Annual meeting PNC Clima
May 22, 2025

CMCC
Centro Euro-Mediterraneo sui Cambiamenti Climatici

Modelling Temperature–Health Relationships under Data Constraints

A methodological assessment using ERA5 and VHR-REA_IT data in urban settings

Domenico D'Ausilio

PhD Candidate in Economics Statistics and Sustainability
University of Naples Parthenope – CMCC (Euro-Mediterranean Center on Climate Change) Foundation

domenico.dausilio001@studenti.uniparthenope.it

Study aim and analytical scope

Research question of the PhD project

How can epidemiological models be designed and evaluated to ensure robust and generalisable estimates of temperature-related health impacts, particularly under conditions of limited, aggregated, or low-resolution data?

Two systematic reviews were conducted to explore:

- (A) how input structures (exposures, confounders, covariates) are specified in epidemiological models;
- (B) how outputs (health outcomes, stratifications, risk metrics) are defined and reported.

Preliminary application

A technical application is conducted for the city of Rome, comparing model estimates obtained using temperature data from two reanalysis sources with different spatial resolutions: ERA5 (~30 km) and VHR-REA_IT (~2 km).

Objective of the analysis

The goal is to assess whether higher-resolution data meaningfully improve epidemiological estimates, and to explore implications for modelling in urban areas with heterogeneous data availability.

Dataset Choice as a Source of Epidemiological Uncertainty

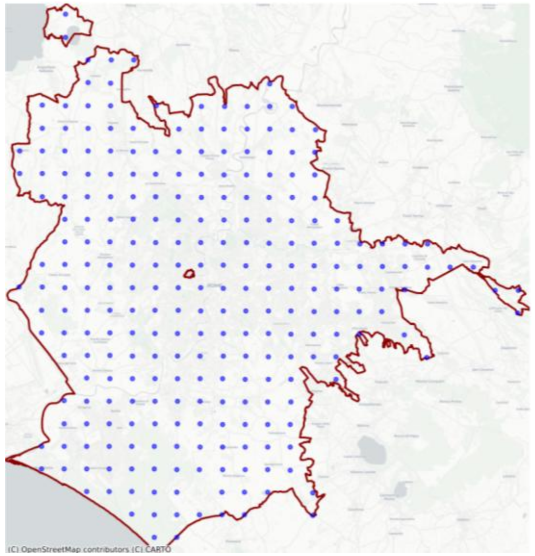
The problem

Climate-health models rely on gridded climate data (e.g. temperature, humidity) as inputs. These data are available at different spatial resolutions — from coarse reanalyses (e.g. ERA5) to fine-scale products (e.g. VHR-REA_IT). Yet, the impact of changing resolution is rarely assessed.

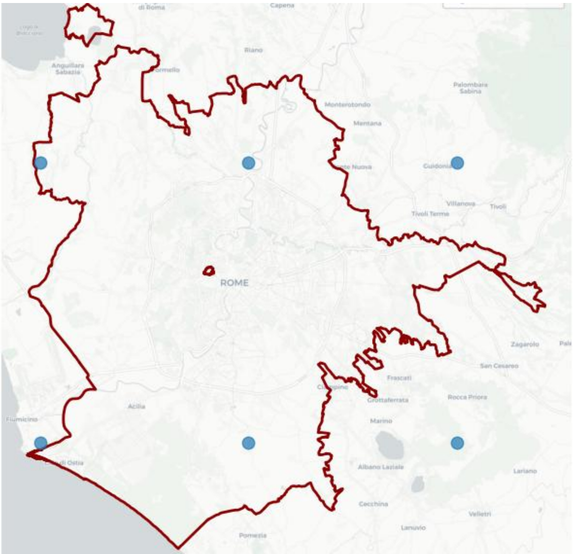
Scale differences

Even within the same location, the number and size of grid points can differ substantially — affecting how exposure is assigned and aggregated.

VHR_REA_IT (ITALY)



