



Sharing is encouraged

Uncertainty-aware downscaling of NO₂ surface levels in urban environments

Andreu Julian-Izquierdo^{1,2}, Cristina Campos¹, Alvaro Criado^{1,3}, Cristina Carnerero¹, Albert Soret¹, Femke C Vossepoel⁴, and Jan M. Armengol^{1,2}

¹Barcelona Supercomputing Center, Earth Sciences Department, Barcelona, Spain ²Department of Fluid Mechanics, Universitat Politècnica de Catalunya, Barcelona, 08034, Spain ³Institute of Environmental Science and Technology (ICTA-UAB) ⁴Faculty of Civil Engineering and Geosciences, TU Delft, Stevinweg 1, 2628 CN Delft, Netherlands



Funded by the European Union

1 MOTIVATION

- Air pollution is a leading environmental risk:** ~7 million premature deaths annually [1]
- Urban complexity:** cities are critical hotspots with high spatial variability at fine scales
- High-resolution gap:** 68% of the Catalan population lives in urban areas where regional models do not resolve air pollutants exposure [2]
- New European Directive 2024/2881:** stricter legal thresholds and sets a strategic support for modelling
- Current limitation:** high-resolution products often lack uncertainty quantification

We aim to develop an uncertainty-aware, high-resolution framework for an integrated urban air quality and exposure assessment

2 INPUT DATA

Monitoring stations

2024 NO₂ annual mean observations from XVP-CA network

CTM

2024 NO₂ annual mean CALIOPE outputs at a 1 km x 1 km resolution

Satellite

L2 NO₂ from Sentinel-5p (TROPOMI)

High-resolution urban morphology

- Traffic intensity information
- Land use
- Terrain elevation
- Industrial point sources

3 METHODOLOGY

Stratified approach
Two different models are trained with **traffic-influenced** and **background-industrial** environments to better capture local emission patterns [3]

Regression trend

Random Forest, GBM, SVM and glmnet
Regression model and hyperparameters are selected based on LOOCV performance
Overfitting analysis performed

