

Population exposure to temperature extremes in Spain: Integrating high-resolution climate and historical population data

Dariya Ordanovich¹ Ana Casanueva² Diego Ramiro¹

¹Institute of Economics, Geography and Demography, Spanish National Research Council, Madrid, Spain

²Applied Mathematics and Computer Sciences Department, University of Cantabria, Santander, Spain

Abstract

This work develops a standardized set of *population-weighted exposure indicators* for administrative units in Spain, designed to align gridded climate data with official statistics. The workflow links daily climate fields with high-resolution population rasters anchored to census years and reconciles aggregates to National Institute of Statistics (INE) totals through time-varying scaling. Population interpolation between anchors is treated as the required intermediate step to align population and climate on a daily basis; administrative labels are attached to environmental grid cells using population shares with a near-tie safeguard, and nominally empty cells fall back to geometric overlap. Building on this alignment, the indicators comprise: (i) core exposure (population-weighted minimum, maximum, mean, and apparent temperature, plus total population); (ii) absolute extremes in populated cells; (iii) threshold counts such as very hot days, freezing conditions, and tropical nights; (iv) population-weighted degree-day summaries; (v) extreme-event metrics for heatwaves and coldwaves based on concurrent exceedance of cell-specific percentiles using a 1971–2000 baseline, reporting population affected, daytime and nighttime magnitude, overall magnitude, event days, and episodes; and (vi) exposure variability within units. Computation is implemented with ERA5-Land and a Spain-wide 2.5 km daily observational grid (1961–2022+) to enable cross-source comparability. All inputs and outputs are versioned, with provenance for anchors, scaling factors, and boundary harmonization. The expected outcome is a set of comparable, administratively aligned, and policy-relevant indicators that represent conditions experienced by residents more accurately than simple area averages, particularly in urban, coastal, and complex orographic contexts.

1. Introduction

Environmental datasets are typically provided on regular spatial grids, whereas demographic data and policy operate on administrative units. Area-averaged climate summaries can therefore misrepresent conditions experienced by residents when population is unevenly distributed. We address this by developing population-weighted exposure indicators that are directly usable at administrative levels and comparable across environmental sources and periods.

2. Data Sources

Population. Population rasters are drawn from HIPGDAC-ES (1900–2021), which compiles 100 m grids from cadastral and harmonized municipal census data; census-year anchors provide the temporal basis for alignment across the climate window.¹ INE administrative totals provide reconciliation checkpoints and trajectories.

Administrative boundaries. Outputs are reported for census tracts (where applicable), municipalities, provinces, autonomous communities, and the national level. Provinces and autonomous communities are treated as fixed; finer divisions are harmonized to reflect historical changes.

Gridded climate data and statistics. The workflow is assembled and tested with ERA5-Land daily fields and a Spain-wide 2.5 km daily observational grid (1961–2022+) to evaluate cross-source comparability.

3. Methodological Workflow

Overview of the Framework

The pipeline integrates three domains: gridded population rasters (space-fixed but time-discrete), administrative geometries (time-varying boundaries), and environmental grids (space-fixed but time-continuous). Its main objectives are:

1. to generate daily, spatially continuous population estimates aligned with environmental grids,
2. to ensure that aggregated totals match official INE statistics at multiple administrative levels,
3. to provide population-weighted exposure metrics for any environmental variable.

Anchors and Climate Window

HIPGDAC-ES rasters exist for census years only (anchor dates, July 1 of each anchor year). The climate data, by contrast, cover a continuous daily window. The algorithm selects the smallest set of anchors that fully span the climate period. For each interval between anchors, cell-level population is interpolated daily.

Aggregation and Administrative Assignment

At each anchor, HIPGDAC pixels are aggregated to the environmental grid cells. Every pixel carries information about its administrative affiliation, allowing computation of province-level population shares per cell. A near-tie rule retains all provinces whose share lies within 5% of the maximum. For zero-population cells, a geometric fallback based on area overlap assigns administrative labels, ensuring coverage of all cells.

