Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Analysis of Air Quality Evolution in the Basque Autonomous Community Using Radial Basis Function Interpolation and Unsupervised Learning

Heber Hernández[1], Elisabete Alberdi[2], Irantzu Álvarez[3], Aitor Goti[4*], Luis Álvarez[1], Xabat Oregi[5]

[1]Civil Mining Engineering, Faculty of Engineering, Santo Tomás University, (Ingeniería Civil en Minas, Facultad de Ingeniería, Universidad Santo Tomás), Santiago, Chile
[2]Department of Applied Mathematics, University of the Basque Country (UPV/EHU), Bilbao, Spain
[3]Department of Graphic Design and Engineering Projects, University of the Basque Country (UPV/EHU), Bilbao, Spain
[4]Department of Mechanics, Design and Industrial Management, Faculty of Engineering, University of Deusto, Bilbao, Spain
[5]Department of Architecture-ENEDI Research Group, University of the Basque Country (UPV/EHU), Donostia, Spain
E-mail: aitor.goti@deusto.es

Received: 20 October 2025; **Revised:** 24 December 2025; **Accepted:** 9 January 2026

Abstract: This study presents results from a spatial modeling and analysis process of six different air pollutants measured over a ten-year period at up to forty-three monitoring stations located in the three provinces of the Basque Autonomous Community (BAC) (Spain). The main objective was to generate detailed maps showing the evolution of these pollutants that cover the entire area using geostatistical techniques. These maps are intended to serve as a basis for both short-term and medium-term environmental studies, while also examining how pollutant levels have changed before and after the COVID-19 pandemic era. Additionally, the paper explores the factors that may explain the differences observed during these two periods. To further analyze the spatial patterns, the study employs the Fuzzy C-Means clustering algorithm to partition the region into four distinct zones based on the concentrations of key pollutants: PM_{10}, NO_2, and O_3. These pollutants were selected due to their high sampling density, spatial coverage, and complementary sources (traffic emissions, combustion processes, and ozone photochemistry), making them representative indicators of the region's atmospheric state. The findings reveal significant changes in air quality during the COVID-19 pandemic, particularly in NO_2, Benzene, and CO levels, which sharply declined due to reduced vehicular traffic. However, the behavior of PM_{10} and O_3 was more complex, influenced by diverse sources and atmospheric chemistry. The reduction in NO_2 emissions during lockdowns led to a counterintuitive increase in O_3 concentrations in some areas, highlighting the non-linear responses of atmospheric systems to emission changes. This study underscores the importance of multi-faceted air quality management strategies that account for the intricate interplay of different emission sources and atmospheric processes. The insights gained from this unique period should inform the development of more effective, evidence-based air quality policies for a healthier and more sustainable future.

***Keywords*:** air quality, spatial analysis, kriging, Basque Autonomous Community (BAC)

MSC: 62P12

DOI: https://doi.org/10.37256/cm.7220268951

Abbreviation

C_6H_6	Benzene
CO	Carbon Monoxide
GDP	Gross Domestic Product
KKT	Karush-Kuhn-Tucker
NO_2	Nitrogen Dioxide
NO_x	Nitrogen Oxides
O_3	Ozone
PM_{10}	Particulate Matter less than 10 microns
RBF	Radial Basis Functions
SO_2	Sulfur Dioxide
VOCs	Volatile Organic Compounds

1. Introduction

Air composition is a sensitive indicator of human activity, reflecting the intensity and nature of societal and economic processes. The concentration of various atmospheric pollutants is directly influenced by anthropogenic activities, ranging from industrial operations to transportation patterns. Consequently, the rigorous analysis of air quality data offers a powerful tool for understanding the environmental consequences of our collective habits and the efficacy of environmental policies. In the Basque Autonomous Community (BAC), a region characterized by a dynamic mix of industrial zones and urban development, monitoring these atmospheric changes is crucial for sustainable planning and resource management. BAC operates a dense network of stations covering its three provinces, with the highest density in the Bilbao metropolitan area and other urban-industrial corridors such as Donostia-San Sebastián and Vitoria-Gasteiz [1]. Key regulated pollutants include NO, NO_2, NO_x and PM_{10}, complemented by Ozone (O_3) and other species in specific studies [2–4].

The literature on the BAC describes a region with relatively good average air quality in European terms but with important spatial contrasts, episodic events, and evolving drivers of pollution. Considering its main measured pollutants, and spatial patterns, Alberdi et al. generated continuous maps of NO, NO_2, NO_x and PM_{10} across the BAC using kriging-based spatial interpolation, revealing hot spots closely aligned with traffic and industrial areas and lower concentrations in rural and mountainous zones [2]. This spatial heterogeneity is reinforced by a regional-scale deposition study showing that trees remove substantial amounts of NO_2, O_3, SO_2, CO and PM_{10}, especially in forested rural areas, even though most health impacts occur in cities [5]. Coniferous forests, covering about 25% of the territory, account for roughly 21% of modeled pollutant removal because of high canopy density and long leaf lifespan [5]. Regarding evolution, long-term Spain-wide analyses (including the BAC) show marked declines in PM_{10}, $PM_{2.5}$, SO_2 and CO, and moderate decreases in NO_2/NO_x since the early 2000s, largely driven by EU industrial and power-plant directives and vehicle standards [6]. Ozone, however, has remained stable at rural sites and increased at many urban and industrial locations, reflecting complex photochemistry and limited NO_x reductions [6]. A national 25-year analysis further suggests that recent weather changes have partially offset emission-based improvements, especially for PM_{10} ("weather penalty") [7]. For the BAC specifically, several strands of evidence highlight the role of episodes and external contributions: 1) Two major ozone episodes in 2001 and 2003 exceeded population information thresholds at urban and rural sites [8]. Mesoscale modeling showed that blocking anticyclones favored long-range import of pollution from Atlantic France, the western Mediterranean, southern France, the Ebro Valley and occasionally Madrid, superimposed on local emissions [8]. 2) During the COVID-19 lockdown (March-June 2020), regional road traffic dropped by roughly 50% while industrial activity only fell about 20%. Urban NO_x decreased by around 45%, but PM_{10} declined by only 10-21%, and O_3 decreased modestly or changed heterogeneously, with several exceedances of World Health Organization (WHO) guidelines linked to regional transport under anticyclonic conditions [3]. This underscores that even under strong traffic reductions, transboundary aerosol and ozone episodes persist. At the sub-regional scale, monitoring in industrial Gipuzkoa (2017-2020) documented baseline levels of Polychlorinated Dibenzo-p-Dioxins and furans and dioxin-like Dioxins and Polychlorinated Biphenyls

(PCBs) that are similar to or lower than other European industrial areas, but with clear seasonal (higher in autumn-winter), weekly (higher on weekdays) and spatial contrasts, and higher levels near planned energy recovery infrastructure [9]. Finally, in terms of eco-benefits related to air quality, the dry deposition study estimates that trees in the BAC remove on average about 12-13 $kg \cdot ha^{-1} \cdot yr^{-1}$ of O_3 and PM_{10}, and smaller amounts of NO_2, SO_2 and CO, with a monetized health benefit of ~ €60 million per year (~ 0.09% of regional GDP in 2016) [5]. Despite these benefits, the authors stress that most pollution exposure-and thus health burden-remains concentrated in urban areas, so forest sinks cannot substitute for emission reduction policies [5].

Summarizing, literature on the BAC describes a region with relatively good average air quality in European terms but with important spatial contrasts, episodic events, and evolving drivers of pollution. The BAC operates a dense network covering urban-industrial corridors. Long-term trends show declines in particulate matter and primary pollutants, though ozone remains a challenge linked to transport and photochemistry. Recent studies highlight the role of forests in pollutant removal and the impact of episodes like COVID-19 or blocking anticyclones on regional air quality. Nevertheless, remaining gaps include detailed long-term trend analyses specific to the BAC, and explicit attribution of changes to sectoral or societal habits or reasons.

With these context conditions, this research is a direct continuation and expansion of our previous work, which pioneered the use of geostatistical methods to map the spatial distribution of key pollutants like NO_x and PM_{10} in the region [2]. That study successfully demonstrated the utility of radial basis functions for creating a detailed snapshot of air quality, and serves as a reference point for future studies. This current research is aimed to build upon that foundation to explore the temporal evolution of air quality, particularly in light of the significant societal shifts of the last several years.

The global COVID-19 pandemic, and the associated public measures, triggered an unprecedented disruption in daily life, fundamentally altering patterns of mobility, commerce, and industrial production. This period serves as a unique, large-scale case study, providing a rare opportunity to assess how such drastic changes in human behavior are reflected in the atmospheric environment. By analyzing air quality data before, during, and after this period, we can gain valuable insights into the potential for lasting changes in societal habits. This paper extends our prior analysis in three critical dimensions:

• Temporal Scope: We broaden the study from a single snapshot to a full decade (2015-2024). This allows for a robust comparison between the five years of activity pre-COVID and the five years that encompass the pandemic and the subsequent "new normal".

• Pollutant Range: The analysis is expanded from two pollutants to six: Benzene, NO_2, O_3, PM_{10}, SO_2 and CO. This provides a more comprehensive and nuanced view of the atmospheric response to changes in different societal sectors.