Residual kriging

Spatial interpolation of residuals
Observations - predictions

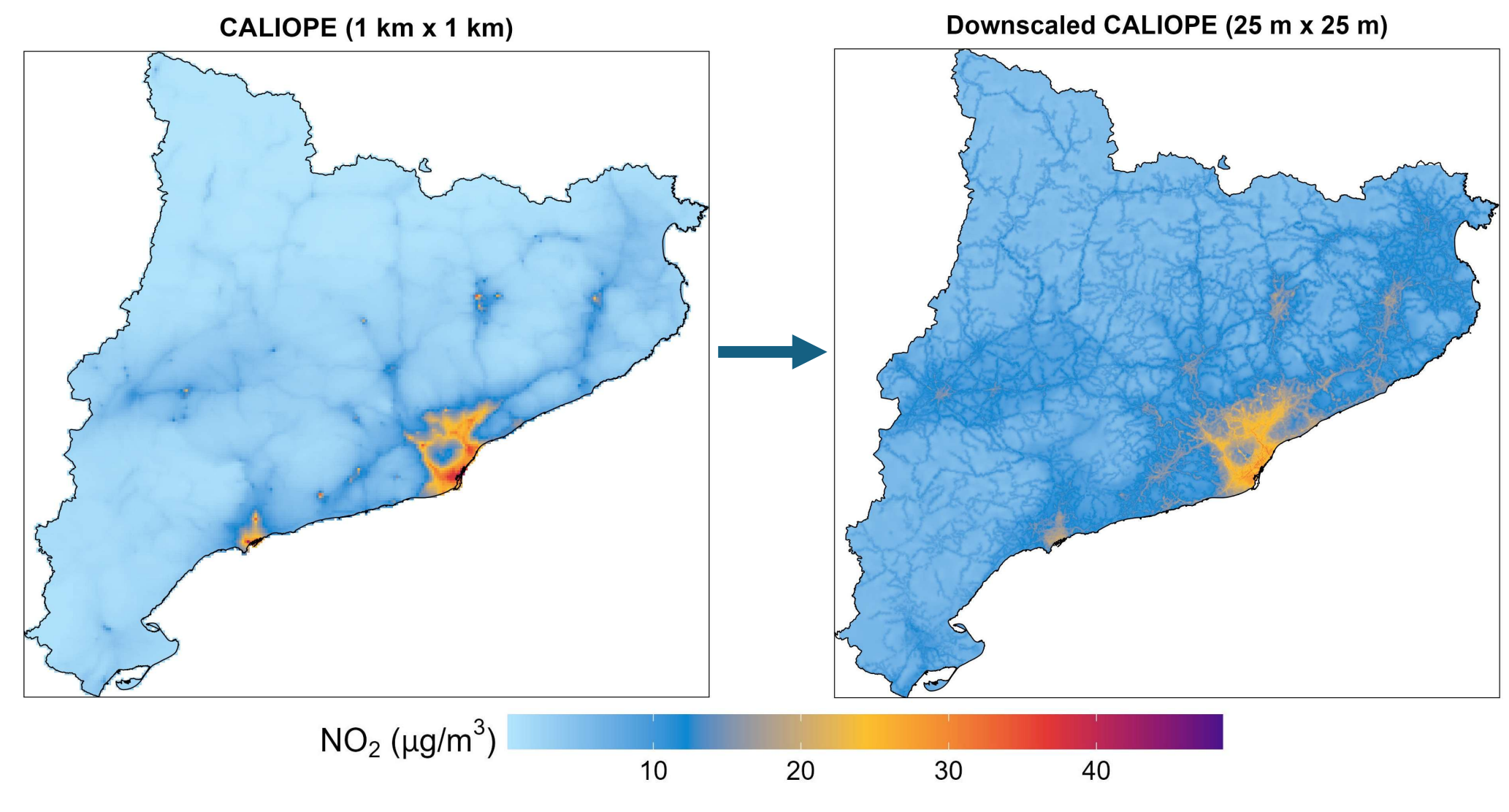
Outputs

Traffic model **Background model**

- Weighting process:
 $M = w \cdot M_{Traffic} + (1 - w) \cdot M_{Background}$
- Leave-One-Out Cross-Validation (LOOCV)

↓

Non-uniform mesh
Resolution down to 25 m near emission sources
Progressively coarser towards the regional background



