Weighting (IDW), to geostatistical methods, such as Kriging and its variants [9, 10]. Among these, IDW is frequently adopted due to its conceptual simplicity, low computational cost, and reasonable performance under moderate spatial variability [11]. However, numerous studies have shown that interpolation performance is strongly influenced by station density, network configuration, and environmental heterogeneity [6, 12]. Thus, the apparent simplicity of distance-based interpolation may conceal spatially structured errors, especially in landscapes where altitude, coastal influence, land–sea contrasts, and mesoscale climatic processes generate sharp gradients over short distances [13].

Most evaluations of interpolation methods rely on global accuracy metrics, such as root-mean-square error (RMSE) or mean absolute error (MAE), derived from cross-validation experiments. While useful, these metrics offer only a

limited view of how interpolation alters the underlying statistical structure of temperature fields. The inherent smoothing applied by many interpolation techniques can reduce spatial variability, attenuate local extrema, and potentially introduce systematic distortions rather than purely random errors [14]. Such modifications are particularly consequential for the analysis of climate extremes. A method may perform well according to average-error metrics while still distorting the distributional tails and temporal persistence that define extreme-event indicators. This distinction is critical for environmental data science, because climate-impact assessments depend not only on accurate central tendencies, but also on the faithful representation of rare and persistent events [15].

Heat waves are typically defined through exceedance of high-temperature thresholds over multiple consecutive days, making their identification highly sensitive to small changes in temperature estimates [16]. Consequently, even modest differences between interpolation methods may substantially affect the detection, duration, and frequency of heat wave events. Yet, the impact of spatial interpolation on derived climate indicators, and on extreme temperature metrics in particular, remains insufficiently explored, especially across multiple spatial scales and under realistic data-sparse conditions typical of many regions where climate adaptation decisions are urgently needed. This represents a relevant methodological gap, because underestimating heat-wave frequency or persistence may lead to biased assessments of exposure, vulnerability, and adaptation needs.

This study addresses this gap by systematically investigating how spatial interpolation influences both temperature estimation and derived climate indicators, from individual stations to municipalities and regional aggregates. We focus on the State of São Paulo, Brazil, a densely populated and environmentally heterogeneous region characterized by coastal plains, escarpments, inland plateaus, strong altitudinal gradients, and uneven meteorological-station coverage. First, we assess interpolation performance using a leave-one-out cross-validation (LOOCV) framework and characterize spatial patterns of model skill via unsupervised learning techniques, specifically principal component analysis (PCA) and density-based clustering (DBSCAN). This analytical pipeline links conventional model evaluation to spatially explicit patterns of uncertainty and can be extended to other environmental variables and geographic regions. Second, we interpolate hourly temperature fields to municipality centroids, thereby emulating realistic data-sparse conditions in which many impact assessments are conducted. Third, we aggregate results at intermediate regional scales to examine how interpolation-induced effects propagate and potentially attenuate or amplify across space.

Beyond conventional accuracy metrics, we explicitly analyze how interpolation modifies key statistical properties of temperature time series, including variability, distribution tails, and temporal persistence. Particular emphasis is placed on the detection and characterization of heat waves, enabling a direct evaluation of how different interpolation strategies influence extreme temperature indicators. By contrasting nearest-neighbor estimates with IDW configurations, we test whether methods that improve average interpolation accuracy also preserve the temporal structure of heat extremes. This multi-scale, multi-metric approach moves beyond standard interpolation assessment by treating spatial prediction not merely as a preprocessing step, but as a source of methodological uncertainty in climate data production. Our results demonstrate that spatial interpolation is not a neutral preprocessing step but an active transformation of the climate signal. Spatial interpolation modifies the representation of temperature distributions, often reducing variability and affecting the persistence of threshold-based extreme events. These effects are spatially structured by environmental heterogeneity and station coverage, and they become more pronounced when moving from station observations to municipality-level estimates. Crucially, the same smoothing that improves conventional accuracy metrics can lead to substantial underdetection of heat-wave frequency and duration, with discrepancies reaching up to 199 heat-wave days in some municipalities when IDW is compared with a nearest-neighbor baseline [17]. By integrating model evaluation, spatial prediction, and multi-scale analysis, this work provides new evidence on the role of interpolation in climate studies and highlights the sensitivity of extreme temperature indicators to methodological choices, with direct implications for climate impact assessments, climate services, urban resilience planning, and decision-making based on interpolated datasets.

2 Study Area and Data

2.1 Study Area

The study focuses on the State of São Paulo, located in southeastern Brazil, extending approximately from 19°S to 25°S latitude and from 44°W to 53°W longitude, and covering an area of about 248,000 km^2. Elevation ranges from sea level along the Atlantic coast to more than 2,700 m in the Serra da Mantiqueira mountain range (Fig. 1).

The state exhibits pronounced physiographic heterogeneity, with a transition from coastal plains to steep escarpments and extensive inland plateaus. This complex terrain strongly influences atmospheric circulation and local climatic variability. Temperature patterns are particularly sensitive to altitude gradients and proximity to

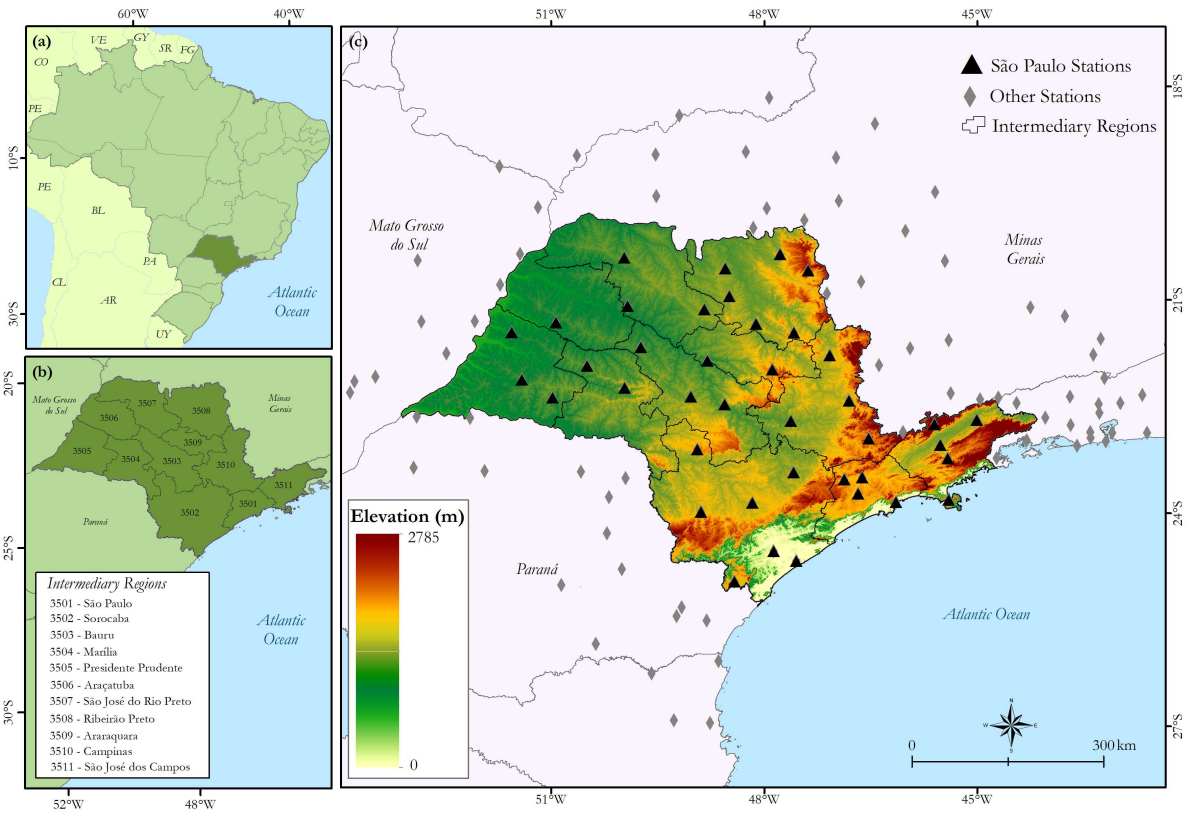


Fig. 1 Study area overview. (a) Location of São Paulo state in Brazil. (b) Intermediate Geographic Regions (IBGE) in and around São Paulo. (c) Elevation map showing meteorological stations (triangles) and intermediate region centroids (squares).

the Atlantic Ocean, resulting in sharp spatial contrasts over short distances. For example, the Vale do Paraíba, located between the Serra do Mar and the Serra da Mantiqueira, presents strong thermal variability driven by topographic and mesoscale effects.

Climatically, São Paulo is characterized by predominantly tropical and subtropical regimes, with a marked seasonal cycle associated with the South American Monsoon System. The combination of complex relief, diverse climatic conditions, and heterogeneous station distribution makes the region a suitable testbed for evaluating spatial interpolation methods and their impact on municipality-scale climate indicators.

2.2 Meteorological Data and Preprocessing

Hourly meteorological observations were obtained from the automatic weather station network operated by the Brazilian National Institute of Meteorology (INMET). Annual compressed files containing station-level CSV records were downloaded and processed. Each file includes station metadata

Table 1 Meteorological stations located in São Paulo State used in this study, including geographic coordinates, altitude, station foundation date, and annual percentages of missing hourly temperature observations from 2021 to 2025, highlighting the heterogeneous data availability across the monitoring network.