• Interpretive Focus: We shift from a descriptive mapping exercise to an analytical one. The core objective is to use the observed trends in air quality as a proxy to investigate the evolution of societal habits and to explore whether the lessons learned during the pandemic can inform future strategies for sustainable development.

Thus, the paper is structured as follows. Section 2 presents the methodology followed. Section 3 presents the application case, and the obtained results. Section 4 the discussion of the key insights, including the theoretical reasons leading to these results. Finally, section 5 concludes the paper by summarizing the contributions and outlining directions for future research.

2. Methodology

The study employs an integrated approach combining spatial interpolation, unsupervised learning, and statistical analysis to characterize air pollution in the Basque Country during the period 2015-2024. The combination of approaches of this type was selected as, evidently, integrated frameworks consistently outperform single-method approaches in prediction accuracy, spatial coverage [10], and interpretability, enabling robust handling of missing data, spatial heterogeneity, and complex temporal trends [11], being increasingly adopted in both academic research and practical air quality management [12]. The methodology is structured in interconnected stages (Figure 1), designed to address the spatiotemporal heterogeneity of the data and the uncertainty in delimiting pollution zones.

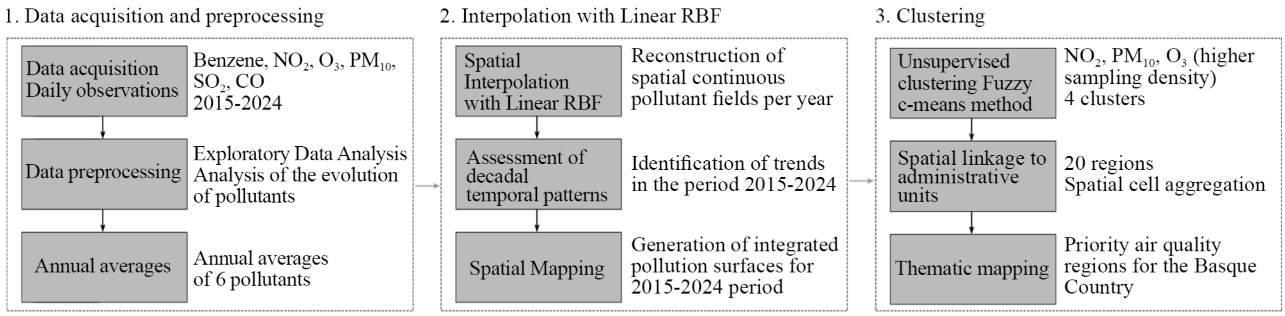


Figure 1. Workflow diagram ofthe analysis of pollutants in the Basque Country

The methodology initiates with the acquisition of daily air quality time-series data from a discrete network of monitoring stations, which is then temporally aggregated into annual mean vectors for the 2015-2024 period. These data have been obtained from the statistics published by the Environment Section of the Department of Industry, Energy Transition and Sustainability of the Basque Government, and is open access on the Open Data Euskadi portal.

Over the ten-year study period, air quality data from a total of 69 monitoring stations of the Basque Air Quality Control Network were analyzed, although the number of active stations varied over time, with 55 stations currently in operation. The stations are located in rural, urban and suburban environments, with a higher density in urban areas, reflecting the primary objective of the network to monitor population exposure to traffic-related, industrial and background air pollution. Not all pollutants are measured at all stations; while some pollutants are monitored across the entire network, others are measured only at specific station types (e.g., industrial or traffic-oriented sites), leading to an uneven spatial sampling that is inherent to the regulatory design of the monitoring network and directly influences the spatial representativeness of the interpolated surfaces.

A spatial interpolation using a linear Radial Basis Function (RBF) model [13–15] is applied to these annual vectors to reconstruct continuous pollutant concentration fields, enabling the assessment of decadal temporal trends and anomalies [2]. These annual fields are subsequently integrated into a single historical average surface, also via RBF interpolation, which represents the mean pollutant distribution over the entire decade. An unsupervised classification is then performed on this integrated surface using a fuzzy c-means clustering algorithm [16, 17] to partition the spatial domain into a set of distinct pollution zones, each characterized by a unique pollutant profile and environmental influences. Finally, these fuzzy-segmented zones are spatially linked to administrative units through a process of spatial cell aggregation to determine predominant cluster assignments, culminating in the generation of a thematic map that delineates priority air quality regions for the Basque Country.

The main points of the methodological section are detailed in the following subsections:

2.1 *Data acquisition and preprocessing*

Daily raw data on air pollutant concentrations were obtained from a network of monitoring stations with a non-uniform spatial distribution and from official records in the Basque Country (Spain).

2.2 *Problem formulation*

Given a set of n scattered points $\left(x_i \in R^d,\ i = 1,\ \ldots,\ n\right)$ with associated values $\left(f_i \in R,\ i = 1,\ \ldots,\ n\right)$, an interpolant function has to be find $\left(s : R^d \rightarrow R\right)$ in such a way that $s\left(x_i\right) = f_i,\ i = 1,\ \ldots,\ n$.

Function $s(x)$ is built as a linear combination of RBF centered at the nodes x_i:

$$s(x) = \sum_{=1}^{n} \lambda_i \phi\left(||x - x_i||\right) \tag{1}$$

being $\phi : R^+ \to R$ a linear radial basis function. In our case, $||\cdot||$ whe Euclidean norm, will be used and λ_i cthe oefficients are determined by solving the linear system that comes from:

$$Z(x_k) = \sum_{=1}^{n} \lambda_i \phi\left(||x - x_i||\right), \ 1 \le k \le n. \tag{2}$$

For the linear interpolation, the RBF $\phi(r) = r$ has been chosen because it guarantees simplicity and continuity.

The interpolation matrix ($A \in R^{nxn}$), with elements $A_{ij} = \phi\left(\left|\left|x_i - x_j\right|\right|\right)$, results in a linear system $A\lambda = f$ where $\lambda = (\lambda_1, \ldots, \lambda_n)^\top$ and $f = (f_1, \ldots, f_n)^\top$. The non-singularity of A is verified by means of the definite positivity condition of ϕ. To ensure numerical stability, Lower-Upper (LU) decomposition is applied.

2.3 *Fuzzy clustering c-means*

Optimization approach

Given a data set $x_i \in R^d, \ i = 1, \ldots, n$ and a number of clusters c, the objective is to minimize the cost function:

$$J_m(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^m ||x_i - v_j||^2, \tag{3}$$

where $U = (u_{ij})$ is the fuzzy membership matrix being $u_{ij} \in [0, 1]$ and $v_j \in R^d, \ j = 1, \ldots c$ are the centroids, and $(m > 1)$ is the defuzzification parameter.

Optimality conditions

The updates of U and V are derived by Karush-Kuhn-Tucker (KKT) conditions:

1. **Updating belongings:**

$$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{||x_i - v_j||}{||x_i - v_k||} \right)^{\frac{2}{m-1}} \right)^{-1}. \tag{4}$$

2. **Updating centroids:**

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m}. \tag{5}$$

The algorithm iterates until $\left|\left|V^{(t+1)} - V^{(t)}\right|\right| < \varepsilon$ is verified, given a tolerance $\varepsilon > 0$. Convergence to a local minimum is guaranteed by Bezdek's theorem.

Parameters c and m are selected. The number of clusters c is selected using the fuzzy partition index $FP = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^2$, maximizing the partition clarity. And the parameter m is chosen considering that it typically takes values in $m \in [1.5, 3.0]$, being, $m = 2$ as a standard for balance between ddefuzzificationand stability.

Afterwards, the validation process is followed. The quality of the clustering is evaluated using the Xie-Beni index [18]:

$$XB = \frac{J_m}{n \cdot \left|\left|v_j - v_k\right|\right|^2}, \tag{6}$$

where low values indicate high intra-cluster cohesion and inter-cluster separation.

2.4 *Geospatial integration and final mapping*

The clusters are linked to the regions of the Basque Country by assigning each administrative unit the predominant pollution level, calculated as the spatial mode of the cells within its boundaries. The final product consists of a thematic map that classifies the regions into a set of categories: "High", "Medium", "Low", and "Very Low", providing a visual tool for environmental management.

3. Application

A database containing daily observations of air pollutants recorded in the Basque Country over the past 10 years (2015-2024) is available,at the Environment Section of the Department of Industry, Energy Transition and Sustainability of the Basque Government, and wisopen access on the Open Data Euskadi portal [19]. The data come from monitoring stations distributed throughout the region. For this study, the following pollutants and the number of stations available for each are considered as main variables: benzene (7 stations), NO_2 (43), O_3 (33), PM_{10} (40), SO_2 (29), and CO (16). Although other pollutant measurements exist, they have not been included in the study because they are only available at specific locations and are therefore not representative of the region.

Table 1 describes the following statistics for each pollutant, considering all the stations that measured them shown in it: annual average, standard deviation, minimum, maximum and first, second, and third quartiles.