Comparison of Uncertainty Quantification (UQ) methods

A. Bootstrapping

Observational error following FAIRMODE recommendations [4] + Resample input data

Iteration convergence analysis

N iterations Downscaling methodology

Distribution of probability of the downscaled NO₂ concentrations

B. Residual modelling

Random Forest → residual ~ similarity index [5] + covariates

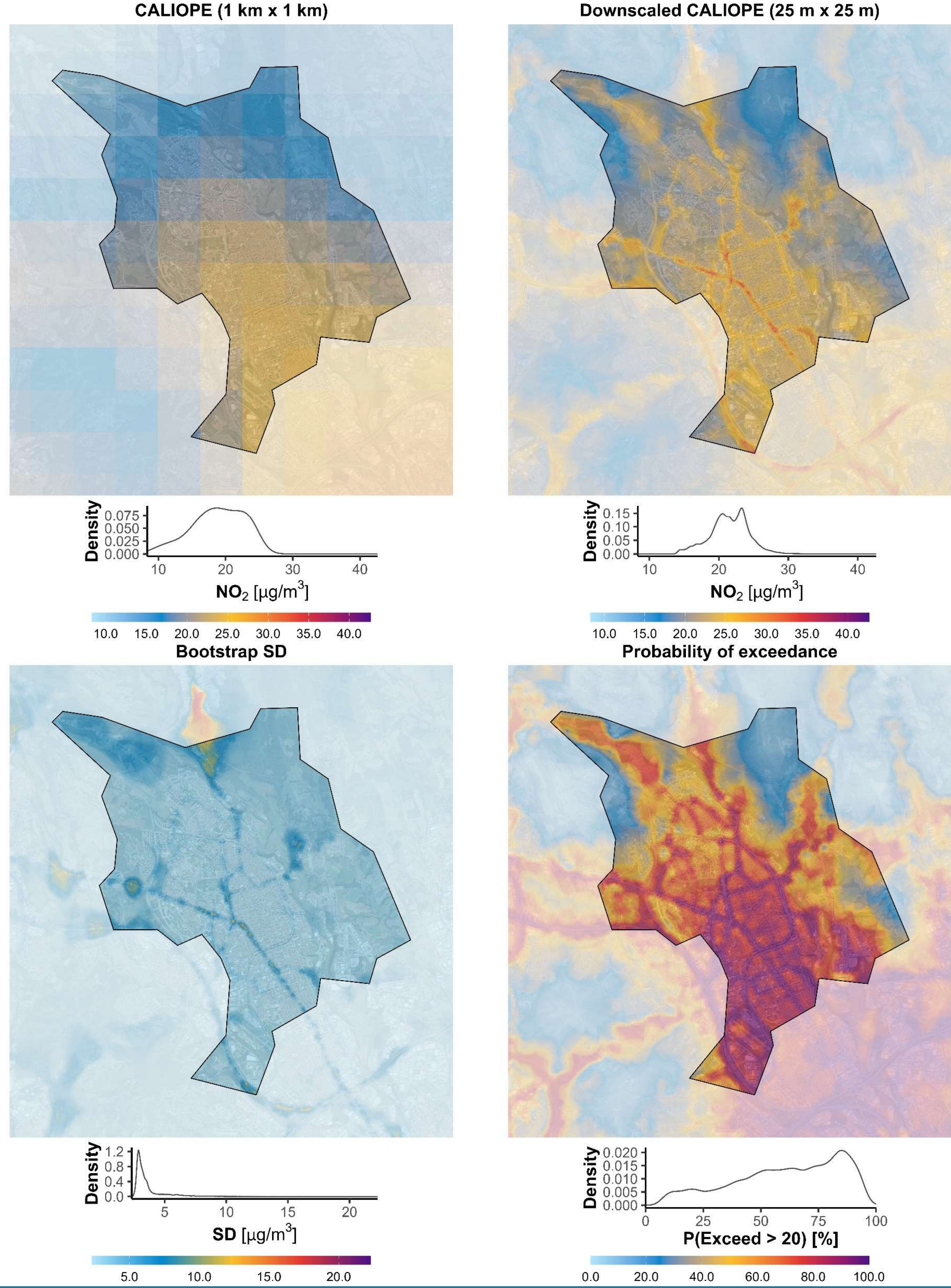
Probability of exceedance

Following the legal threshold set by the EU (20 µg/m³)

4 RESULTS

High-resolution UQ and NO₂ concentration maps for Catalonia's 947 municipalities (2024 annual mean)