						Missing Data (%)				
ID	Station	Lat	Lon	Alt	Foundation	2021	2022	2023	2024	2025
A701	Sao Paulo - Mirante	-23.4963	-46.6201	785.6	25/07/06	0.0	0.0	2.4	0.2	0.5
A705	Bauru	-22.3581	-49.0289	636.2	30/08/01	1.2	12.5	0.0	0.0	0.1
A706	Campos do Jordao	-22.7503	-45.6039	1663.0	13/03/02	7.4	12.7	0.7	0.1	0.1
A707	Presidente Prudente	-22.1199	-51.4086	431.9	04/02/03	0.4	16.4	0.1	2.4	0.5
A708	Franca	-20.5844	-47.3825	1002.7	12/12/02	0.6	0.6	0.1	50.0	48.8
A711	Sao Carlos	-21.9804	-47.8839	859.3	04/09/06	0.1	12.8	0.1	17.9	19.2
A712	Iguape	-24.6717	-47.5459	2.7	20/07/06	52.7	24.0	0.0	0.3	0.0
A713	Sorocaba/Iperó	-23.4261	-47.5856	609.3	22/08/06	55.1	34.6	0.0	0.0	54.9
A714	Itapeva	-23.9819	-48.8858	743.3	25/07/06	7.7	11.0	0.0	0.0	11.2
A715	Sao Miguel Arcanjo	-23.8519	-48.1647	675.7	15/08/06	30.5	39.1	26.8	46.5	0.2
A716	Ourinhos	-22.9490	-49.8946	443.5	29/08/06	100.0	-	37.8	11.1	0.1
A718	Rancharia	-22.3728	-50.9747	398.8	01/09/06	21.7	32.3	0.0	0.0	0.3
A725	Avare	-23.1017	-48.9411	776.4	22/09/06	91.0	32.6	1.2	15.3	28.7
A726	Piracicaba	-22.7031	-47.6233	566.5	26/09/06	100.0	19.1	0.0	10.9	26.1
A727	Lins	-21.6661	-49.7347	460.6	20/09/06	64.9	31.7	0.0	18.0	35.1
A728	Taubate	-23.0417	-45.5208	582.3	20/12/06	84.5	4.5	0.0	27.2	0.8
A729	Votuporanga	-20.4032	-49.9660	510.4	04/12/06	1.8	-	-	-	-
A733	Jales	-20.1650	-50.5950	460.4	22/08/07	57.7	41.1	30.5	8.3	0.2
A734	Valparaiso	-21.3192	-50.9303	381.9	30/08/07	13.4	36.3	2.1	0.0	0.1
A735	Jose Bonifacio	-21.0857	-49.9204	408.4	03/09/07	90.4	6.6	37.9	87.5	27.8
A736	Ariranha	-21.1331	-48.8406	525.4	13/11/07	66.6	46.7	0.0	0.0	4.6
A737	Ibitinga	-21.8556	-48.7997	496.8	09/11/07	33.5	3.9	0.0	36.7	28.0
A738	Casa Branca	-21.7806	-47.0753	734.1	25/06/07	100.0	2.1	0.0	7.8	12.8
A739	Itapira	-22.4151	-46.8053	634.9	05/11/07	100.0	34.2	3.9	8.2	5.7
A740	Sao Luis do Paraitinga	-23.2283	-45.4169	862.3	01/11/07	82.9	15.0	16.7	24.1	33.7
A741	Barra Bonita	-22.4711	-48.5575	533.7	24/04/08	64.0	20.5	0.0	0.1	0.2
A744	Braganca Paulista	-22.9519	-46.5306	891.0	20/12/17	14.1	14.3	3.3	3.2	1.6
A746	Barra do Turvo	-24.9628	-48.4164	659.9	03/07/08	100.0	25.0	0.0	0.1	12.7
A747	Pradopolis	-21.3383	-48.1139	540.4	22/04/08	0.0	0.0	0.0	2.2	0.1
A748	Barretos	-20.5592	-48.5450	534.4	19/06/10	22.5	53.0	4.4	39.6	27.4
A753	Ituverava	-20.3597	-47.7753	610.6	17/07/08	57.6	15.2	0.1	0.2	1.3
A755	Barueri	-23.5239	-46.8694	776.5	29/03/11	0.2	6.5	4.6	8.7	1.0
A762	Dracena	-21.4577	-51.5523	720.0	03/11/16	16.5	0.3	0.0	39.8	23.1
A763	Marilia	-22.2352	-49.9651	660.0	15/05/17	0.0	0.0	1.3	0.4	0.2
A764	Bebedouro	-20.9492	-48.4900	790.0	27/10/16	51.0	18.6	0.9	-	4.7
A765	Bertioga	-23.8447	-46.1433	5.0	01/02/17	73.5	16.2	0.1	0.2	0.2
A766	Registro	-24.5331	-47.8642	35.0	09/02/17	78.3	18.4	2.0	19.8	92.5
A767	Sao Sebastiao	-23.8108	-45.4025	24.0	25/10/17	98.6	-	-	-	-
A768	Tupa	-21.9273	-50.4903	498.0	12/05/17	4.1	9.9	4.9	4.3	3.9
A769	Cachoeira Paulista	-22.6889	-45.0056	586.0	20/10/17	100.0	13.1	1.9	3.7	41.8
A770	Sao Simao	-21.4611	-47.5794	620.0	03/07/19	1.4	1.4	2.8	3.7	2.1
A771	Sao Paulo - Interlagos	-23.7245	-46.6775	771.0	14/03/18	0.0	0.0	1.9	11.2	0.3

(geographic coordinates, altitude, and operational information) followed by hourly measurements.

In this study, air temperature observations were analyzed for the period from January 2021 to December 2025. Data availability varies substantially across stations due to operational interruptions and missing records, requiring systematic preprocessing prior to analysis.

To improve spatial consistency and reduce interpolation artifacts near domain boundaries, a buffer region with a radius of 250 km around São Paulo State was included. This allows interpolation procedures to incorporate neighboring stations located outside the state, increasing spatial support and reducing edge effects.

The final dataset includes meteorological stations located both within São Paulo State and in surrounding regions, providing a more robust basis for spatial interpolation under heterogeneous and incomplete observational coverage. In total, the dataset comprises 160 stations, with 42 stations located within São Paulo State. Table 1 summarizes the percentage of missing temperature observations for the São Paulo stations during the 2021–2025 period and reveals substantial differences in data completeness among stations and years. While several stations present relatively complete records throughout the study period (e.g., São Paulo–Mirante, Marília, and São Simão), others exhibit prolonged interruptions and high percentages of missing data. In some cases, missing data exceed 50% in specific years, and a few stations contain periods with nearly complete absence of observations.

These gaps highlight the operational limitations and temporal instability commonly observed in meteorological monitoring networks. Importantly, missing records are not spatially or temporally uniform, meaning that the set of stations available for interpolation changes continuously over time. This characteristic directly affects spatial interpolation, especially for hourly estimation, since neighboring observations available at one time step may not be available at the next.

3 Methodology

3.1 Data Processing and Feature Construction

The methodological framework follows a data-driven pipeline designed to address spatial temperature estimation under sparse and incomplete observations (Figure 2). The workflow integrates data preprocessing, model evaluation, spatial prediction, and temporal aggregation into a unified framework for climate analysis.

The pipeline consists of four main steps: (i) preprocessing and feature construction from hourly station data, (ii) evaluation of interpolation performance using cross-validation, (iii) spatial prediction at the municipality level, and (iv) temporal aggregation for the derivation of climate

indicators. Additional analytical steps, including unsupervised learning, are used to characterize model performance patterns.

Hourly temperature observations from meteorological stations were first quality-controlled and organized into consistent time series. From these data, spatial and performance-related features were constructed, including geographic coordinates, altitude, and distances between stations.

Interpolation performance was assessed using a leave-one-out cross-validation (LOOCV) procedure. In this approach, each station is temporarily removed from the dataset and its temperature is estimated using the remaining stations. The predicted value is then compared with the observed measurement, enabling a systematic evaluation of interpolation errors across the study domain.

For benchmarking purposes, the Nearest Neighbor (NN) method was adopted as a baseline, and multiple configurations of the Inverse Distance Weighting (IDW) method were tested. The evaluated configurations considered radii of 150, 200, and 250 km and distance-decay exponents $\alpha = 2$ and $\alpha = 3$.

The search radii were defined considering the spatial distribution of stations and the temporal completeness of the hourly observational records. Since interpolation was performed at an hourly resolution prior to the derivation of daily temperature indicators, the availability of observations varied continuously in space and time due to missing data and operational interruptions in the station network.

As a consequence, the set of stations available for interpolation could differ between consecutive hours. Preliminary analyses showed that search radii smaller than 150 km frequently resulted in situations where no valid observations were available for interpolation at specific locations and times. The selected radii (150, 200, and 250 km) therefore represent a compromise between spatial representativeness and data availability, allowing temperature estimates to be generated continuously for all locations and time steps throughout the study period. For each interpolation configuration, all stations with valid observations within the specified search radius were used in the estimation. At least one valid observation was available for every interpolation instance considered in the analysis. When only a single station was available within the search radius, the interpolated value was equivalent to the observation from that station.

Because IDW relies exclusively on spatial distance, it does not explicitly account for environmental factors such as elevation or land–sea contrasts. As a consequence, its performance may vary across regions with strong spatial gradients or heterogeneous terrain.

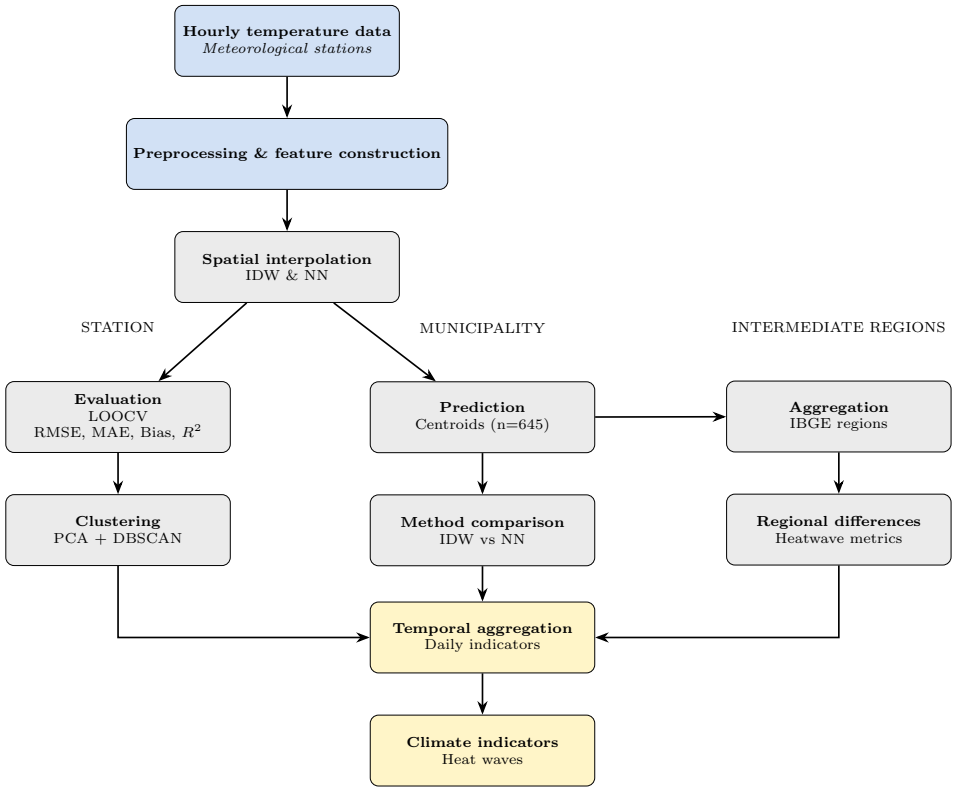


Fig. 2 Schematic representation of the data-driven workflow adopted in this study. The pipeline starts from hourly temperature observations at meteorological stations and includes preprocessing and feature construction. Two complementary branches are then developed: (i) a model evaluation framework based on leave-one-out cross-validation and unsupervised learning (PCA + DBSCAN) to characterize spatial patterns of interpolation performance, and (ii) a spatial prediction framework used to estimate temperature at municipality centroids. The interpolated hourly series are subsequently aggregated into daily indicators (mean, minimum, and maximum temperatures), enabling the analysis of climate extremes such as heat waves. The workflow highlights the interaction between model evaluation and spatial prediction, emphasizing how interpolation choices can influence the representation of temperature variability and extremes.

Finally, hourly temperature observations were interpolated to the centroids of the 645 municipalities of São Paulo state using the evaluated IDW configurations and the NN approach. The resulting municipality-level hourly series were aggregated into daily indicators, including mean, minimum, and maximum temperatures.

3.2 Spatial Interpolation Methods

Spatial temperature interpolation was performed using the Inverse Distance Weighting (IDW) method, which estimates values at unsampled locations as a distance-weighted average of nearby observations.

The interpolated value $Z(x)$ at location x is defined as

$$Z(x) = \frac{\sum_{i=1}^{n} Z_i w_i}{\sum_{i=1}^{n} w_i}, \tag{1}$$

where Z_i is the temperature observed at station i and w_i is the corresponding weight. Weights were

computed as

$$w_i = \begin{cases} d(x, x_i)^{-\alpha}, & d(x, x_i) \leq R_{\max}, \\ 0, & d(x, x_i) > R_{\max}, \end{cases} \tag{2}$$

where $d(x, x_i)$ is the distance between the interpolation point and station i, α controls the distance decay, and $R_{\max}$ defines the maximum search radius.