Table 1. Descriptive statistics by pollutant and year (2015-2024)

Year	Pollutant	Count	Mean	Std	Min	0.25	0.50	0.75	Max
2015	Benzene (μg/m^3)	7	0.55	0.19	0.33	0.44	0.51	0.63	0.89
2015	NO_2 (μg/m^3)	43	21.75	9.32	5.41	15.18	22.13	26.58	54.52
2015	O_3 (μg/m^3)	33	47.88	14.80	15.00	38.00	46.32	58.22	79.60
2015	PM_{10} (μg/m^3)	40	18.13	3.75	11.52	15.84	18.32	20.48	26.33
2015	SO_2 (μg/m^3)	29	5.71	1.26	3.16	4.97	5.45	6.75	8.56
2015	CO (mg/m^3)	16	0.26	0.04	0.17	0.24	0.26	0.28	0.33
2016	Benzene (μg/m^3)	7	0.54	0.26	0.17	0.40	0.52	0.68	0.93
2016	NO_2 (μg/m^3)	43	19.51	8.24	2.90	14.08	19.40	26.24	41.63
2016	O_3 (μg/m^3)	29	52.23	12.50	32.36	44.99	46.86	60.05	76.04
2016	PM_{10} (μg/m^3)	40	15.88	3.12	9.71	13.75	16.31	18.05	21.65
2016	SO_2 (μg/m^3)	29	4.43	1.49	1.32	3.57	4.73	5.27	7.25
2016	CO (mg/m^3)	16	0.24	0.05	0.14	0.21	0.24	0.28	0.31
2017	Benzene (μg/m^3)	7	0.65	0.36	0.24	0.42	0.57	0.85	1.20
2017	NO_2 (μg/m^3)	43	19.73	8.50	3.10	14.12	19.95	25.48	43.02
2017	O_3 (μg/m^3)	28	50.83	14.27	20.75	40.26	47.95	61.51	77.26
2017	PM_{10} (μg/m^3)	40	16.77	3.46	9.42	14.21	17.06	18.90	24.88
2017	SO_2 (μg/m^3)	28	4.28	1.38	2.00	3.10	4.36	5.45	6.60
2017	CO (mg/m^3)	16	0.26	0.05	0.19	0.23	0.27	0.29	0.34

Table 1. (cont.)

Year	Pollutant	Count	Mean	Std	Min	0.25	0.50	0.75	Max
2018	Benzene ($\mu g/m^3$)	8	0.60	0.27	0.28	0.37	0.59	0.82	0.93
2018	NO_2 ($\mu g/m^3$)	46	17.82	7.82	3.26	12.30	18.10	23.83	38.99
2018	O_3 ($\mu g/m^3$)	29	54.35	12.52	38.06	44.27	52.18	64.29	77.84
2018	PM_{10} ($\mu g/m^3$)	44	15.38	3.37	9.36	13.37	14.62	17.38	22.78
2018	SO_2 ($\mu g/m^3$)	27	3.96	1.66	1.22	2.96	3.60	5.22	8.62
2018	CO (mg/m^3)	16	0.25	0.04	0.20	0.22	0.25	0.27	0.31
2019	Benzene ($\mu g/m^3$)	11	0.39	0.22	0.12	0.26	0.38	0.46	0.90
2019	NO_2 ($\mu g/m^3$)	48	16.86	7.33	2.36	11.36	16.30	22.45	35.84
2019	O_3 ($\mu g/m^3$)	32	55.44	12.98	38.71	45.06	51.97	64.12	81.42
2019	PM_{10} ($\mu g/m^3$)	46	15.14	3.43	8.51	12.74	14.46	16.84	22.71
2019	SO_2 ($\mu g/m^3$)	29	4.26	1.73	1.38	2.71	4.21	5.58	7.75
2019	CO (mg/m^3)	16	0.25	0.05	0.16	0.24	0.25	0.27	0.34
2020	Benzene ($\mu g/m^3$)	11	0.38	0.22	0.16	0.26	0.29	0.46	0.87
2020	NO_2 ($\mu g/m^3$)	48	13.64	5.63	2.14	9.52	13.88	17.14	26.62
2020	O_3 ($\mu g/m^3$)	32	52.59	10.13	36.66	43.68	50.48	60.51	71.99
2020	PM_{10} ($\mu g/m^3$)	46	14.67	3.23	7.44	12.30	14.23	16.80	21.18
2020	SO_2 ($\mu g/m^3$)	29	4.03	1.94	1.38	2.47	3.58	4.92	10.45
2020	CO (mg/m^3)	16	0.25	0.06	0.15	0.22	0.24	0.28	0.36
2021	Benzene ($\mu g/m^3$)	13	0.40	0.22	0.21	0.26	0.32	0.39	1.04
2021	NO_2 ($\mu g/m^3$)	49	14.09	5.99	1.70	9.53	14.27	18.45	28.79
2021	O_3 ($\mu g/m^3$)	33	53.62	10.49	37.26	44.86	52.36	63.50	72.55
2021	PM_{10} ($\mu g/m^3$)	47	15.39	3.09	8.22	12.73	15.77	17.25	22.69
2021	SO_2 ($\mu g/m^3$)	30	4.10	1.96	1.04	2.71	4.00	5.16	11.69
2021	CO (mg/m^3)	16	0.28	0.07	0.18	0.23	0.26	0.30	0.45
2022	Benzene ($\mu g/m^3$)	13	0.42	0.23	0.25	0.31	0.35	0.41	0.96
2022	NO_2 ($\mu g/m^3$)	49	14.26	6.09	2.50	9.70	14.06	18.42	29.05
2022	O_3 ($\mu g/m^3$)	33	55.17	10.95	38.77	45.81	53.12	62.38	76.52
2022	PM_{10} ($\mu g/m^3$)	47	17.00	3.25	9.58	14.88	16.80	19.02	25.27
2022	SO_2 ($\mu g/m^3$)	30	4.28	1.74	1.54	3.02	4.08	5.36	9.29
2022	CO (mg/m^3)	16	0.27	0.06	0.17	0.22	0.25	0.32	0.38
2023	Benzene ($\mu g/m^3$)	13	0.39	0.21	0.20	0.26	0.34	0.40	0.95
2023	NO_2 ($\mu g/m^3$)	49	13.72	6.02	2.31	9.64	13.47	17.69	29.14
2023	O_3 ($\mu g/m^3$)	33	55.21	11.12	35.78	46.87	53.80	61.80	78.81
2023	PM_{10} ($\mu g/m^3$)	47	15.53	3.08	8.17	13.58	15.31	17.55	22.92
2023	SO_2 ($\mu g/m^3$)	30	4.98	2.09	1.98	3.54	4.67	6.01	12.25
2023	CO (mg/m^3)	16	0.27	0.05	0.18	0.23	0.27	0.29	0.35
2024	Benzene ($\mu g/m^3$)	12	0.38	0.25	0.09	0.26	0.31	0.38	0.95
2024	NO_2 ($\mu g/m^3$)	49	12.63	5.30	2.66	9.06	12.07	16.43	26.09
2024	O_3 ($\mu g/m^3$)	33	53.40	9.58	38.80	45.17	53.43	59.97	73.40
2024	PM_{10} ($\mu g/m^3$)	47	14.18	3.20	7.17	12.01	14.03	15.74	20.93
2024	SO_2 ($\mu g/m^3$)	30	4.28	1.64	2.27	3.20	4.05	4.96	10.84
2024	CO (mg/m^3)	16	0.27	0.04	0.21	0.25	0.27	0.29	0.37

Primary combustion pollutants (benzene, NO_2, PM_{10}, SO_2) show strong positive correlations with each other (see Figure 2), indicating common emission sources. O_3 has negative correlations with these primary pollutants, as would be expected due to chemical processes in the atmosphere.

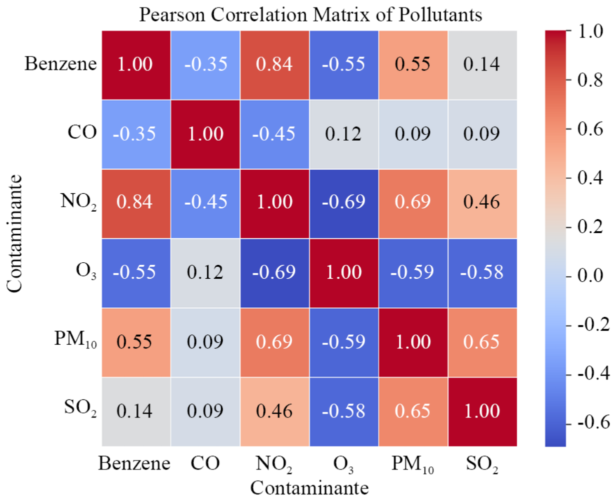
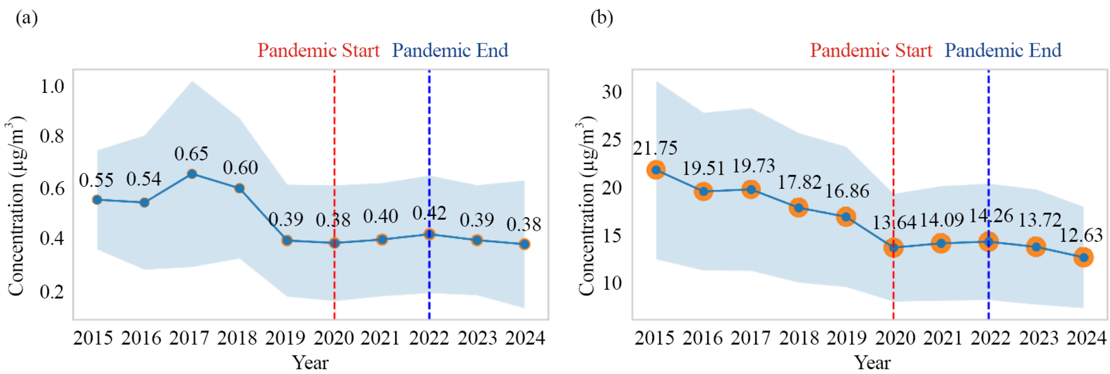