ERA5



Dataset comparison

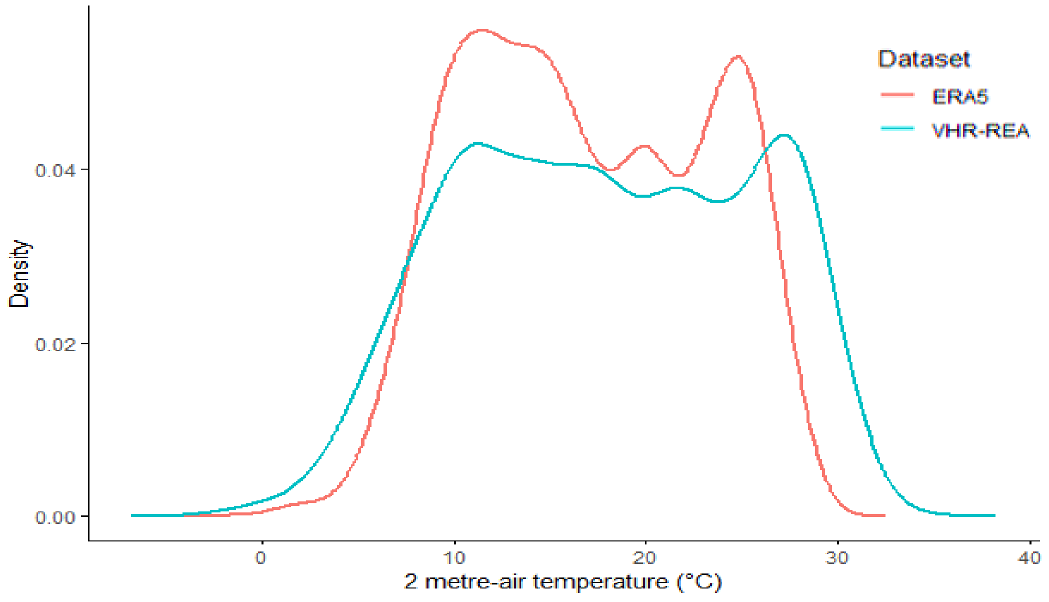
ERA5 (*Hersbach et al., 2020*)	VHR-REA_IT (*Raffa et al., 2021*)
Global reanalysis (ECMWF–C3S)	Very High Resolution Dynamical Downscaling of ERA5 Reanalysis over Italy (COSMO-CLM)
Spatial resolution: **~31 km**	Spatial resolution: **~2.2 km**
Temporal resolution: **Hourly**	Temporal resolution: **Hourly**
Coverage: **Global** (1950–present)	Coverage: **Italy** (1989–present)
Assimilates global observations (satellite, in situ) via **4D-Var**	Inherits boundary conditions from ERA5; integrates high-res urban physics
No explicit urban parameterisation	Includes **TERRA-URB urban parameterisation**
Excellent for **large-scale analysis** and comparability	Optimised for **urban and local-scale studies**
Smoothed topography and land cover	Captures **topographic and land-use heterogeneity** (e.g., cities, coasts)

🔍 Methodological Insight

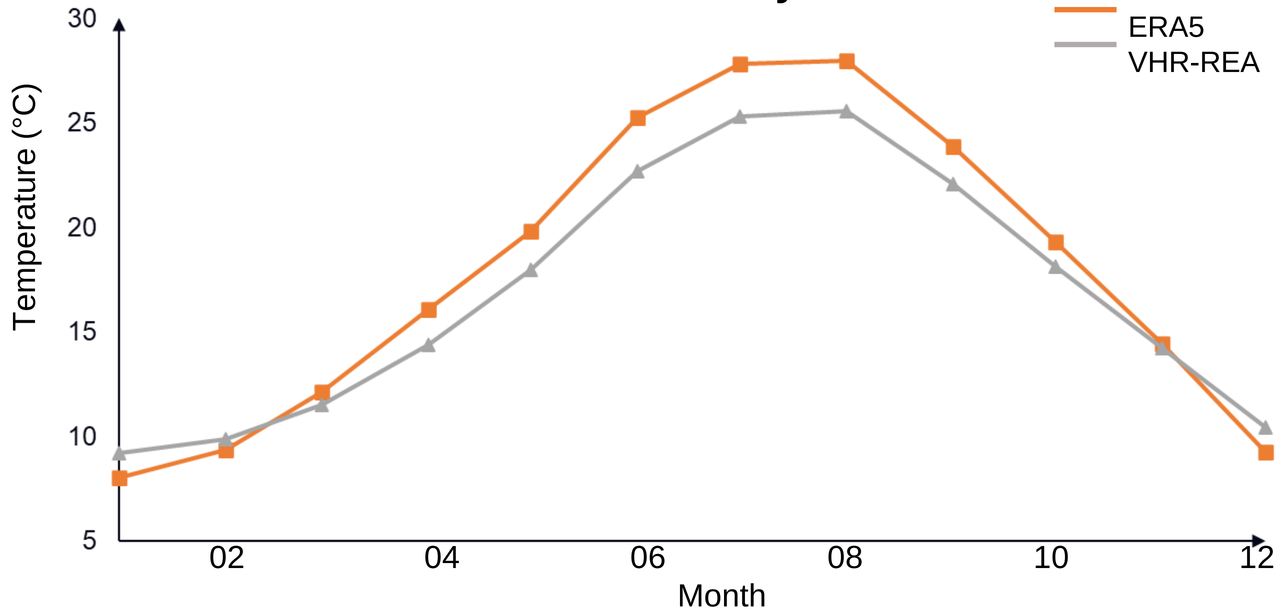
- Thanks to its finer spatial resolution and urban parameterization, **VHR-REA_IT better represents topographic detail and urban climate dynamics.**
- ERA5, may flatten urban heterogeneity due to its coarser grid and lack of explicit urban modelling.