3.3 Cross-Validation and Performance Evaluation

Interpolation accuracy was quantified using standard performance metrics applied to cross-validation predictions. Interpolated values ($\hat{y}_t$) were compared with observed temperatures (y_t). Let T denote the number of observations.

Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2}. \tag{3}$$

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|\hat{y}_t - y_t|. \tag{4}$$

Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}. \tag{5}$$

Bias

$$\text{Bias} = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t). \tag{6}$$

These metrics provide complementary information on interpolation accuracy, variability, and systematic deviation.

3.4 Dimensionality Reduction and Unsupervised Clustering

To identify groups of stations with similar model performance under heterogeneous conditions, a combined approach based on Principal Component Analysis (PCA) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) was employed.

PCA was applied to the standardized feature set, including model performance metrics (RMSE, MAE, Bias, R^2) and geographic attributes (latitude, longitude, altitude). This transformation reduces redundancy and captures the dominant sources of variability in a lower-dimensional space.

The first three principal components were retained, as they preserve most of the variability while maintaining relevant spatial and performance-related information.

Clustering was then performed using DBSCAN, which identifies groups of points in regions of high density and labels isolated points as noise. This method is suitable for environmental datasets, where clusters may have irregular shapes and outliers may arise due to localized variability.

For a point $\mathbf{x}_i$, the ϵ-neighborhood is defined as

$$\mathcal{N}_\epsilon(\mathbf{x}_i) = \left\{ \mathbf{x}_j \in \mathbb{R}^d \;\middle|\; \|\mathbf{x}_j - \mathbf{x}_i\| \leq \epsilon \right\}, \quad (7)$$

and a point is considered a core point if

$$|\mathcal{N}_\epsilon(\mathbf{x}_i)| \geq \text{min_samples}. \quad (8)$$

This step provides a data-driven characterization of spatial patterns in interpolation performance, complementing the cross-validation analysis.

3.5 Spatial Prediction at Municipality Level

Following the evaluation stage, the selected interpolation methods were applied to estimate temperature values at municipality centroids.

This step represents a spatial inference process, extending point-based observations to a continuous spatial domain. The resulting municipality-level time series provide a consistent dataset for regional analysis and enable the propagation of interpolation uncertainties identified during cross-validation.

Additionally, differences between interpolation methods were computed at the municipality level (e.g., mean and maximum differences in daily temperature), and used as input features for subsequent clustering and regional analyses.

3.6 Intermediate Region-Level Aggregation and Evaluation

To assess the regional consistency of interpolation effects, municipality-level results were aggregated according to the Intermediate Geographic Regions defined by IBGE.

For each region, summary statistics of temperature variables and derived indicators were computed for both observed (station-based) and interpolated datasets. In addition, differences between interpolation methods (e.g., IDW and nearest neighbor) were aggregated to quantify regional-scale discrepancies.

This step enables the evaluation of how interpolation effects propagate across spatial scales, linking local estimation errors to broader regional patterns in temperature distributions and climate indicators.

3.7 Temporal Aggregation and Climate Indicators

After spatial interpolation, the hourly temperature series at municipality centroids were aggregated into daily indicators, including mean, minimum, and maximum temperatures.

This aggregation step can be interpreted as temporal feature engineering, enabling the analysis of daily variability and providing the basis for the construction of climate indicators.

3.8 Heat Wave Definition

The objective of this study is not primarily to propose a climatological definition of heat waves, but rather to evaluate how spatial interpolation methods influence the representation of threshold-based climate indicators. For this reason, a fixed temperature threshold was adopted to ensure methodological consistency across municipalities and interpolation methods, enabling direct comparison between interpolation approaches.

Heat waves were defined based on thresholds applied to daily maximum temperature series. Events were identified as sequences of consecutive days with daily maximum temperature exceeding 35°C, allowing the characterization of both event duration and persistence.

The same definition was applied consistently to all municipalities and interpolation methods. This framework allows the assessment of how different spatial interpolation approaches influence the detection of extreme temperature events.

Overall, the methodology can be interpreted as a data science pipeline integrating preprocessing, model evaluation, spatial prediction, unsupervised learning, and temporal feature engineering for climate indicator extraction.

Because heat wave identification relies on fixed temperature thresholds and consecutive-day criteria, the detected events are inherently sensitive to small variations in temperature estimates. As a result, differences introduced by spatial interpolation may affect not only temperature values, but also the frequency, duration, and persistence of detected heat wave events.

Beyond assessing how interpolation method selection influences climate indicators, the proposed framework also provides a systematic approach for identifying the most suitable interpolation strategy for heatwave quantification according to different application domains and decision-making contexts, such as public health, biodiversity conservation, urban planning, and climate adaptation studies.

4 Results

4.1 Station-Level Model Evaluation and Performance Analysis

Interpolation performance was assessed using a leave-one-out cross-validation (LOOCV) framework, with the resulting error metrics used to characterize spatial patterns in model behavior. Table 2 summarizes the performance of the evaluated interpolation configurations. Overall, all IDW-based approaches outperform the nearest-neighbor (NN) baseline, which exhibits the highest error (RMSE = 2.652 °C) and the lowest

explanatory power ($R^2 = 0.778$). The inclusion of distance-based weighting and a finite search neighborhood consistently improves interpolation accuracy across all metrics.

For a fixed decay parameter α, increasing the search radius leads to improved performance. For instance, with $\alpha = 2$, RMSE decreases from 2.016 °C ($R = 150$ km) to 1.973 °C ($R = 250$ km). A similar trend is observed for $\alpha = 3$, although with slightly smaller gains. This behavior suggests that expanding the spatial neighborhood improves representativeness by incorporating a larger set of surrounding observations.

For a given radius, configurations with $\alpha = 2$ systematically outperform those with $\alpha = 3$. Lower values of α produce a smoother decay of influence with distance, allowing a broader contribution from neighboring stations. The best overall performance is obtained for $(R, \alpha) = (250\text{ km}, 2)$, with RMSE = 1.973 °C and $R^2 = 0.877$. Bias values remain small across all configurations ($|\text{Bias}| < 0.17$ °C), indicating limited systematic deviation.

Figure 3 illustrates the spatial distribution of meteorological stations across the Intermediate Geographic Regions defined by IBGE. Station-level results are detailed in Table 3, where interpolation accuracy is reported for each INMET station, sorted by RMSE, comparing NN and the selected IDW configuration ($R = 250$ km, $\alpha = 2$).

At the station scale, IDW generally provides lower RMSE values and reduced bias compared to NN. For example, RMSE decreases from 5.56 to 4.60 in Campos do Jordão (A706) and from 5.39 to 1.79 in Bragança Paulista (A744). Similarly, Presidente Prudente (A707) shows a reduction from 3.73 to 1.88. These improvements indicate a more accurate representation of local temperature variability.

The coefficient of determination (R^2) also increases for many stations under IDW, reflecting improved explanatory capability. Stations with moderate performance under NN often exhibit noticeable gains when spatial weighting is applied.

Spatial patterns are also evident when results are grouped by Intermediate Geographic Region.

Table 2 Performance of the IDW configurations evaluated in the study. The table reports the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), coefficient of determination (R^2), and Bias for the nearest-neighbor baseline and for each combination of search radius R (km) and distance-decay exponent α. Lower RMSE and MAE and higher R^2 indicate better interpolation accuracy.

		$\alpha = 2$			$\alpha = 3$		
	NN	$R = 150$	$R = 200$	$R = 250$	$R = 150$	$R = 200$	$R = 250$
RMSE	2.652	2.016	1.976	1.973	2.071	2.038	2.027
MAE	1.909	1.479	1.454	1.454	1.514	1.492	1.485
R2	0.778	0.872	0.877	0.877	0.865	0.869	0.870
Bias	-0.090	-0.136	-0.146	-0.165	-0.127	-0.136	-0.149

Areas with higher station density, such as Campinas and São José dos Campos, tend to exhibit lower interpolation errors. In contrast, regions with sparse coverage or more complex terrain generally present higher errors, although IDW still improves performance relative to NN.

4.1.1 Clustering of Stations

The first three principal components explain 38.3%, 29.3%, and 14.4% of the total variance, respectively, resulting in a cumulative variance of 82.0%. This indicates that most of the variability in performance metrics and spatial descriptors is preserved in the reduced feature space. Although the first two components account for 67.6% of the variance, the inclusion of the third component improves cluster separability by retaining information related to elevation and localized performance differences.

DBSCAN was applied to the three-dimensional PCA space using $\epsilon = 0.8$ and min_samples $= 4$, resulting in three groups: two well-defined clusters and a set of noise points (cluster -1). Table 4 summarizes the descriptive statistics and results of the Kruskal–Wallis tests for all evaluated metrics across clusters.

Cluster -1 groups stations with markedly higher errors (MAE and RMSE), lower and more variable R^2, and stronger negative bias. These characteristics indicate systematic difficulties in model performance, likely associated with heterogeneous environmental conditions and localized effects not captured by the interpolation.

Clusters 0 and 1 exhibit overall satisfactory performance, with low errors and high R^2. Cluster 1 shows slightly better performance and lower dispersion, suggesting more stable predictive behavior. In contrast, cluster 0 presents marginally higher errors and is associated with higher mean altitudes, indicating a potential influence of topographic complexity.

The Kruskal–Wallis test indicates statistically significant differences among clusters for all evaluated metrics. These differences are primarily driven by contrasts involving cluster -1.

Error metrics (MAE and RMSE) are significantly higher in cluster -1 ($p < 0.001$), confirming its characterization as a group of poorly predicted stations. In contrast, clusters 0 and 1 do not differ significantly in terms of error magnitude.

Bias also differs among clusters ($p = 0.016$), with cluster -1 exhibiting stronger negative bias, while cluster 0 shows a slight positive bias. Differences between clusters 0 and 1 remain small.

For R^2, cluster -1 presents lower and more variable values ($p < 0.001$), whereas clusters 0 and 1 maintain consistently high performance.

Altitude differences are also significant ($p = 0.005$), particularly between clusters 0 and 1, suggesting that elevation plays a secondary but

Table 3 Station-level interpolation accuracy metrics sorted by RMSE. The table shows the ten stations with the largest errors and the ten with the smallest errors. Results compare the nearest-neighbor (NN) method with the IDW configuration $(R, \alpha) = (250\ \text{km}, 2)$. Regions correspond to IBGE Intermediate Geographic Regions.