Figure 2. Pearson correlation matrix of pollutants

The pandemic had a clear and differential impact on pollutants: during 2020, a notable reduction was observed in indicators associated with traffic, mainly in NO_2 (with an average of 13.64 µg/m^3 in 2020 and 16.86-21.75 µg/m^3 in 2015-2019) and PM_{10} (with an average of 14.67 µg/m^3 in 2020 and 15.1-18.1 µg/m^3 in pre-pandemic stage). This decline was the most pronounced between 2019 and 2020, consistent with the drastic reduction in mobility. In contrast, O_3 maintained high and relatively stable averages (≈ 52-55 µg/m^3) and did not show a sustained decline during 2020; this is consistent with the complex effects of the decline in NO_x on ozone photochemistry. Pollutants associated with industrial processes or point sources (SO_2 and benzene) showed smaller variations in the average (SO_2 went from 4-5 µg/m^3 in recent years; benzene was already low and stable) although isolated highs appeared in some years. In the post-pandemic period (2021-2024), a partial recovery of NO_2 and PM_{10} can be seen compared to 2020, but the averages and maximums did not return to the highest levels observed in 2015-2018 (for example, NO_2 averaged 12.6 µg/m^3 in 2024 vs. 21.8 µg/m^3 in 2015), which suggests an incomplete recovery of activity or structural changes in mobility and emissions (teleworking, cleaner fleets, urban policies). In summary, he pandemic produced a pronounced temporary reduction in traffic-related pollutants, smaller and more variable effects on SO_2 and benzene, and persistent or even opposite behavior in O_3; The post-pandemic period shows partial recovery but without returning to pre-COVID peaks, indicating possible lasting changes in emissions patterns (see Figure 3).



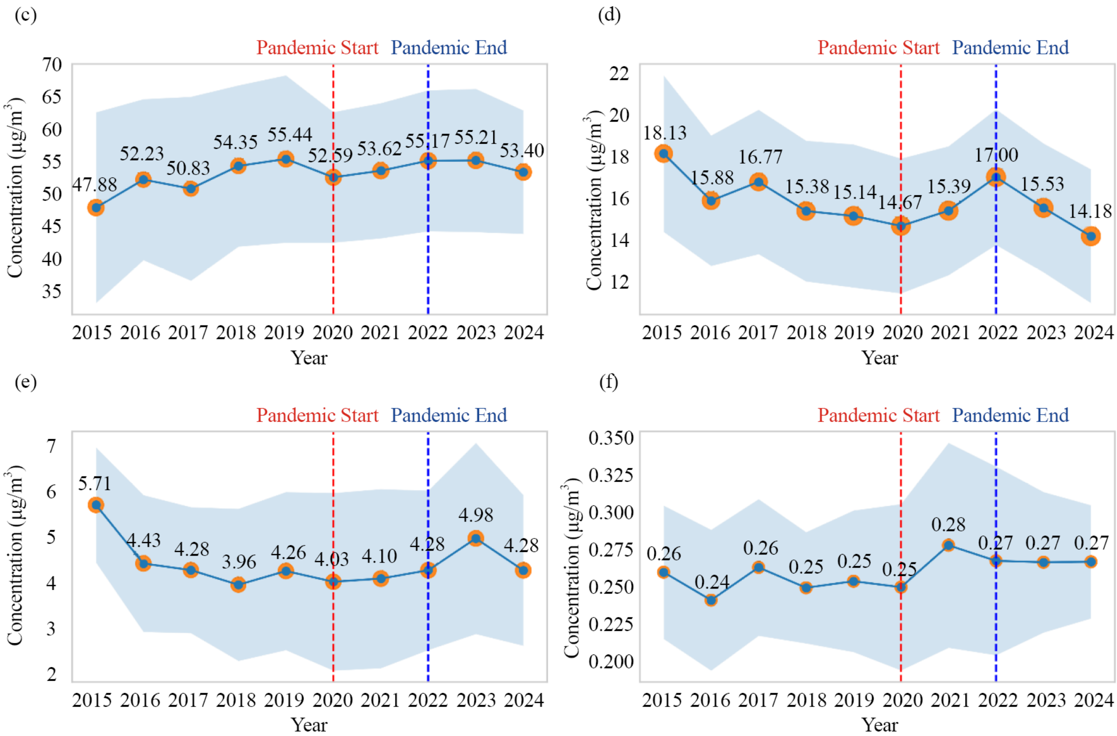


Figure 3. Benzene (a), NO_2 (b), O_3 (c), PM_{10} (d), SO_2 (e) and CO (f)

Based on the annual averages of pollutants by season, an interpolation is performed over the extension of the Basque Country (see Figures 4-9). The spatial interpolation was performed using a RBF with a linear kernel. The mapping was projected in the ETRS89/UTM zone 30 N (EPSG:25830) coordinate reference system. The interpolation grid covered the study area defined by the bounding box [X: 481,073-592,449; Y: 4,707,389-4,805,932], with a total dimension of 400 × 400 cells. This configuration resulted in a spatial resolution (cell size) of approximately 279 meters (East-West) by 247 meters (North-South).

To better understand the spatial distribution of air pollution and identify regions with similar environmental characteristics, the study focused on three key pollutants, PM_{10}, NO_2, and O_3, selected from the six originally analyzed. These pollutants were chosen due to their high sampling density, spatial coverage, and complementary sources (traffic emissions, combustion processes, and ozone photochemistry), making them representative indicators of the region's atmospheric state. Using these variables, the Basque Country was partitioned into four distinct zones, each reflecting relatively unique pollution profiles and behaviors.

The clustering process used is the fuzzy c-means method, in which certain key hyperparameters are selected for configuration and analysis. A range of k values, from 2 to 10, is tested to determine the optimal number of clusters. Furthermore, the diffusivity parameter (m) is set to 2, which allows controlling the degree of overlap and fuzziness between groups, facilitating smoother membership in the partitions. The algorithm is configured with a convergence error of 0.005 and a maximum of 1,000 iterations, ensuring sufficient convergence accuracy. For each value of k, the objective or target function (JM), which measures the clustering quality, is calculated, and these values are analyzed to identify the most appropriate structure in the data.