Input differences

Probability Distribution Function



Multi-annual cycle



◇ **ERA5 exhibits a sharper peak around central temperatures**, indicating a narrower overall temperature range.

◇ **VHR-REA_IT shows fatter tails**, suggesting improved representation of temperature extremes.

◇ This reflects the **enhanced spatial resolution and urban sensitivity of the VHR-REA_IT dataset**.

◇ **ERA5 consistently reports higher monthly temperatures**, especially during summer months (June–August).

◇ **VHR-REA_IT smooths the annual cycle** slightly

◇ These discrepancies may **affect exposure assessment** and **heat–health model calibration**

Data sources and modelling approach

Period of analysis: 2011-2022

Data sources

Mortality data

- Daily all-cause mortality for the municipality of Rome
- Source: ISTAT, (2024)

Meteorological data

- Daily mean 2-metre air temperature
- Daily mean dew point temperature
- Relative humidity estimated from the Magnus-Tetens approximation (Alduchov & Eskridge, 1996)
- Source: VHR-REA_IT (Raffa et al., 2021); ERA5 (Hersbach et al., 2020)

Modelling approach

Statistical model

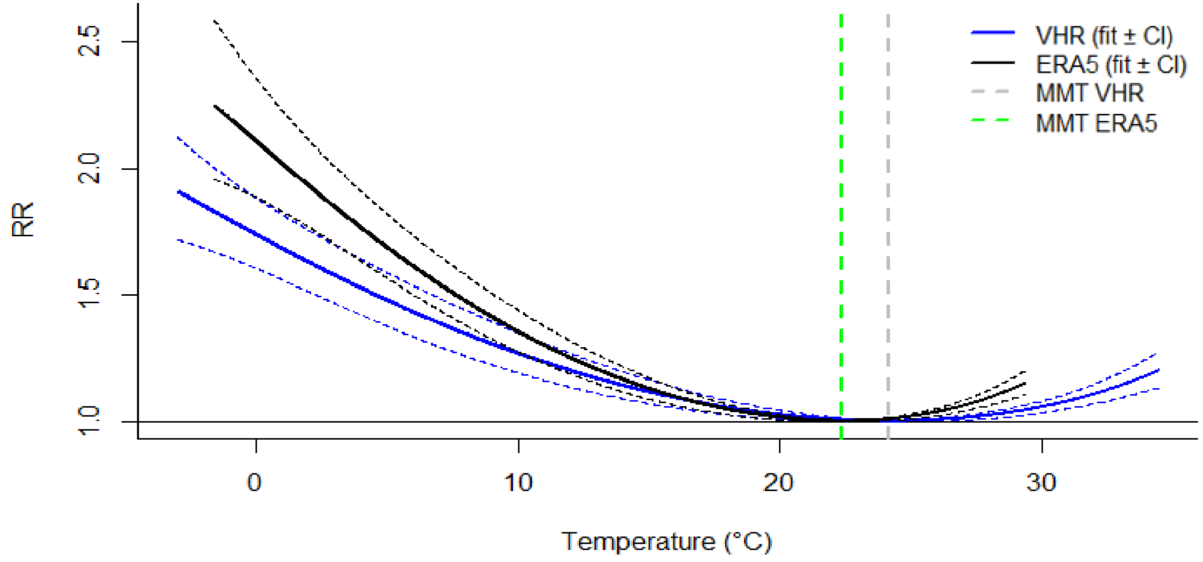
- Distributed Lag Non-Linear Model (*DLNM*) specified within a Generalised Additive Model (*GAM*)
- Natural cubic splines for temperature and lag
- Lag period: 28 days