			RMSE		MAE		R2		Bias	
Station		Region	NN	(250,2)	NN	(250,2)	NN	(250,2)	NN	(250,2)
A706	Campos do Jordao	São José dos Campos	5.56	4.60	4.87	4.37	-0.41	0.04	4.65	4.36
A744	Braganca Paulista	Campinas	5.39	1.79	4.93	1.39	-0.15	0.87	-4.88	-1.08
A718	Rancharia	Presidente Prudente	3.91	3.31	2.63	2.20	0.67	0.76	2.18	1.63
A767	Sao Sebastiao	São José dos Campos	3.87	3.42	3.43	3.16	-3.70	-2.68	-3.38	-3.16
A765	Bertioga	São Paulo	3.80	3.58	3.40	3.12	0.03	0.14	-2.91	-2.69
A707	Presidente Prudente	Presidente Prudente	3.73	1.88	2.47	1.34	0.46	0.86	-1.90	-0.83
A728	Taubate	São José dos Campos	3.69	2.80	3.01	2.50	0.50	0.71	-2.82	-2.40
A746	Barra do Turvo	Sorocaba	3.66	2.53	3.25	2.11	0.50	0.76	3.08	1.99
A769	Cachoeira Paulista	São José dos Campos	3.22	2.20	2.72	1.80	0.61	0.82	-2.41	-1.60
A753	Ituverava	Ribeirão Preto	2.93	1.90	2.40	1.56	0.77	0.90	-0.31	-0.18
		... stations omitted for brevity ...								
A770	Sao Simao	Ribeirão Preto	1.74	1.84	1.22	1.49	0.90	0.88	-0.45	-1.29
A747	Pradopolis	Ribeirão Preto	1.74	1.38	1.22	1.05	0.91	0.95	0.43	-0.16
A735	Jose Bonifacio	São José do Rio Preto	1.66	1.31	1.20	0.98	0.91	0.95	0.18	-0.22
A764	Bebdouro	Ribeirão Preto	1.59	1.66	1.12	1.25	0.93	0.93	-0.03	-0.11
A766	Registro	Sorocaba	1.56	1.77	1.11	1.45	0.90	0.88	-0.33	-1.24
A736	Ariranha	São José do Rio Preto	1.56	1.37	1.11	1.02	0.93	0.95	-0.06	-0.12
A705	Bauru	Bauru	1.52	1.16	1.13	0.86	0.92	0.95	0.34	0.18
A741	Barra Bonita	Bauru	1.46	1.42	1.09	1.11	0.93	0.93	-0.35	-0.63
A701	Sao Paulo - Mirante	São Paulo	1.42	1.32	1.02	0.98	0.91	0.92	-0.64	-0.74
A755	Barueri	São Paulo	1.42	1.09	1.02	0.81	0.92	0.95	0.62	0.19

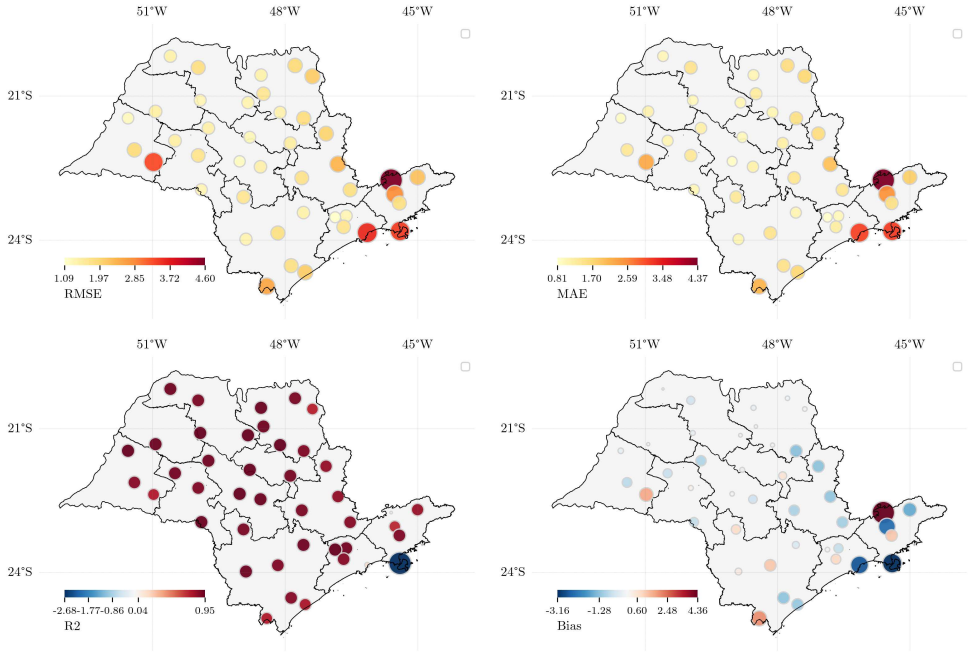


Fig. 3 Spatial distribution of RMSE, MAE, R2 and Bias obtained with the IDW configuration $(R, \alpha) = (250\ \text{km}, 2)$ for INMET meteorological stations in São Paulo state. The map highlights regional differences in interpolation accuracy, with larger errors occurring mainly in coastal and mountainous areas and lower errors across the interior plateau.

Table 4 Summary of descriptive statistics by cluster for all evaluated metrics, including Kruskal–Wallis p-values.

	Cluster	Mean	Std	Median	p-valor (KW)
MAE	-1	2.223	0.965	1.799	
	0	1.289	0.345	1.296	0.0002
	1	1.192	0.307	1.080	
RMSE	-1	2.582	0.933	2.196	
	0	1.647	0.376	1.679	0.0004
	1	1.585	0.465	1.421	
Bias	-1	-0.877	2.058	-1.243	
	0	0.301	1.005	0.401	0.0163
	1	-0.263	0.593	-0.220	
R^2	-1	0.365	1.053	0.782	
	0	0.890	0.051	0.900	0.0001
	1	0.916	0.045	0.932	
Altitude	-1	557.180	521.598	586.000	
	0	737.954	100.582	771.000	0.0048
	1	533.063	109.944	517.905	

non-negligible role in structuring model performance.

Figure 4 provides a distributional view of model performance across clusters. Cluster -1 exhibits higher median errors and a wider spread, with numerous extreme values, indicating unstable and less reliable predictions.

In contrast, clusters 0 and 1 show tighter distributions and lower median errors, reflecting more consistent model behavior. Differences between

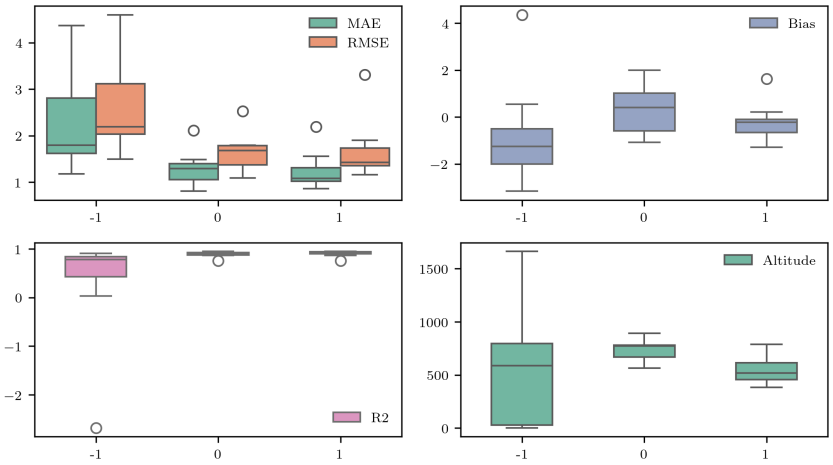


Fig. 4 Boxplots of performance metrics (MAE, RMSE, Bias, and R^2) across clusters identified by DBSCAN in the PCA-transformed space. The distributions highlight clear differences in error magnitude, variability, and model skill among clusters. Cluster -1 (outliers) exhibits higher dispersion and extreme values, while clusters 0 and 1 show more compact distributions and consistently better performance.

these clusters are subtle, although cluster 1 generally presents slightly lower variability and higher R^2 values.

The dispersion observed in cluster -1 reinforces the presence of heterogeneous conditions and supports its interpretation as a group of stations where the interpolation approach struggles to capture local variability. These results demonstrate that clustering in the reduced feature space effectively separates distinct regimes of model performance, linking statistical behavior to underlying geographic and environmental variability.

These results demonstrate that clustering in the reduced feature space effectively separates distinct regimes of model performance, linking statistical behavior to underlying geographic and environmental variability.

4.1.2 Spatial Patterns

The spatial distribution of clusters (Figure 5) reveals that poorly performing stations (cluster -1) are scattered across the region rather than concentrated in specific geographic zones. This suggests that model limitations are not driven solely by large-scale spatial gradients.

Clusters 0 and 1 display greater spatial coherence, indicating that regions with similar geographic and topographic characteristics tend to share comparable levels of model performance. In contrast, the dispersed pattern of outliers highlights the role of local-scale variability, which becomes particularly relevant under heterogeneous conditions.

Altitude ranges differ substantially among clusters. Cluster -1 exhibits the widest range,

spanning from very low to high elevations, reinforcing the role of environmental heterogeneity in model instability.

Clusters 0 and 1 show more constrained altitude intervals, with cluster 0 concentrated at higher elevations and cluster 1 at moderately lower elevations. While altitude alone does not fully explain model performance, it appears to modulate prediction difficulty when combined with local variability and station representativeness.

4.2 Municipality-Level Spatial Prediction and Evaluation

4.2.1 Municipality Clustering

To investigate whether spatial patterns observed at the station level are consistent at the municipality scale, we analyzed the differences between daily maximum temperatures obtained using the IDW configuration ($R = 250$ km, $\alpha = 2$) and the Nearest Neighbor (NN) approach.

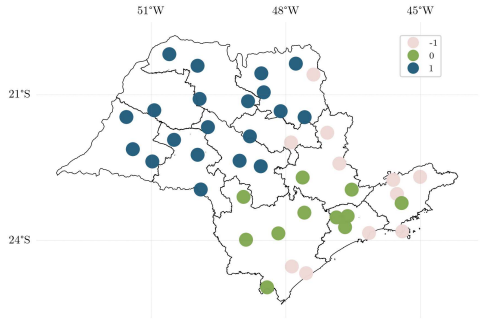


Fig. 5 Spatial distribution of clusters identified by DBSCAN on PCA-transformed. Colors indicate cluster membership: green (-1, outliers), blue (0, moderate performance), and gray (1, best performance).

For each municipality, summary statistics of these differences (mean and maximum) were computed and combined with spatial descriptors, including centroid coordinates, minimum distance to the nearest station, and altitude. All variables were standardized prior to analysis.

Principal Component Analysis (PCA) was applied to reduce dimensionality, and the first three components were retained. Clustering was subsequently performed using DBSCAN with parameters $\epsilon = 0.651$ and min_samples $= 18$.

The resulting clusters (Figure 6) closely mirror the clustering obtained for meteorological stations, indicating strong consistency between station-based and municipality-level analyses.

Municipalities classified as noise (cluster -1) are predominantly located in areas with very low or very high elevations and greater distance from observation stations. These areas are associated with larger discrepancies between interpolation methods, suggesting increased uncertainty and reduced model reliability.

Cluster 1 is mainly concentrated in the northern portion of the study area, where spatial patterns are more homogeneous and interpolation performance is more stable. The remaining clusters exhibit intermediate behavior, reflecting gradual transitions between these regimes.

This consistency suggests that interpolation errors are not randomly distributed but spatially

structured, reflecting underlying environmental heterogeneity and station coverage.

4.3 Intermediate Region-Level Evaluation

Daily temperature data were analyzed for both weather stations (observed) and municipality centroids (interpolated using IDW with $R = 250$ km and $\alpha = 2$). The variables of interest were daily maximum, minimum, and mean temperatures, grouped by intermediate region.

Table 5 presents descriptive statistics for daily maximum, mean and minimum temperatures. Interpolated municipality values reproduce station-based maxima reasonably well across all regions, with differences in mean values generally below 1°C. This indicates that IDW preserves the central tendency of maximum temperatures. However, a consistent smoothing effect is observed in the lower tail. For example, in S.J. dos Campos, the minimum of daily maxima increases from 9.80°C (station) to 16.00°C (municipality), indicating a substantial reduction in variability. Similar patterns are observed in other regions, although less pronounced. At the upper tail, extreme maximum temperatures are well captured, with small differences in absolute maxima, suggesting that heat extremes are relatively robust to interpolation.