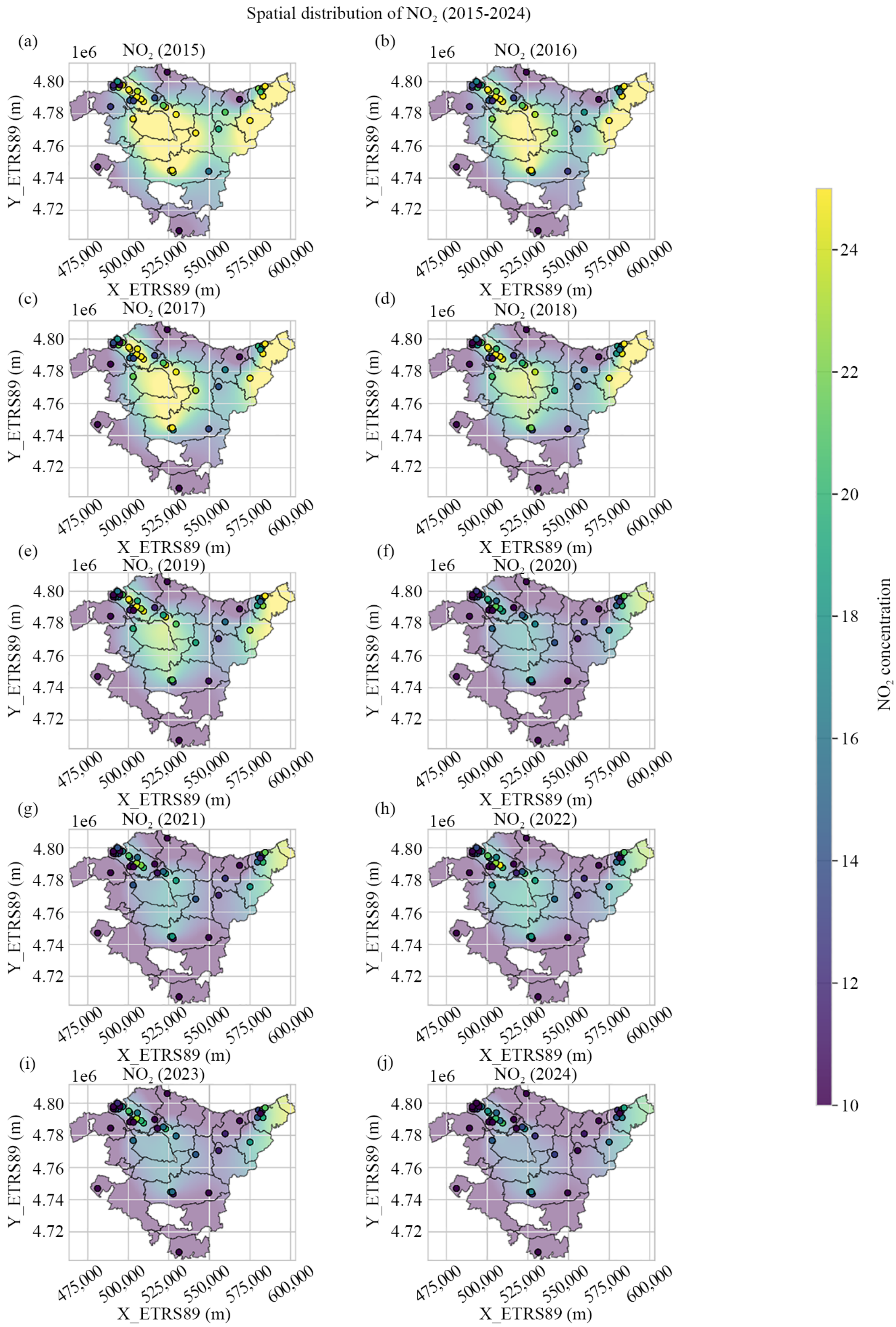


Figure 4. Spatial distribution of NO_2 (2015-2024)

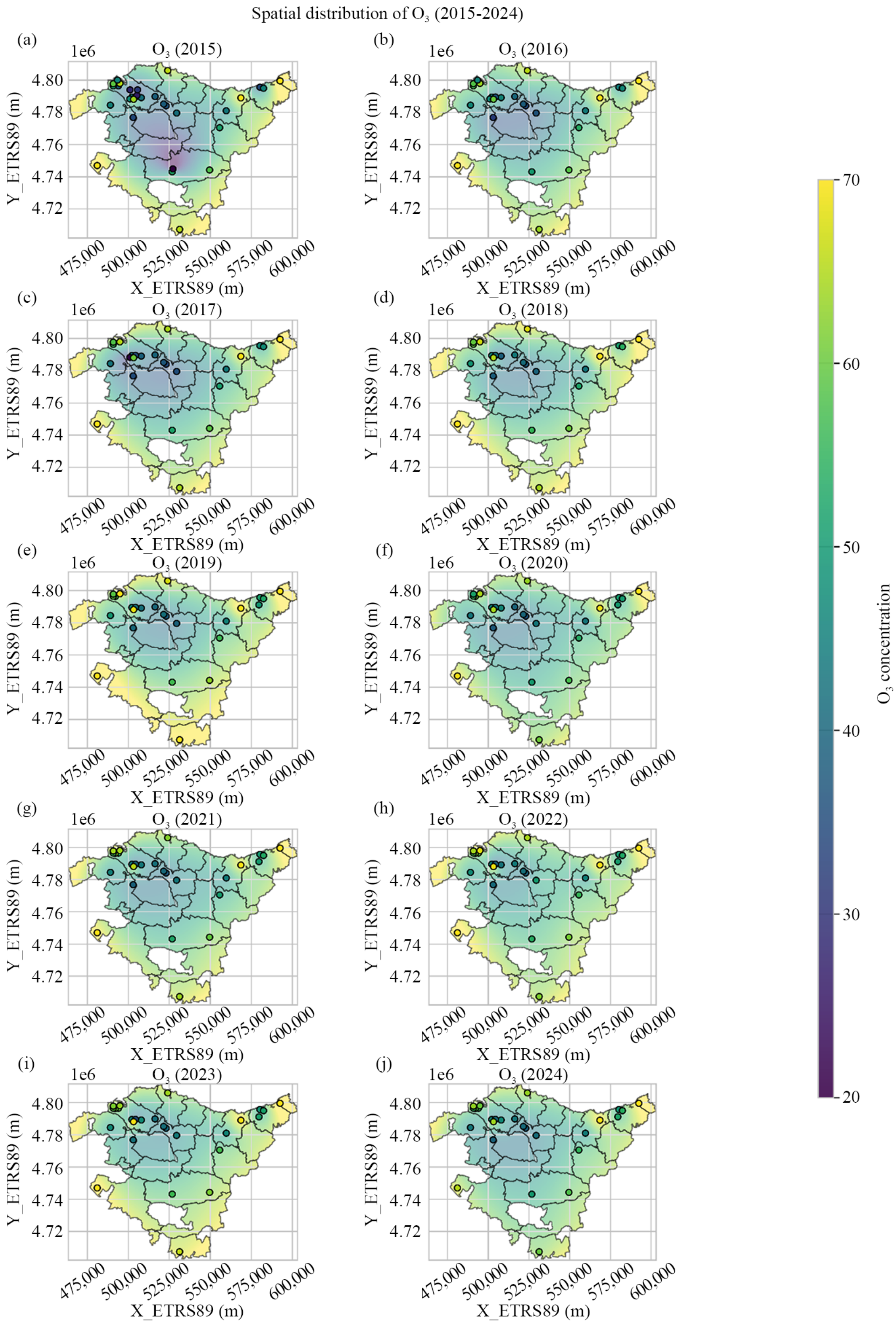


Figure 5. Spatial distribution of O_3 (2015-2024)

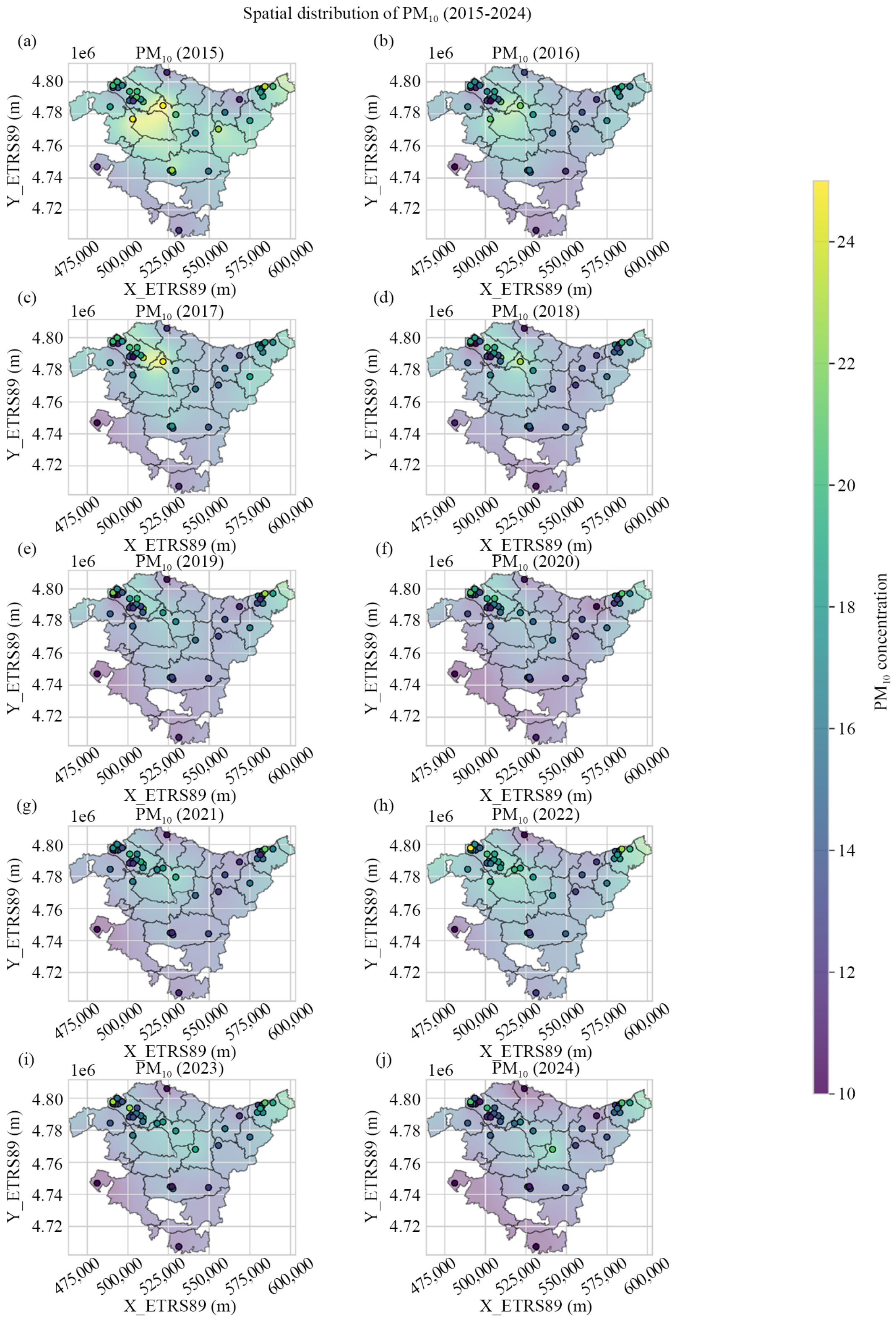


Figure 6. Spatial distribution of PM_{10} (2015-2024)