Covariates

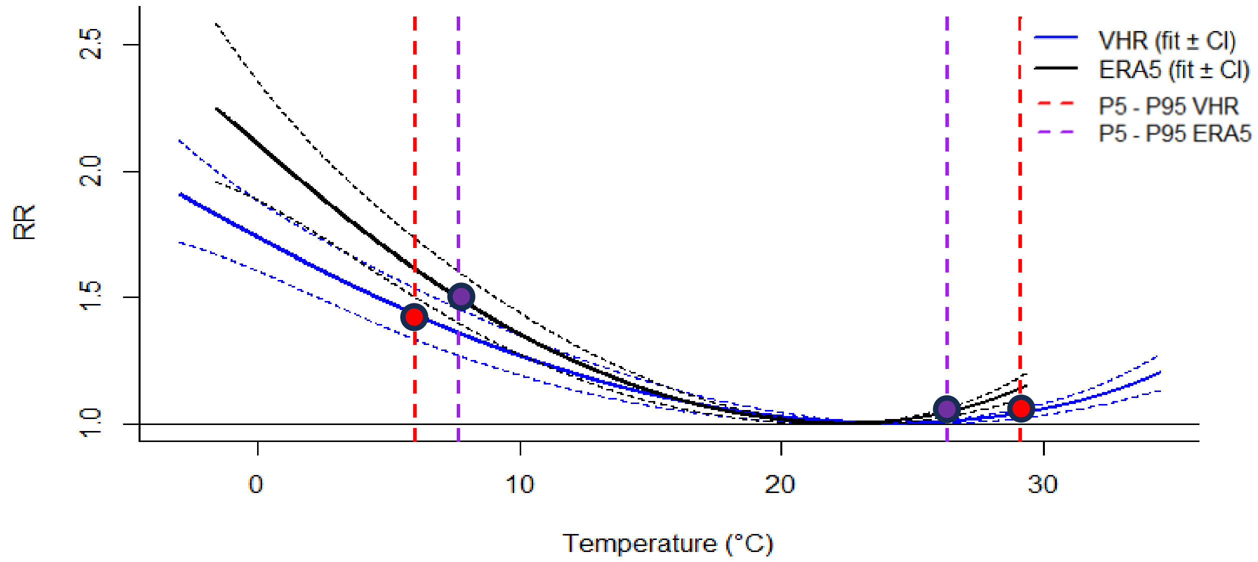
- Main exposure: daily mean 2-metre air temperature
- Confounder: Relative humidity
- Temporal adjustments: Day Of the Week (*DOW*), Seasonality, and Day Of the Year (*DOY*)

Differences in the estimates

Minimum mortality temperatures and percentiles



Extreme temperatures



VHR
MMT=24.2°C
MMP=73.85

ERA5
MMT=22.39°C
MMP=74.39

VHR
P5=6.07°C
P95=29.2°C

ERA5
P5=7.69°C
P95=26.4°C

Sensitivity analyses and model robustness checks

To ensure the robustness of our estimates, we performed a comprehensive set of sensitivity analyses covering model specification, estimation method, and temporal stability.

- Tested multiple degrees of freedom (3–6) and lag structures in the DLNM cross-basis
- Selected models based on AIC and compared RR estimates from the 5 best-fitting models
- Performed bootstrap resampling to obtain empirical confidence intervals
- Compared Generalised Additive Models (GAM) with Generalised Linear Models (GLM)
- Applied leave-one-year-out analysis to assess temporal robustness

Summary of findings and public health relevance

Key findings

- U-shaped temperature–mortality relationship
- Resolution has effect on both MMT, and extreme risk estimates

Implications

- There is an improvement in using finer data, however coarser data can be useful in low-resource settings
- Model performance may rely more on specification and covariates than on input resolution

Future directions

- Include additional environmental cofactors such as rainfall and air pollution
- Explore methods to reconstruct reliable estimates through downscaling techniques for both input and output data
- Broaden the analysis to assess the generalisability and stability of results in low-data urban settings

References

1. Alduchov, O. A., & Eskridge, R. E. (1996). Improved Magnus form approximation of saturation vapor pressure. Journal of Applied Meteorology (1988-2005), 601-609.
2. Gasparrini, A., Armstrong, B., & Kenward, M. G. (2010). Distributed lag non-linear models. Statistics in medicine, 29(21), 2224-2234.
3. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. Quarterly journal of the royal meteorological society, 146(730), 1999-2049.
4. Istituto Nazionale di Statistica (ISTAT). (2024). Codici statistici delle unità amministrative territoriali: comuni, città metropolitane, province e regioni. Retrieved from https://www.istat.it/it/archivio/6789
5. Raffa, M., Reder, A., Marras, G. F., Mancini, M., Scipione, G., Santini, M., & Mercogliano, P. (2021). VHR-REA_IT dataset: very high resolution dynamical downscaling of ERA5 reanalysis over Italy by COSMO-CLM. *Data*, *6*(8), 88.


cmcc
Centro Euro-Mediterraneo
sui Cambiamenti Climatici
www.cmcc.it