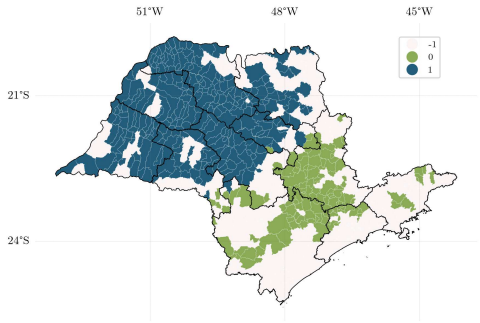


Fig. 6 Spatial distribution of municipality clusters obtained using DBSCAN applied to PCA-transformed features combining interpolation differences and geographic attributes. Colors indicate cluster membership, highlighting regions with similar behavior in terms of discrepancy between IDW and Nearest Neighbor methods. Areas classified as noise (cluster -1) are primarily associated with extreme elevations and greater distance from stations, where interpolation uncertainty is higher.

Mean temperatures are generally well reproduced (Table 5), with differences typically below 1°C, indicating that the overall thermal regime is captured adequately. Nevertheless, variability is reduced in the interpolated data, particularly at the lower tail. In S.J. dos Campos, the minimum mean temperature increases from 3.57°C (station) to 8.00°C (municipality), reflecting a loss of cold variability. In Campinas, the interpolated mean is slightly lower than the observed mean, reinforcing the presence of regional biases.

In contrast to maximum temperatures, minimum temperatures exhibit more pronounced discrepancies (Table 5). The interpolation systematically attenuates cold extremes. For instance, in S.J. dos Campos, the minimum temperature increases from -2.50°C (station) to -1.78°C (municipality). In Campinas, the difference is even more pronounced, with a minimum of 1.10°C at stations compared to -1.45°C in interpolated values, indicating not only smoothing of extremes but

Table 5 Daily maximum, mean, and minimum temperatures for each intermediate region, comparing observed station data and interpolated municipality values.

		Maximum				Mean				Minimum			
Region	Source	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max
Araraquara	Municipalities	30.20	3.43	14.10	39.95	22.24	3.11	9.06	29.97	16.08	3.75	0.30	23.18
	Stations	29.43	3.74	11.70	39.70	22.05	3.31	8.81	30.38	16.44	3.85	0.20	28.50
Araçatuba	Municipalities	32.08	3.76	15.77	42.08	23.97	3.31	9.24	31.93	17.75	4.01	-0.40	25.45
	Stations	31.58	3.98	13.20	42.00	24.46	3.52	8.23	33.10	18.81	4.28	-0.60	32.30
Bauru	Municipalities	30.54	3.79	13.14	40.43	22.38	3.30	8.62	30.76	16.29	3.70	0.41	23.94
	Stations	30.39	3.95	13.10	40.50	23.09	3.47	9.68	32.18	17.21	3.57	2.10	26.40
Campinas	Municipalities	29.33	3.41	15.48	39.46	20.74	3.09	7.82	28.78	13.23	3.75	-1.45	20.51
	Stations	28.98	3.86	12.40	39.50	21.45	3.33	7.48	30.72	15.46	3.99	1.10	30.40
Marília	Municipalities	30.18	3.81	12.95	39.69	22.69	3.42	8.15	30.92	16.46	3.83	-0.42	24.39
	Stations	29.92	3.91	12.10	39.80	23.19	3.43	8.70	32.24	17.67	3.54	1.90	26.60
P. Prudente	Municipalities	31.29	3.90	14.03	40.91	23.80	3.53	8.59	31.96	16.16	4.95	-2.79	26.28
	Stations	30.95	3.98	13.00	40.90	23.50	3.60	8.26	31.67	15.86	5.31	-3.80	25.30
Ribeirão Preto	Municipalities	31.35	3.24	16.81	41.24	22.60	2.94	9.78	29.96	15.55	4.03	-0.47	22.10
	Stations	31.51	3.27	16.90	41.30	22.84	2.98	10.00	30.32	15.59	4.13	-0.60	22.90
S.J. do Rio Preto	Municipalities	31.60	3.34	16.66	41.65	23.69	3.03	9.96	31.03	16.73	4.13	0.97	23.97
	Stations	31.42	3.37	16.70	41.70	24.22	3.13	11.89	33.47	17.16	4.21	0.90	24.30
S.J. dos Campos	Municipalities	27.73	3.92	16.00	38.14	19.54	3.16	8.00	27.58	11.83	3.74	-1.78	20.10
	Stations	27.15	4.90	9.80	38.80	18.67	3.80	3.57	28.51	11.02	3.95	-2.50	23.90
Sorocaba	Municipalities	28.85	4.12	15.14	40.77	20.50	3.35	7.16	28.26	14.16	4.13	-2.02	21.82
	Stations	28.50	4.35	14.70	41.20	20.70	3.60	8.38	30.80	14.22	4.29	-2.20	24.80
São Paulo	Municipalities	27.03	3.99	15.36	37.43	20.25	3.23	7.30	28.69	15.38	3.42	1.25	22.21
	Stations	27.54	4.05	15.50	39.00	20.68	3.27	8.04	29.05	15.74	3.38	1.80	22.30

also distortions in distribution shape. Additionally, the mean minimum temperature in Campinas is substantially lower in the interpolated data (13.23°C) compared to station observations (15.46°C), indicating the presence of systematic bias in addition to smoothing effects.

4.3.1 Distributional Comparison

Figure 7 presents boxplots of daily maximum, mean, and minimum temperatures, allowing a direct comparison of the full distributions.

For maximum temperatures, distributions are similar across regions, with close agreement in medians and upper quartiles. However, municipality data exhibit slightly narrower interquartile ranges, indicating reduced variability. For mean temperatures, central tendencies remain consistent, but variability is systematically reduced, particularly for cooler conditions. The most pronounced differences occur for minimum temperatures. The boxplots show a clear upward shift in the lower tails for municipality data, indicating attenuation of cold extremes. In regions such as S.J. dos Campos and Campinas, both minimum values and lower quartiles are higher in the interpolated data, confirming that extreme low temperatures are not fully captured.

These results reinforce that IDW interpolation preserves central tendencies while reducing variability, with the strongest effects observed in the lower tails. Additionally, some regions exhibit differences not only in spread but also in median values, suggesting that interpolation may introduce biases beyond simple smoothing, particularly for minimum temperatures.

Overall, the boxplots reinforce the findings from Tables 5, highlighting that interpolation preserves central tendencies but reduces variability, with the strongest effects observed in the lower tails of the temperature distributions. This behavior is consistent with the spatial averaging inherent to IDW, which tends to dampen localized extremes, particularly under heterogeneous terrain and sparse station coverage.

Among the analyzed regions, S.J. dos Campos consistently exhibits the largest discrepancies, particularly for minimum and mean temperatures. This suggests a strong influence of local-scale variability, likely associated with complex topography and heterogeneous climatic conditions.

Other regions, such as Campinas, also present systematic differences, particularly in minimum temperatures, indicating that interpolation errors are not limited to extremes but may reflect broader representativeness issues.

4.4 Climate Indicators: Heat Wave Analysis

The impact of spatial interpolation on climate indicators was evaluated by comparing heat wave

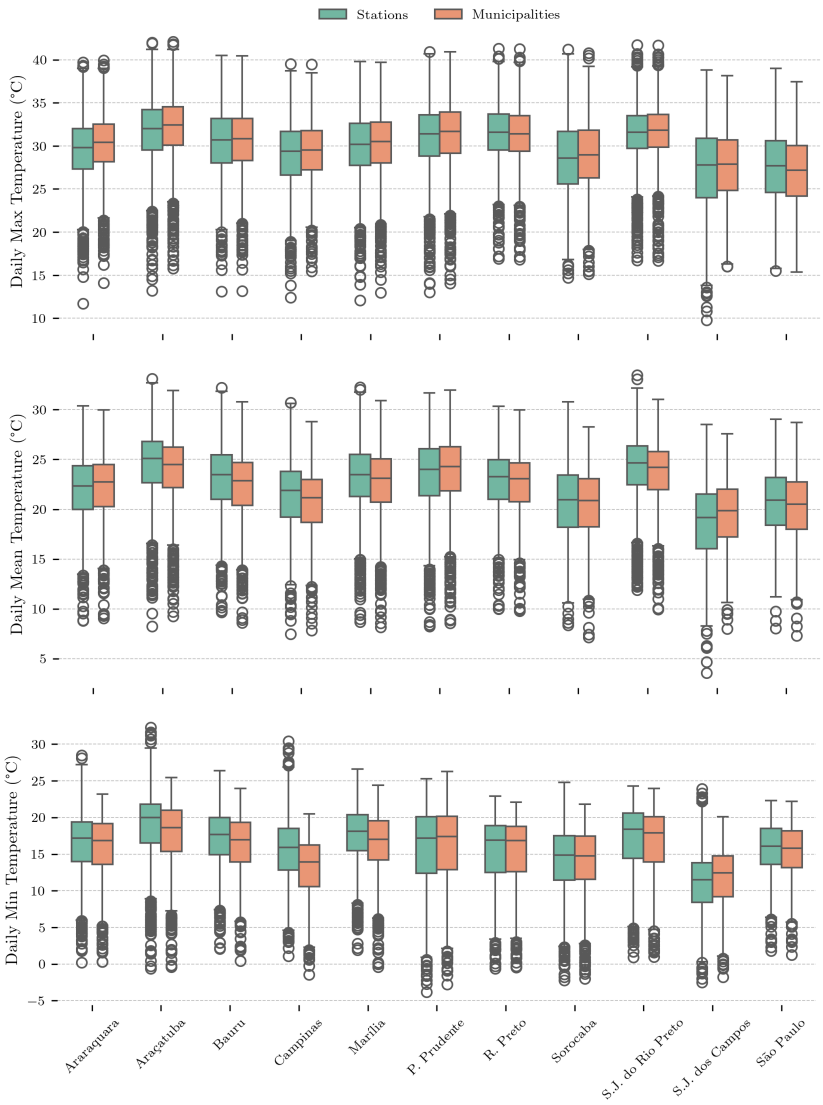


Fig. 7 Boxplots of daily maximum (top), mean (middle), and minimum (bottom) temperatures for each intermediate region, comparing observed station data and interpolated municipality values. The distributions highlight differences in central tendency, variability, and extremes between the two data sources.

detection based on the nearest-neighbor (NN) method and the inverse distance weighting (IDW) configuration ($R = 250$ km, $\alpha = 2$).

Table 6 reveals a consistent and substantial difference between the two approaches. Across all municipalities, the NN method identifies a significantly higher number of heat wave events and total heat wave days than the IDW interpolation.

While differences are already evident for short-duration events (3- and 4-day heat waves), they become markedly larger for longer events (5+ days), where NN detects approximately two to

Table 6 Comparison of heat wave metrics using the nearest-neighbor (NN) method and the inverse distance weighting (IDW) configuration with search radius $R = 250$ km and exponent $\alpha = 2$.