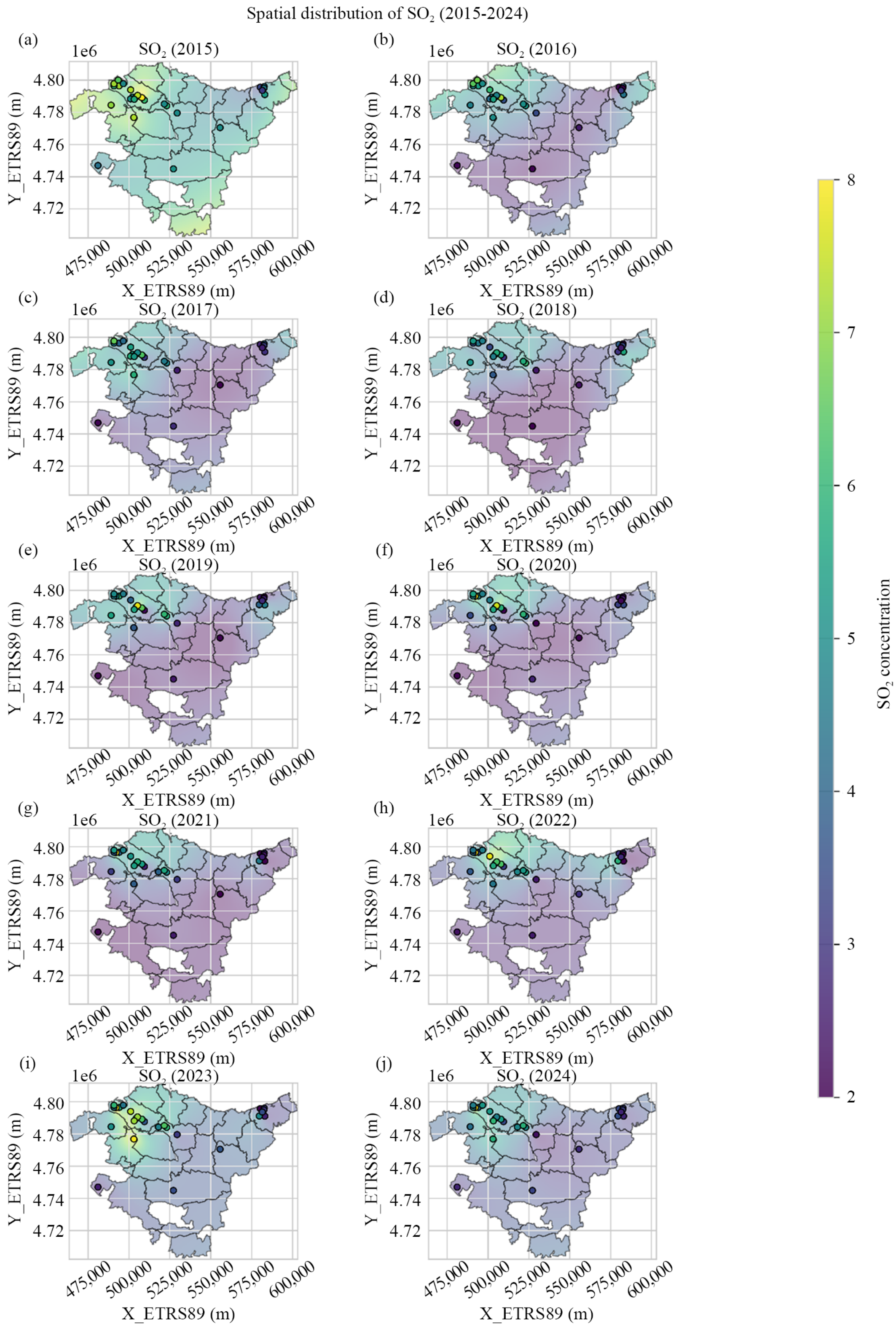


Figure 7. Spatial distribution of SO_2 (2015-2024)

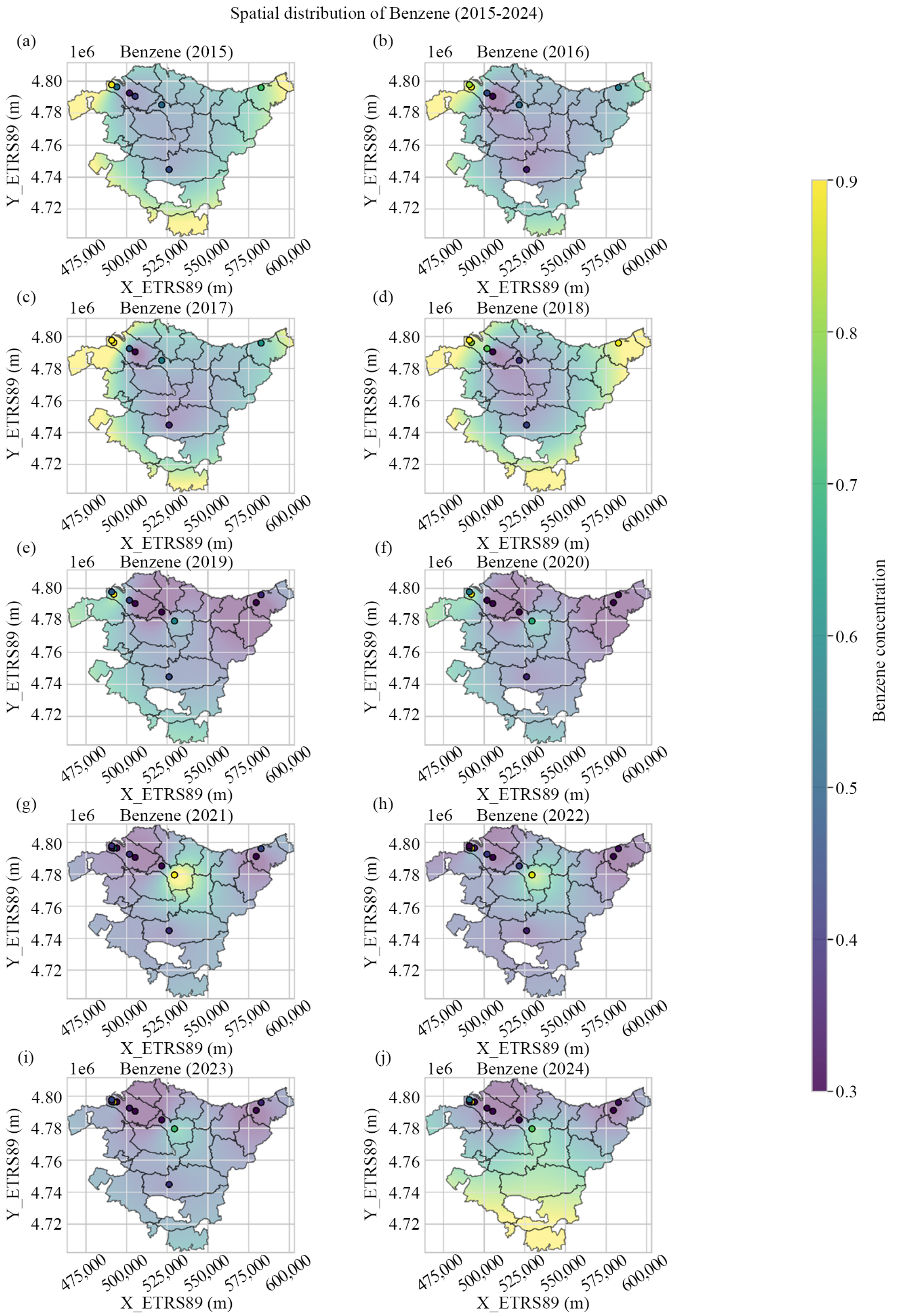


Figure 8. Spatial distribution of benzene (2015-2024)