The Sabadell case of study



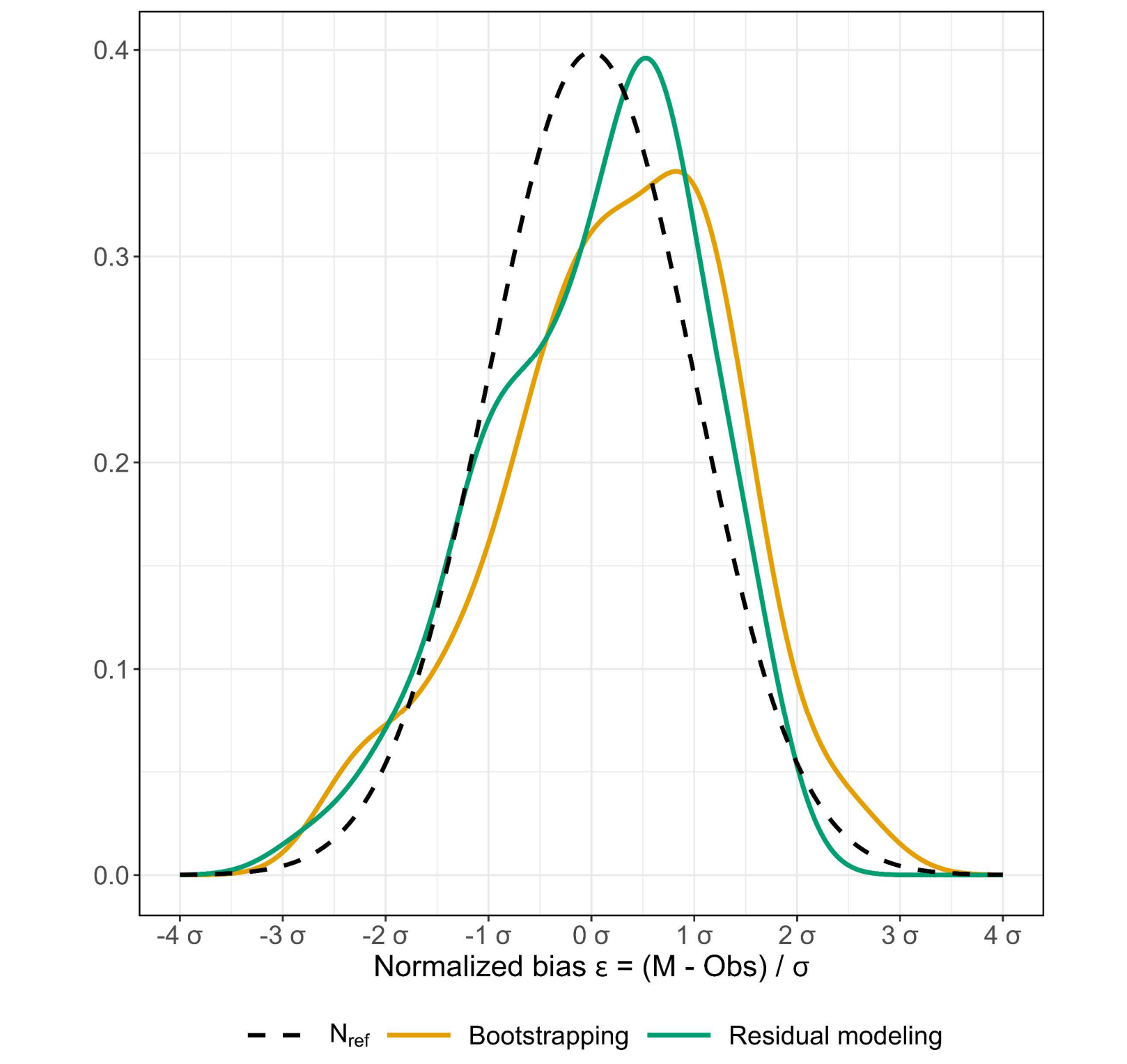
A. Global performance statistics for the NO₂ annual mean

	MB (µg/m ³)	R ²	RMSE (µg/m ³)
CALIOPE	0.31	0.69	5.24
Downscaled	0.38	0.83	3.37

B. Uncertainty Quantification

	±1σ (68 %)	±2σ (95 %)	±3σ (99,7 %)
Bootstrapping	59.4	90.6	100
Residual M.	67.2	95.3	98.4

- Both UQ techniques slightly overestimate observations.
- The residual modelling approach shows a better fit to the expected distribution, while bootstrapping yields a higher error variance. However, both correctly capture the distribution width.



5 CONCLUSIONS

- High-resolution downscaling:** the integration of satellite data and traffic information enhances previous regional models, successfully capturing street-level NO₂ patterns.
- Uncertainty quantification validation:** residual modelling performs better at monitoring sites, while bootstrapping better captures the spatial patterns of uncertainty.
- High uncertainty at NO₂ hotspots:** The absence of monitoring stations in high-pollution hotspots restricts model training on peak values and hinders direct validation of urban traffic extremes, increasing uncertainty.
- Health & exposure possibilities:** the integration of UQ and high-resolution data enables robust health studies in areas previously limited by coarse information, establishing a new standard for local-scale NO₂ exposure assessment.
- Policy framework:** Uncertainty Quantification is crucial for informed decision-making.

References

[1] World Health Organization (2021). *WHO Global Air Quality Guidelines: Particulate Matter (PM_{2.5} and PM₁₀), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. World Health Organization.

[2] Institut d'Estadística de Catalunya (Idescat) (2024). *Xifres oficials de població*. Generalitat de Catalunya.

[3] Hengl, T. (2007). *A practical guide to geostatistical mapping of environmental variables*. European Communities.

[4] Thunis, P., & Cuvelier, C. (2023). *DELTA User Guide*. European Commission, Joint Research Centre (JRC).

[5] Sahigara, F., Ballabio, D., Todeschini, R., & Consonni, V. (2013). Defining a novel k-nearest neighbours approach to assess the applicability domain of a QSAR model for reliable predictions. *Journal of Cheminformatics*, 5, 27.

Acknowledgements

This work has received funding from the European Union's Horizon Europe Research and Innovation Programme under the project "UrbanAIR" (grant agreement No. 101188131).