		3-day		4-day		5+day		Total HW days		
Municipality	Region	NN	(250,2)	NN	(250,2)	NN	(250,2)	NN	(250,2)	Diference
Pereira Barreto	Araçatuba	19	11	13	5	30	11	325	126	199
Ilha Solteira	Araçatuba	19	10	13	7	30	11	325	126	199
Murutinga do Sul	Araçatuba	19	11	13	7	30	11	325	134	191
Andradina	Araçatuba	19	12	13	6	30	13	325	147	178
Itapura	Araçatuba	19	10	13	8	30	13	325	149	176
Bilac	Araçatuba	21	5	8	3	18	6	212	68	144
Inúbia Paulista	P. Prudente	21	6	8	5	18	5	212	73	139
Piacatu	Araçatuba	21	4	8	5	18	6	212	73	139
Gabriel Monteiro	Araçatuba	21	4	8	5	18	6	212	73	139
Lucélia	P. Prudente	21	9	8	6	18	6	212	92	120

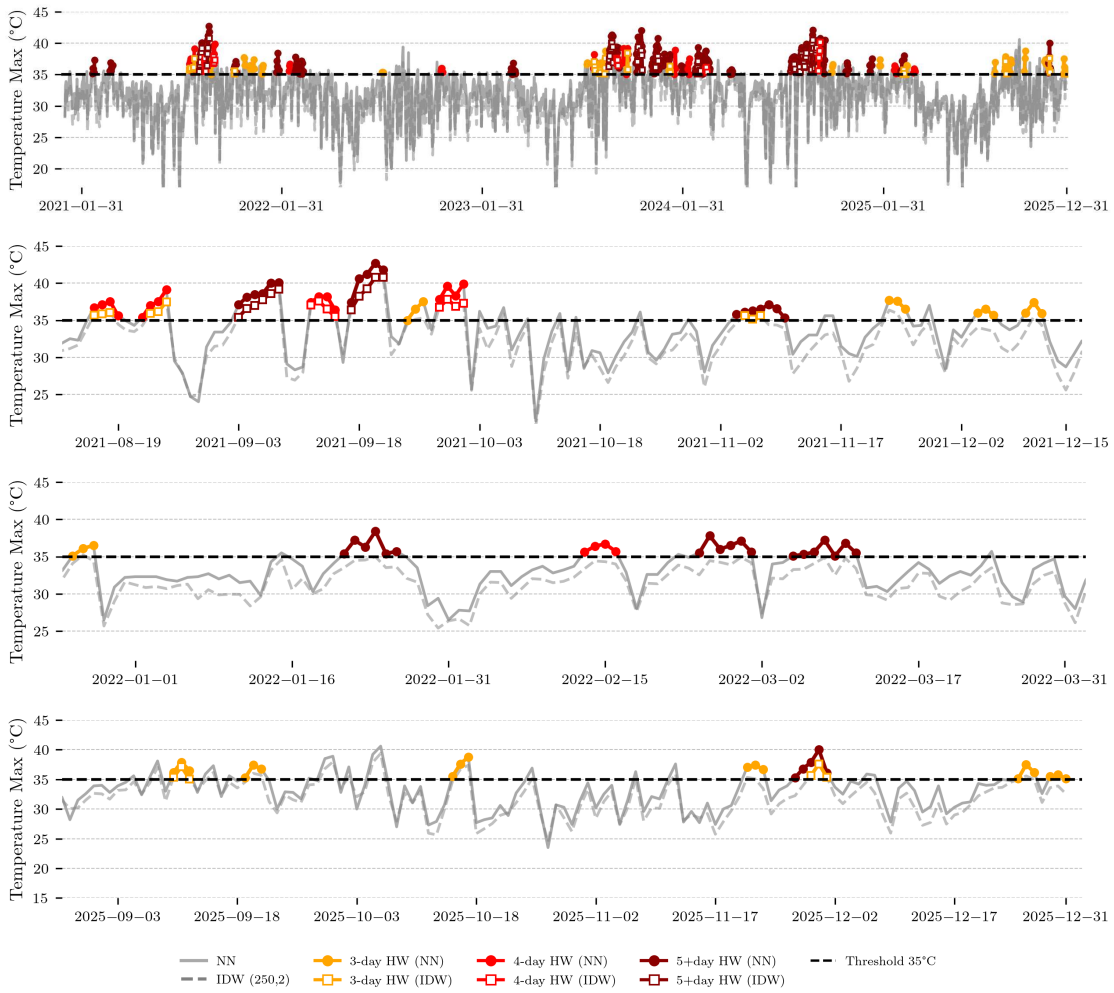


Fig. 8 Daily maximum temperature and identified heat waves for the municipality of Pereira Barreto (3537404) under the nearest-neighbor (NN) and inverse distance weighting (IDW, $R = 250$ km, $\alpha = 2$) methods. The panels display the full period (2021–2025) and selected subperiods highlighting differences in heat wave detection. Pereira Barreto, along with Ilha Solteira (3520442), exhibits the largest discrepancy in the total number of heat wave days between the two methods. Colors indicate heat wave duration (3-, 4-, and 5+ days), while markers distinguish methods (circles for NN and squares for IDW). Solid and dashed lines represent NN and IDW temperature series, respectively, and the horizontal dashed line denotes the 35°C threshold.

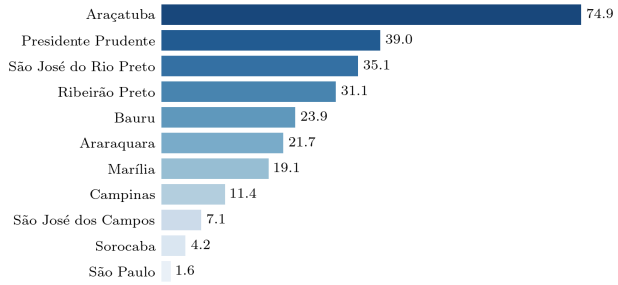


Fig. 9 Regional averages of the difference in total heat wave days between the nearest-neighbor (NN) and inverse distance weighting (IDW, $R = 250$ km, $\alpha = 2$) methods. The NN approach systematically identifies more heat wave days, particularly in western regions such as Araçatuba, indicating substantial sensitivity to the choice of spatial aggregation method.

three times more occurrences. This translates into large discrepancies in cumulative heat wave days. For example, municipalities such as Pereira Barreto and Ilha Solteira exhibit differences of up to 199 days, while even the smallest gap in the sample remains above 120 days.

These results indicate that IDW interpolation systematically reduces both the frequency and persistence of heat waves. This behavior is directly linked to the smoothing effects identified in the previous section. By averaging spatial information, IDW attenuates temperature variability, particularly reducing the persistence of consecutive high-temperature days. As a result, sequences that would qualify as heat waves under NN are often shortened or no longer exceed the required duration thresholds under IDW.

In contrast, the NN method preserves local extremes and temporal continuity, leading to higher sensitivity in detecting

Figure 8 illustrates these differences for the municipality of Pereira Barreto. The NN approach identifies more frequent and longer-lasting heat wave events, particularly during periods of sustained high temperatures. In contrast, the IDW method smooths temperature variations, resulting in fewer detected events and shorter durations.

This effect is especially evident in the highlighted subperiods, where sequences classified as long heat waves (5+ days) under NN are either shortened or not detected under IDW. These

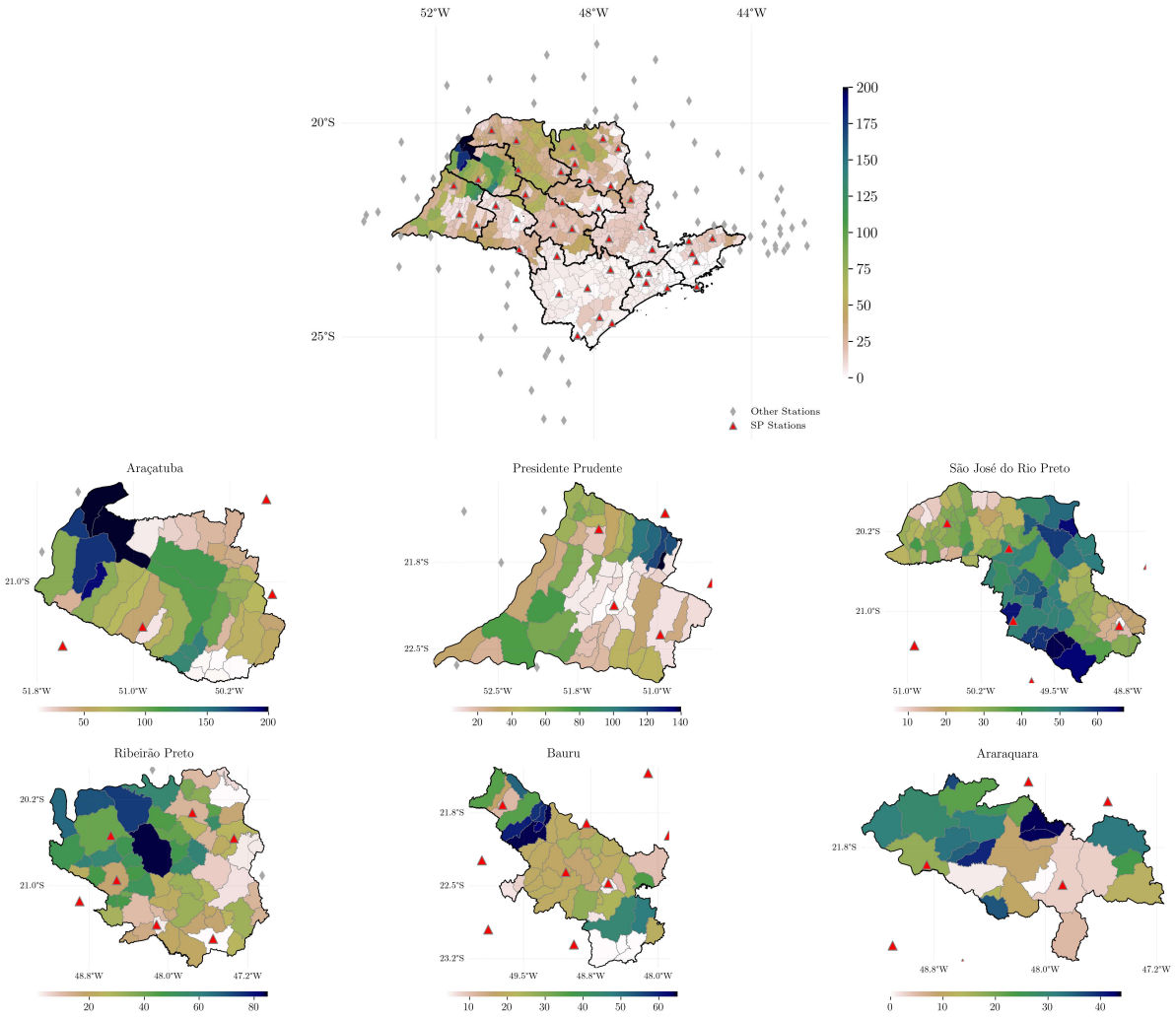


Fig. 10 Heatwave maps for the state of São Paulo (top) and selected regions (bottom panels).

visual patterns are consistent with the quantitative differences reported in Table 6.

The regional analysis (Figure 9) and the corresponding spatial patterns (Figure 10) confirm that these discrepancies are spatially structured. Western regions, such as Araçatuba, exhibit the largest differences in total heat wave days, indicating that the sensitivity of heat wave detection to interpolation choice is not uniform across the study area.

This spatial variability reflects the interaction between station density, environmental heterogeneity, and interpolation behavior.

Overall, these results demonstrate that the choice of spatial interpolation method has a direct and non-negligible impact on climate-indicator estimation. In particular, methods that introduce

stronger spatial smoothing, such as IDW, may modify the frequency and duration of detected heat-wave events. These differences arise because even small changes in estimated temperatures can affect the persistence of threshold exceedances used in heat-wave identification.

This finding has important implications for climate impact assessments, as it suggests that interpolation choices can influence the characterization of extreme events, potentially affecting risk evaluation and decision-making processes.