¹Goerlich, F. (2025). *HIPGDAC-ES: Historical Population Grid Data Compilation for Spain (1900–2021)*. *Scientific Data*. doi:10.1038/s41597-025-04533-8.

Temporal Interpolation and INE-Constrained Adjustment

For each cell, daily values between anchors are interpolated linearly:

$$\hat{P}_d^{\text{interp}}(g) = P_{\text{cell}}^{(y_0)}(g) + \frac{P_{\text{cell}}^{(y_1)}(g) - P_{\text{cell}}^{(y_0)}(g)}{N} \Delta d.$$

INE provides quarterly official totals at the provincial level. To reconcile interpolated values with these totals, a time-varying scale factor $s_A(t)$ is derived:

$$s_A(t) = \frac{\text{INE}_A(t)}{M_A(t)},$$

where $M_A(t)$ is the sum of interpolated cell populations for province A at time t . Between INE checkpoints, $s_A(t)$ is linearly interpolated. The final adjusted cell population is

$$\hat{P}_d(g, A) = s_A(d) \hat{P}_d^{\text{interp}}(g, A).$$

This adjustment preserves the relative spatial distribution within provinces while ensuring statistical consistency across hierarchical administrative levels.

Population-weighted indicator families

- **Core Exposure:** Metrics that describe the typical thermal environment experienced by the population within the administrative unit, weighted by population. Includes daily population-weighted minimum, maximum, mean, and apparent temperature, as well as total population. Useful for general understanding of cold or heat exposure.
- **Absolute Extremes:** Metrics capturing the most extreme temperatures within populated cells. Includes absolute daily maximum and minimum temperatures, populations exposed to specific thresholds (e.g., $T_{max} > 35^\circ\text{C}$ or $T_{min} < 0^\circ\text{C}$), and population-weighted degree days above defined heat or night-time temperature thresholds. Useful for identifying direct exposure to extreme conditions.
- **Extreme Events:** Metrics that identify and quantify significant heatwaves or coldwaves. Includes occurrence flags, population exposed, and magnitude indices (daytime, nighttime, overall) weighted by population. Useful for assessing widespread and intense extreme temperature events affecting human health.
- **Cumulative Exposure:** Metrics that quantify the total duration and frequency of extreme events over a period. Includes the number of days with heatwaves or coldwaves and the number of distinct episodes. Useful for understanding the temporal accumulation of extreme exposures.
- **Exposure Variability:** Metrics that measure the spatial heterogeneity of temperatures across populated cells. Includes standard deviations of daily maximum and minimum temperatures, and the population-weighted diurnal temperature range. Useful for highlighting areas with internal variability or uneven exposure within the administrative unit.

4. Ongoing Work and Expected Results

Development proceeds across all components in parallel. Current work completes daily alignment of population with climate, administrative assignment with a near-tie safeguard, and reconciliation to INE totals, while assembling the indicator families at municipal, provincial, and autonomous-community levels. The expected results are: (i) systematic divergence of population-weighted indicators from area-averaged values where residents cluster in warmer locations, affecting the timing

and apparent intensity of exposure; (ii) threshold counts and degree-day summaries that better reflect conditions in densely inhabited areas; (iii) heatwave and coldwave indicators capturing population affected and diurnal aspects, with nighttime components particularly relevant where tropical nights are frequent; (iv) cross-source comparability that confirms broad patterns while documenting source-dependent differences; and (v) packaged, versioned outputs and event files directly linkable to datasets reported on administrative units for applications in epidemiology, climate services, and adaptation planning.

5. Considerations

Within-year population dynamics are approximated by interpolation; event metrics depend on baseline selection; and historical boundary changes introduce uncertainty at fine spatial scales. These aspects are being quantified through sensitivity analyses and cross-source evaluation to ensure stable and interpretable administrative-level indicators.