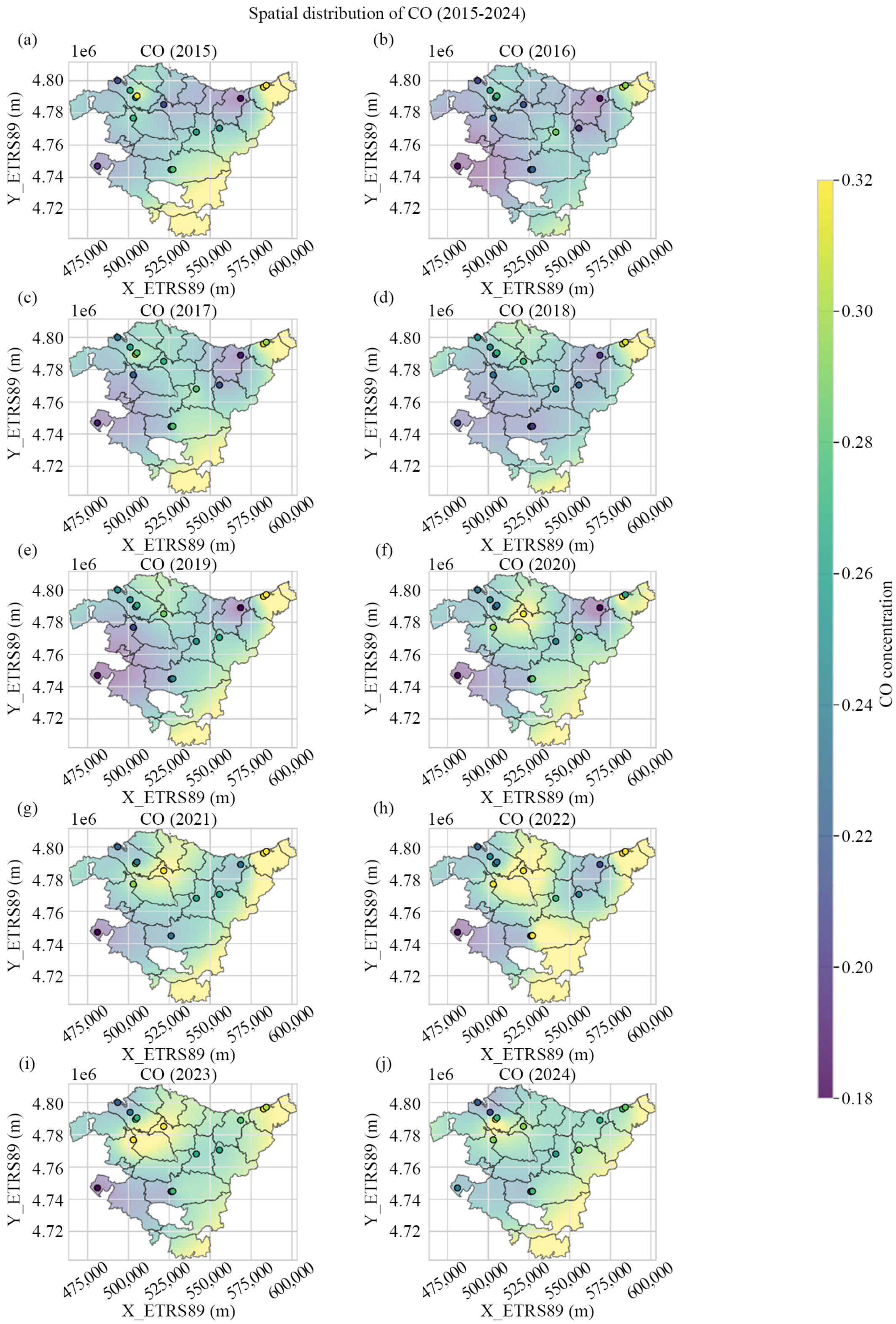


Figure 9. Spatial distribution of CO (2015-2024)

The areas identified through fuzzy cluster analysis, using levels of the pollutants NO_2, O_3, and PM_{10}, reflect different air pollution profiles in the Basque Country (see Figure 10). The classification into four zones is based on the interpretation of the results of the elbow method and silhouette analysis (see Figures 11 and 12), which is further validated by the fuzzy Partition Coefficient (PC = 0.57) and the Xie-Beni index (XB = 0.195).

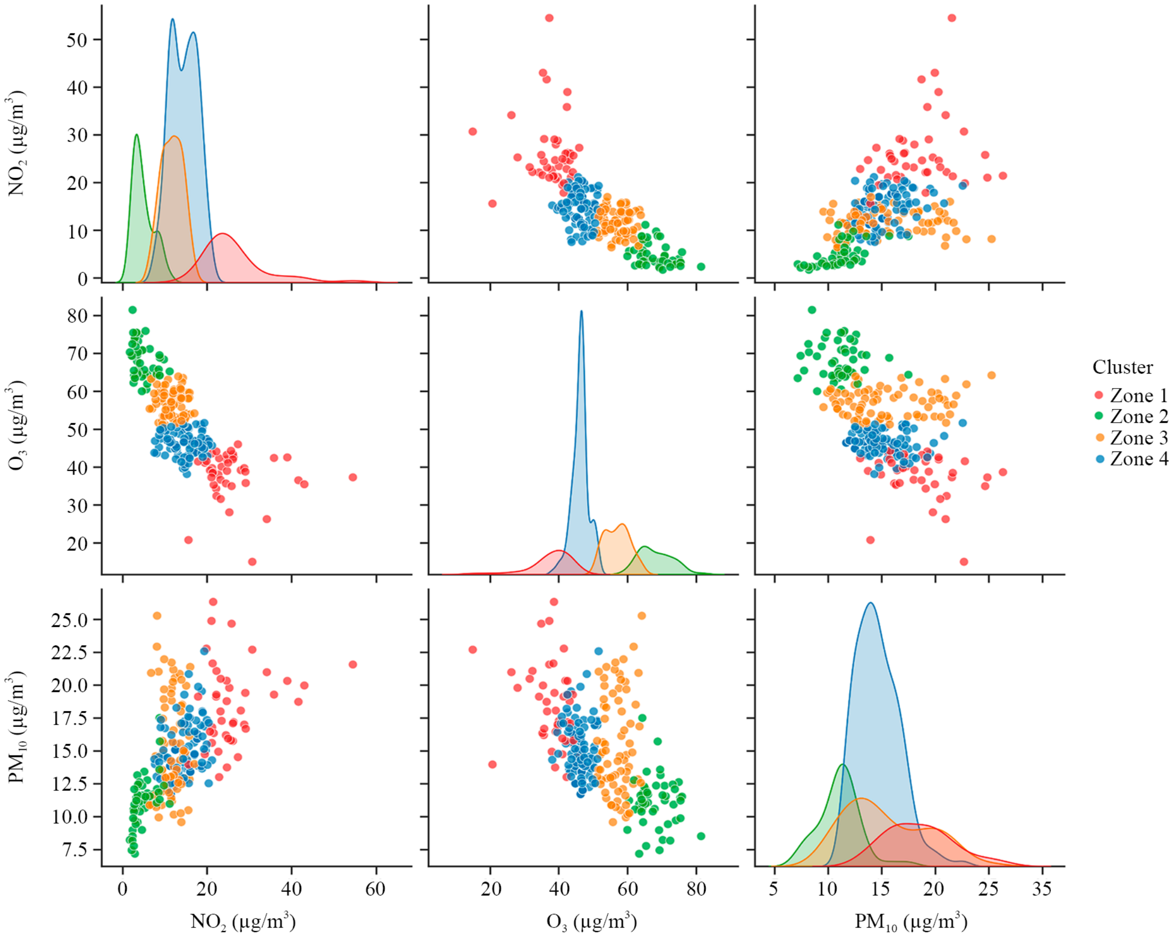


Figure 10. Relationship of pollutants by cluster (k = 4)

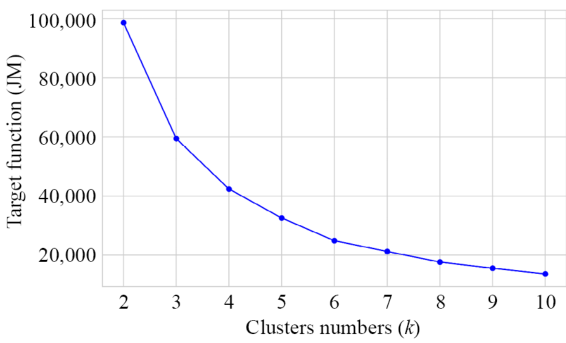


Figure 11. Fuzzy elbow clustering method

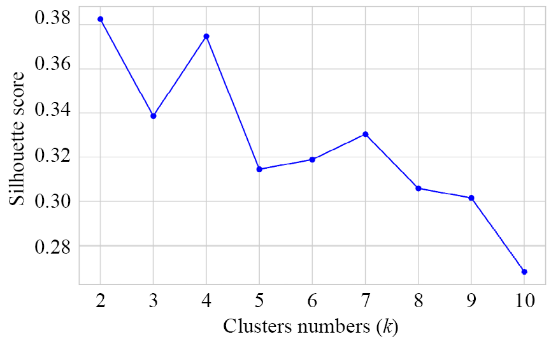


Figure 12. Fuzzy silhouette clustering method

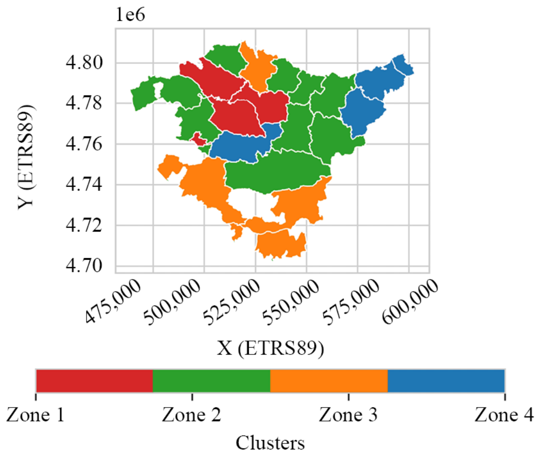


Figure 13. Air pollution zones by region

Figure 13 shows the identified four zones, while Table 2 shows the average pollutant concentrations by area over the last 10 years. Analyzing the results in Table 2, it is observed that Zone 1 presents high levels of NO_2 (21.13 µg/m^3) and O_3 (58.42 µg/m^3), as well as a relatively high concentration of PM_{10} (16.64 µg/m^3), which suggests areas with greater influence of emissions from traffic and industrial sources, characterized by common sources of combustion and transportation. Zone 2 exhibits moderate levels of NO_2 (8.18 µg/m^3) and O_3 (63.11 µg/m^3), with the lowest value in PM_{10} (12.65 µg/m^3), indicating regions with lower vehicular load and industrial occupancy, although with the presence of ozone at high levels. Zone 3 presents intermediate values for all pollutants (NO_2: 13.14 µg/m^3, O_3: 53.67 µg/m^3, PM_{10}: 14.46 µg/m^3), characterizing areas with moderate exposure conditions and possibly mixed influences from dispersed sources. Finally, zone 4 shows a similar trend to Zone 3 but with slightly higher levels of NO_2 (17.61 µg/m^3), lower levels of O_3 (45.49 µg/m^3), and similar levels of PM_{10} (17.55 µg/m^3), probably associated with areas with lower vehicular mobility and reduced industrial processes.