5 Discussion

5.1 Comparison with previous studies

The results demonstrate that spatial interpolation performance is strongly influenced by environmental heterogeneity and station representativeness. The identification of distinct performance clusters indicates that interpolation errors are spatially structured rather than randomly distributed, reflecting the interaction between station-network geometry, terrain complexity, and local climatic gradients. Similar behavior has been reported in previous studies conducted in regions with sparse observational coverage and complex topography, where the spatial representativeness of individual stations is reduced [18–20].

Distance-based interpolation methods such as IDW were found to systematically smooth temperature fields, modifying both variability and extremes. The attenuation of local variability observed in this study is consistent with previous geostatistical investigations showing that interpolation methods may introduce structured smoothing effects rather than purely random errors [21]. These findings are particularly relevant because interpolated temperature products are frequently used as surrogate observations in climate-impact and risk-assessment studies conducted at administrative scales [22]. Our results indicate that the same interpolation procedure capable of improving average agreement with observations may simultaneously suppress local variability and temporal persistence, which are fundamental characteristics for the identification of extreme events such as heat waves.

Despite these limitations, interpolated datasets remain an important alternative in regions with sparse or incomplete observational networks. They provide a feasible framework for estimating spatial temperature patterns and deriving climate indicators at municipality and regional scales, supporting the analysis of heat waves and other climate-related phenomena in data-scarce environments.

5.2 Spatial structure of interpolation performance

At the station level, poorly performing locations are associated with heterogeneous environments, indicating that interpolation based solely on distance is insufficient to capture localized variability. This limitation is amplified in regions with complex topography or sparse observational coverage, where spatial gradients are sharper and less well represented by neighboring stations.

Such behavior is expected in landscapes where altitude, coastal influence, land–sea contrasts, and mesoscale circulation produce abrupt thermal transitions over relatively short distances. In these contexts, geographic proximity alone may be a weak proxy for climatic similarity [23, 24]. Consequently, distance-based interpolation may perform adequately in relatively homogeneous interior regions, while failing to reproduce localized conditions in mountainous, coastal, or transitional areas.

At the municipality level, interpolation preserves central tendencies but systematically reduces variability. This effect is particularly evident for minimum temperatures, where cold extremes are attenuated and, in some cases, the distribution is distorted. Mean temperatures may therefore appear robust to interpolation, giving the impression that interpolated surfaces are reliable. However, many climate indicators depend not only on averages, but also on thresholds, tails, and temporal persistence. Consequently, preserving mean conditions is insufficient when the interpolation process modifies the lower or upper tails of the distribution.

Taken together, these findings reveal consistent patterns across temperature metrics:

- **Maximum temperatures:** Accurately represented in terms of mean and upper extremes, with moderate smoothing of lower-tail variability.
- **Minimum temperatures:** Most affected variable, with clear attenuation of cold extremes and occasional systematic bias. This behavior is particularly relevant for applications sensitive to low-temperature events, such as agricultural planning and frost-risk assessment, since reduced variability in interpolated data may lead to underestimation of risk.
- **Mean temperatures:** Well reproduced in terms of central tendency, but with reduced variability.
- **Spatial variability:** Regions with higher environmental heterogeneity exhibit larger discrepancies, highlighting limitations of centroid-based interpolation.

These results indicate that IDW does not simply introduce random deviations around observed conditions, but reshapes the empirical distribution of temperature in ways that depend on both the

metric considered and the environmental context. The stronger sensitivity of minimum temperatures may reflect the greater influence of local-scale processes on nocturnal and early-morning thermal conditions, including cold-air drainage, elevation effects, valley circulation, surface moisture, and land-cover heterogeneity. Such processes are difficult to represent using interpolation schemes based only on horizontal distance, making the lower tail of the temperature distribution particularly vulnerable to spatial smoothing. This has implications beyond heat-wave analysis, including applications related to agriculture, ecosystem stress, and temperature-related health outcomes [25–27].

More broadly, the findings suggest that interpolation uncertainty should be explicitly considered in climate-impact analyses and decision-support frameworks based on threshold-dependent indicators. They also highlight the need for future methodologies capable of propagating interpolation uncertainty into climate indicators, particularly for extreme-event detection in data-sparse environments. Such approaches could provide confidence intervals for heat-wave metrics and improve the interpretation of climate-risk assessments based on interpolated datasets.

5.3 Heat wave detection as a nonlinear amplification mechanism

The sensitivity of extreme temperature metrics to interpolation becomes particularly evident in the analysis of heat waves. Compared with the nearest-neighbor approach, IDW resulted in fewer and shorter detected heat-wave events. This difference is not only quantitative but also structural, because spatial smoothing reduces the persistence of consecutive high-temperature days. Consequently, sequences satisfying the heat-wave criteria under one interpolation method may not be identified under another. These results suggest that interpolation choice can influence not only estimated temperature values but also the classification of climate events.

Importantly, these discrepancies arise even though interpolation errors in maximum temperature are relatively small. This is consistent with a nonlinear amplification mechanism, where heat-wave detection is based on fixed thresholds, and small differences in estimated maximum temperature can substantially alter event classification. Small underestimations may prevent interpolated temperatures from exceeding the threshold, breaking sequences of consecutive days and reducing the number and duration of detected heat-wave events.

This mechanism is particularly evident in the municipalities of Pereira Barreto and Ilha Solteira, where cumulative heat-wave duration differed by up to 199 days between the nearest-neighbor and IDW approaches (Table 6). Such differences illustrate how methodological choices in spatial interpolation may modify the detection and characterization of extreme temperature events, even when differences in the underlying temperature estimates are relatively small, thereby influencing downstream climate-impact assessments.

The effect was particularly pronounced in regions characterized by higher temperature regimes, including Araçatuba, Presidente Prudente, and São José do Rio Preto. In these areas, temperatures frequently approach or exceed the heat-wave threshold, making event classification highly sensitive to small perturbations. Therefore, even modest smoothing can generate disproportionately large differences in estimated heat-wave duration. The western portion of São Paulo illustrates a threshold-amplification mechanism in which small interpolation-induced changes produce large differences in exposure estimates. This also explains why the largest discrepancies are not necessarily found where interpolation errors are highest.

These results indicate that conventional performance metrics such as RMSE, MAE, and bias may be insufficient for evaluating interpolated datasets intended for extreme-event analysis. A method can exhibit acceptable average accuracy while still producing substantial distortions in threshold-based climate indicators. Therefore, interpolation evaluation should consider not only pointwise errors but also impacts on event frequency, duration, persistence, and distributional characteristics.

The regional analysis further indicates that these discrepancies are spatially structured Western regions combine higher temperatures and relatively sparse station coverage, which increases reliance on spatial interpolation and may amplify its effects on derived indicators This result reinforces the importance of evaluating interpolation methods not only globally, but also regionally and according to the intended application. For decision-making, a spatially averaged error metric may be less informative than knowing where and under which climatic regimes the method changes the classification of extreme events [28, 29].

Overall, the impact of interpolation is not uniform across temperature metrics. While average conditions and even maximum values may appear well represented, threshold-based climate indicators are inherently more sensitive to small estimation errors. Accurate representation of extremes therefore requires not only low overall error, but also sufficient spatial resolution and data density to preserve local variability and temporal continuity. This finding questions the assumption that

an interpolator selected on the basis of cross-validation accuracy will necessarily be adequate for downstream climate-impact analyses based on derived indicators. For applications involving extremes, model evaluation should include distributional diagnostics, tail behavior, persistence metrics, and direct validation of derived indicators [30, 31].

5.4 Implications for climate-risk assessment and environmental decision-making

From an applied perspective, these findings highlight that interpolation should not be treated as a neutral preprocessing step. Its influence propagates through the analytical pipeline and can significantly affect downstream results, particularly in extreme event detection and risk assessment.

This has direct implications for applications that increasingly depend on gridded or spatially interpolated climate information, including municipal heat-risk assessments, adaptation planning, early-warning systems, infrastructure management, and public-health preparedness. If interpolation suppresses the persistence of heat events, it may underestimate the duration of population exposure and the cumulative stress associated with prolonged heat conditions [32–34].

These findings are also relevant to broader sustainability agendas, particularly those related to sustainable cities, climate action, and health protection [35]. The results do not directly measure health, energy, agricultural, or social outcomes, but they show that the climate indicators used to inform such sectors may be sensitive to interpolation choices. In data-sparse regions, where interpolated products can provide the main empirical basis for local climate assessment, this sensitivity may affect resource allocation, prioritization of adaptation measures, and the identification of municipalities most exposed to extreme heat [36].

The consistency observed between station-level and municipality-level clustering further suggests that these effects are robust across spatial scales. Uncertainty identified at the station level does not disappear when data are aggregated to municipalities or intermediate regions; instead, it may propagate into systematic spatial patterns of over- or under-detection in climate indicators.

5.5 The trade-off between average accuracy and preservation of extremes

The results suggest that the concept of the "best" interpolation method depends on the purpose of the analysis. If the objective is to minimize average prediction error, IDW may outperform a nearest-neighbor baseline. However, if the objective is to preserve local extremes and consecutive threshold exceedances, stronger smoothing may become

a disadvantage. Rather than identifying a universally superior interpolation method, the findings indicate that method selection should be aligned with the climate indicator and application of interest.

Several limitations should be acknowledged. First, the analysis focuses on nearest-neighbor and IDW approaches, which represent operationally simple methods with contrasting smoothing behavior. More complex geostatistical or hybrid approaches incorporating elevation, land cover, coastal distance, or atmospheric covariates may reduce some of the distortions identified here.

The objective of this study was not to identify the optimal temperature interpolation framework for São Paulo state, but rather to evaluate how commonly used distance-based interpolation methods influence the estimation of climate indicators and extreme-event metrics. The comparison between nearest-neighbor and IDW was therefore designed to isolate the effects of different levels of spatial smoothing under operationally simple approaches that are frequently applied in data-sparse environments. Future studies should investigate whether interpolation methods that explicitly incorporate elevation, land cover, coastal influence, or other environmental covariates can better preserve local variability and improve the representation of threshold-based climate indicators.

This issue becomes particularly evident in eastern São Paulo state, near the Serra da Mantiqueira and the borders with Rio de Janeiro and Minas Gerais. In municipalities such as Queluz and neighboring areas, temperature estimates depend strongly on the Cachoeira Paulista station due to the complex topography and limited spatial coverage of nearby observations. When this station presents missing data, the interpolation tends to rely more heavily on high-elevation stations located in the Mantiqueira region, leading to systematic underestimation of local temperatures. This example illustrates how sparse observational coverage combined with strong topographic gradients can amplify interpolation sensitivity and distort local thermal patterns.

Second, municipality centroids provide a practical and reproducible spatial unit for administrative-scale assessment, but they cannot fully represent within-municipality heterogeneity, particularly in large or topographically complex municipalities, especially in the eastern portion of São Paulo state.

In addition, the municipality-level heat-wave metrics evaluated in this study are derived from interpolated temperature series rather than direct observations at the municipal scale. Consequently, differences in heat-wave frequency and duration should be interpreted as methodological sensitivities associated with the interpolation approach rather than as direct evidence of true climatic

differences. While these results demonstrate how interpolation choices can propagate into climate indicators, the magnitude of the discrepancies remains conditional on the underlying interpolation assumptions and observational network.

Third, the heat-wave definition relies on a fixed threshold and consecutive-day criterion. Alternative percentile-based or region-specific thresholds could modify the absolute number of detected events, although the sensitivity of threshold-based indicators to interpolation effects would likely remain relevant. Fourth, the study period is relatively short and does not capture long-term trends in heat-wave frequency. Finally, the analysis depends on the density, continuity, and quality of the meteorological-station network, which varies across space and time.

Overall, the findings support a shift from interpolation evaluation based solely on average error toward evaluation frameworks that explicitly assess whether interpolated datasets preserve the statistical characteristics required by downstream applications. For climate-risk assessment, this means evaluating not only whether interpolated temperatures are close to observations, but whether they preserve extremes, persistence, spatial contrasts, and event classifications [37]. Spatial interpolation is therefore not simply a technical step in climate-data preparation, it is a methodological decision capable of reshaping the evidence base used for adaptation and environmental policy [38].

6 Conclusions

This study assessed the performance of spatial interpolation methods for temperature estimation under sparse observational conditions and investigated their implications for climate indicators and extreme-event characterization. The results demonstrate that interpolation performance exhibits pronounced spatial heterogeneity. Cluster-based analyses revealed distinct performance regimes, indicating that interpolation errors are spatially structured rather than randomly distributed, reflecting both environmental variability and differences in station representativeness. Distance-based methods, particularly IDW, preserved the central tendency of temperature fields but reduced temperature variability, with the strongest effects observed for minimum temperatures. These results indicate that interpolation can modify the statistical representation of temperature distributions, especially in regions where local-scale processes strongly influence thermal conditions.

The influence of interpolation propagated directly into derived climate indicators. Heat-wave analyses revealed substantial discrepancies between interpolation approaches, with IDW resulting in fewer and shorter detected heat-wave

events relative to the nearest-neighbor method. In several regions, differences exceeded 100 days in cumulative heat-wave duration over the study period. Importantly, these discrepancies emerged through a nonlinear amplification mechanism: although interpolation errors in maximum temperature were comparatively small, threshold-based indicators proved highly sensitive to such deviations. Even minor underestimations in interpolated temperatures could disrupt sequences of threshold exceedances, substantially altering the identification of extreme events.

Collectively, these findings demonstrate that spatial interpolation should not be regarded as a neutral preprocessing procedure. Instead, it constitutes a methodological decision capable of modifying the statistical properties of interpolated temperature datasets and the characterization of climate extremes derived from them. The results also show that satisfactory average interpolation accuracy does not necessarily imply reliable representation of threshold-based climate indicators.

Therefore, differences in municipality-level heat-wave indicators should be interpreted as differences in the representation of climate extremes produced by alternative interpolation approaches, rather than as direct estimates of changes in the underlying occurrence of heat waves.

These findings highlight the importance of carefully selecting and evaluating interpolation methods in climate applications, particularly when the focus involves extreme events and climate-risk assessment. Future work should explore approaches that better preserve local variability and improve the representation of extremes, including hybrid methods and models that incorporate environmental covariates such as elevation, land cover, and coastal influence. In addition, the strong sensitivity of threshold-based indicators to relatively small interpolation errors suggests the need for methodological frameworks that explicitly account for uncertainty in spatially interpolated climate data.

More broadly, this study reinforces the importance of considering how interpolation choices may propagate into downstream climate indicators and risk analyses. These challenges are likely to become increasingly relevant as climate services, adaptation planning, and impact assessments rely more heavily on spatially interpolated datasets at municipal and regional scales.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Fundação de Apoio à Universidade de São Paulo (FUSP) under Grant No. P4321.

Authors' contributions

FFB conceptualized the study, curated the data, conducted the formal analysis and investigation, developed the methodology and software, performed validation and visualization, and wrote the original draft and subsequent revisions of the manuscript. MLF contributed to the conceptualization, formal analysis, investigation, methodology, and supervision, and participated in writing the original draft and reviewing and editing the manuscript. DCB contributed to the conceptualization, formal analysis, investigation, methodology, supervision, visualization, funding acquisition, and participated in writing the original draft and reviewing and editing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

[1] Rezvani, S.M., Almeida, N.M., Falcão, M.J.: Climate adaptation measures for enhancing urban resilience. Buildings **13**(9), 2163 (2023)

[2] Perkins-Kirkpatrick, S.E., Lewis, S.C.: Increasing trends in regional heatwaves. Nature Climate Change **5**, 535–539 (2015)

[3] Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., Zhou, B. (eds.): Climate Change 2021: The Physical Science Basis. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (2021). https://doi.org/10.1017/9781009157896

[4] Ravindra, K., Bhardwaj, S., Gandhi, A.P., Attri, S.D., Mor, S.: Navigating the challenges of a warmer climate: strategies for mitigating impacts and fostering resilience in india. Air, Soil and Water Research **18**, 11786221251325137 (2025)

[5] Bera, M., Nag, P.K.: Global research trends on climate change-induced heat stress and its impacts on health and productivity of outdoor workers: a systematic review. Theoretical and Applied Climatology **157**(1), 1 (2026)

[6] Herrera, S., et al.: Evaluation of daily gridded temperature datasets. Journal of Geophysical Research: Atmospheres (2019)

[7] Lussana, C., et al.: Spatial interpolation of daily temperature in complex terrain. International Journal of Climatology (2018)

[8] Chen, L., Chakraborty, T.C., Casey, J.A., Huang, C.H., Chen, T.H.K.: Heat metrics and thresholds reshape population exposure and inequality signals. Environmental Research Letters (2026)

[9] Li, J., Heap, A.D.: A review of spatial interpolation methods for environmental scientists. Geoscience Australia Record (2011)

[10] Chilès, J.-P., Delfiner, P.: Geostatistics: Modeling Spatial Uncertainty. Wiley, ??? (2012)

[11] Lyra, G.B., et al.: Performance of inverse distance weighting for climate variables. Theoretical and Applied Climatology (2018)

[12] Liu, X., et al.: Effects of station density on interpolation accuracy of temperature. Climate Dynamics (2020)

[13] Fard, P., Chung, M.K.J., Estiri, H., Patel, C.J.: Spatio-temporal interpolation and delineation of extreme heat events in california between 2017 and 2021. Environmental Research **237**, 116984 (2023)

[14] Yamamoto, J.K.: Correcting the smoothing effect of ordinary kriging. Mathematical Geology (2005)

[15] Wagener, T., Reinecke, R., Pianosi, F.: On the evaluation of climate change impact models. Wiley Interdisciplinary Reviews: Climate Change **13**(3), 772 (2022)

[16] Perkins, S.E., Alexander, L.V.: On the measurement of heat waves. Journal of Climate (2013)

[17] Langue, C.G.N., Lavaysse, C., Vrac, M., Flamant, C.: Heat wave monitoring over west african cities: uncertainties, characterization and recent trends. Natural Hazards and Earth System Sciences **23**(4), 1313–1333 (2023)

[18] Aldungarova, A., Utepov, Y., Mukhamejanova, A., Tulebekova, A., Nazarova, A., Tleubayeva, A., Mkilima, T.: Advancing

intermediate soil properties (isp) interpolation for enhanced geotechnical survey accuracy: A review. Engineering Reports **7**(8), 70328 (2025)

[19] Jiang, X., Xiong, X., Wang, W., Ye, X., Chen, X., Wang, Y., Zhang, F.: An improved interpolation algorithm for surface meteorological observations via fuzzy adaptive optimisation fusion. Atmosphere **16**(7), 844 (2025)

[20] Li, S., Zhang, B., Shi, B., Ai, Q., Zeng, Y., Yan, X., Wang, H.: High-precision spatial interpolation of meteorological variables in complex terrain using machine learning methods. Sensors **26**(7), 2167 (2026)

[21] Talvitie, J., Renfors, M., Lohan, E.S.: Distance-based interpolation and extrapolation methods for rss-based localization with indoor wireless signals. IEEE Transactions on Vehicular Technology **64**(4), 1340–1353 (2015)

[22] O'Gorman, E., Hochstrasser, T., Renou-Wilson, F.: Climate impact assessment procedures: time to look beyond compliance? Impact Assessment and Project Appraisal **42**(6), 553–564 (2024)

[23] Skyllingstad, E.D., Vickers, D., Mahrt, L., Samelson, R.: Effects of mesoscale sea-surface temperature fronts on the marine atmospheric boundary layer. Boundary-Layer Meteorology **123**(2), 219–237 (2007)

[24] Tang, R., Yu, Y., Xi, J., Ma, W., Wang, Y.: Mesoscale eddies induce variability in the sea surface temperature gradient in the kuroshio extension. Frontiers in Marine Science **9**, 926954 (2022)

[25] Achilleos, G.A.: The inverse distance weighted interpolation method and error propagation mechanism–creating a dem from an analogue topographical map. Journal of Spatial Science **56**(2), 283–304 (2011)

[26] Souza, E.G., Bazzi, C.L., Khosla, R., Uribe-Opazo, M.A., Reich, R.M.: Interpolation type and data computation of crop yield maps is important for precision crop production. Journal of Plant Nutrition **39**(4), 531–538 (2016)

[27] Abdulmanov, R., Miftakhov, I., Ishbulatov, M., Galeev, E., Shafeeva, E.: Comparison of the effectiveness of gis-based interpolation methods for estimating the spatial distribution of agrochemical soil properties. Environmental Technology & Innovation **24**, 101970 (2021)

[28] Naveau, P., Hannart, A., Ribes, A.: Statistical methods for extreme event attribution in climate science. Annual Review of Statistics

and Its Application **7**(1), 89–110 (2020)

[29] Zhang, H., Chapman, S., Trancoso, R., Toombs, N., Syktus, J.: Assessing the impact of bias correction approaches on climate extremes and the climate change signal. Meteorological Applications **31**(3), 2204 (2024)

[30] Fuentes, M., Henry, J., Reich, B.: Nonparametric spatial models for extremes: Application to extreme temperature data. Extremes **16**(1), 75–101 (2013)

[31] Huser, R., Wadsworth, J.L.: Advances in statistical modeling of spatial extremes. Wiley Interdisciplinary Reviews: Computational Statistics **14**(1), 1537 (2022)

[32] World Health Organization: Heat and Health in the WHO European Region: Updated Evidence for Effective Prevention. World Health Organization Regional Office for Europe, ??? (2021)

[33] Hess, J.J., Errett, N.A., McGregor, G., Busch Isaksen, T., Wettstein, Z.S., Wheat, S.K., Ebi, K.L.: Public health preparedness for extreme heat events. Annual Review of Public Health **44**(1), 301–321 (2023)

[34] Luo, Y., Cheng, X., He, B.J., Dewancker, B.J.: Identification and assessment of heat disaster risk: a comprehensive framework based on hazard, exposure, adaptation and vulnerability. International Journal of Environmental Science and Technology **22**(12), 11275–11294 (2025)

[35] Huang, Q., Li, Z., Wu, J., Khan, I., Rehman, M.Z.: Urbanization, climate change, and energy transition in china: A time-series assessment toward climate-resilient and sustainable cities. Physics and Chemistry of the Earth, Parts A/B/C, 104356 (2026)

[36] Webb, M.A., Hall, A., Kidd, D., Minansy, B.: Local-scale spatial modelling for interpolating climatic temperature variables to predict agricultural plant suitability. Theoretical and Applied Climatology **124**(3), 1145–1165 (2016)

[37] Brönnimann, S., Martius, O., Rohr, C., Bresch, D.N., Lin, K.H.E.: Historical weather data for climate risk assessment. Annals of the New York Academy of Sciences **1436**(1), 121–137 (2019)

[38] Digkoglou, P., Papathanasiou, J.: Application of multiple criteria decision aiding in environmental policy-making processes. International Journal of Environmental Science and Technology **22**(8), 6967–6982 (2025)