Table 2. Average concentrations of pollutants by area over the last 10 years

Cluster	NO_2	O_3	PM_{10}
Zone 1	21.13	58.42	16.64
Zone 2	8.18	63.11	12.65
Zone 3	13.14	53.67	14.46
Zone 4	17.61	45.49	17.55

The partitioning of the region into four distinct pollution zones provides a granular understanding of the spatial heterogeneity in air quality. This segmentation reveals specific areas with elevated pollutant concentrations, which are critical for targeted environmental management and policy intervention. Zone 1 encompasses the regions of Gran Bilbao, Duranguesado, and Arratia-Nervión, characterized by high levels of NO_2 and PM_{10} concentrations, which suggest contributions from both traffic and industrial combustion processes. Major municipalities such as Bilbao, Barakaldo, Getxo, Durango, Amorebieta-Etxano, Areatza, and Orduña, contribute significantly to these pollution levels due to a combination of high traffic density and industrial activity distributed throughout the zone. The presence of elevated O_3, a secondary pollutant formed through photochemical reactions involving NO_x and Volatile Organic Compounds (VOCs), further underscores the complex interplay of emissions in these areas. In contrast, Zone 2 exhibits moderate NO_2 (8.18 $\mu g/m^3$) but notably high O_3 (63.11 $\mu g/m^3$) levels, with the lowest PM_{10} (12.65 $\mu g/m^3$) concentrations. This profile suggests regions with reduced direct emissions from traffic and industry but significant secondary O_3 formation. The high O_3 levels are likely driven by the transport of precursor pollutants (NO_x and VOCs) from nearby urban or industrial areas, combined with favorable meteorological conditions for photochemical reactions. To address this, strategies should focus on mitigating secondary pollutant formation, such as increasing green spaces to enhance the natural removal of O_3 precursors and implementing VOC reduction programs in nearby industrial zones. This zoning approach not only identifies areas requiring immediate attention but also provides a framework for tailored air quality management strategies, ensuring that interventions are both effective and region-specific.

3.1 *Discussion*

The results presented in the preceding section offer a comprehensive look at the air quality dynamics in the BAC over a transformative decade. The descriptive statistics and spatial interpolations reveal distinct patterns in pollutant concentrations, particularly when comparing the pre-COVID (2015-2019) and COVID-era (2020-2024) periods. This section will delve into the interpretation of these findings, drawing upon established scientific literature to provide context and potential explanations for the observed trends.

A central theme emerging from our analysis is the significant impact of the COVID-19 lockdown measures on air quality. This real-world experiment provides a unique lens through which to examine the contributions of various emission sources. Our findings are broadly consistent with studies conducted elsewhere in Spain and across Europe, which have documented marked reductions in traffic-related pollutants following the implementation of mobility restrictions. One by one:

The observed decrease in Nitrogen Dioxide (NO_2) concentrations during the lockdown period aligns with the well-established fact that vehicle exhaust is a primary source of this pollutant. A study in the Basque Country found that urban and interurban road traffic decreased by 49% and 53%, respectively, during the lockdown, leading to a 45% drop in urban NO_2 levels [3]. Our data corroborates this trend, highlighting the direct and immediate impact of reduced traffic on ambient NO_2 concentrations.

Similarly, the reduction in Benzene and Carbon Monoxide (CO) levels can be attributed to the decrease in vehicular traffic, as both are products of incomplete fossil fuel combustion. Benzene is also associated with industrial activities, and while industrial activity in the region saw a more modest reduction of 20% compared to traffic, this likely contributed to the overall decline.

The behavior of Particulate Matter (PM_{10}) is more complex. While a reduction in traffic and industrial emissions would be expected to lower PM_{10} levels, our analysis, like that of other studies, may show a less pronounced decrease than for gaseous pollutants. This is because PM_{10} has a more diverse range of sources, including construction, agriculture, and natural sources like dust. Furthermore, secondary aerosol formation-whereby gaseous pollutants like SO_2 and NO_x are converted into particulate matter in the atmosphere-can offset the reduction from primary emission sources.

The case of Ozone (O_3) is particularly intriguing and underscores the complexity of atmospheric chemistry. Unlike the primary pollutants discussed above, ozone is a secondary pollutant formed through photochemical reactions involving nitrogen oxides (NO_x) and volatile organic compounds (VOCs) in the presence of sunlight. In urban areas with high NO_x emissions (a "VOC-limited" or "NO_x-saturated" regime), the high levels of NO can actually "scavenge" or destroy ozone molecules. Therefore, a sharp reduction in NO_x emissions, as seen during the lockdowns, can lead to a counterintuitive *increase* in local ozone concentrations, a phenomenon known as the "weekend effect" that was observed in some urban areas during the pandemic. This highlights the non-linear and often paradoxical responses of the atmospheric system to changes in emissions.

For Sulfur Dioxide (SO_2), the primary sources are industrial processes and the burning of sulfur-containing fuels in power plants. The observed trends in SO_2 in our study are likely linked to changes in industrial activity levels during the lockdown and the broader, long-term trends in the use of cleaner fuels and emission control technologies in the industrial sector.

It is also important to consider the role of meteorology in modulating pollutant concentrations. Weather patterns, including wind speed and direction, temperature, and rainfall, can have a significant impact on the dispersion and chemical transformation of pollutants. While our study focuses on the long-term trends, it is important to acknowledge that year-to-year variability can be influenced by meteorological factors.

In summary, the analysis of the 2015-2024 period provides a clear signal of the impact of the COVID-19 pandemic on air quality in the BAC. The sharp reduction in traffic-related pollutants like NO_2 and Benzene serves as a powerful demonstration of the potential air quality benefits of reducing our reliance on fossil fuel-powered transportation. However, the more complex behavior of pollutants like PM_{10} and O_3 underscores the need for multi-faceted and nuanced air quality management strategies that account for the intricate interplay of different emission sources and atmospheric processes.

Concerning the analysis of the clusters shown in Figure 13, it can be stated that they are strongly correlated with the interplay between emission sources, orography, and meteorological factors. The Zone 1 Cluster, corresponds directly to densely populated, highly industrialized areas situated in rather deep valleys, where poor natural ventilation leads to pollutant accumulation. Conversely, the Zone 2 Cluster is found in sparsely populated, high-altitude rural regions, which benefit from minimal emissions and excellent wind exposure. Intermediate clusters, such as the Zone 4 Cluster along the coast, reflect areas with good dispersion influenced by cleansing sea breezes, even in moderately populated zones. Finally, the Zone 3 Cluster marks transition zones with moderate industrial activity and average ventilation. This spatial pattern confirms that the physical geography of the CAV dictates the ultimate fate of atmospheric pollutants.

3.2 *Conclusions*

This research meticulously documents the evolution of six pollutants and, more importantly, interprets the observed changes as a reflection of shifting societal dynamics of the BAC. The implications of this research are significant. The "natural experiment" of the COVID-19 lockdown has provided invaluable data on the sources and dynamics of air pollution in the Basque Country. It underscores that targeted policies aimed at reducing vehicular traffic can yield immediate and substantial improvements in air quality. However, it also serves as a caution that air quality management is not a simple matter of reducing one type of emission. The complex and non-linear nature of atmospheric chemistry, especially with regard to ozone and secondary particulate matter, requires a holistic approach that considers the balance of different precursor pollutants.

The research provides a scientifically grounded basis for discussing the environmental impact of the collective habits of the region. For future guidelines, and as the quality of the air is a cornerstone of public health and environmental stability, monitoring of these data crossed with health-related records, may provide enriching information. Thus, the research team will focus its efforts on linking local pollution data with health-related data. As well, the inclusion of a wider array of pollutants, particularly Volatile Organic Compounds (VOCs) and finer particulate matter, would provide an even more detailed picture of the region's air quality. Furthermore, integrating our air quality data with detailed meteorological and traffic models would allow for a more precise quantification of the relative contributions of emissions and weather to the observed pollutant concentrations.

In conclusion, this study not only builds upon our previous work but also contributes to the broader scientific understanding of the environmental impacts of the COVID-19 pandemic. The insights gained from this unique period should be leveraged to inform the development of more effective, evidence-based air quality policies for a healthier and more sustainable future.

Finally, it is worth mentioning that he validity of this research has been as well used to reinforce the plausibility of these techniques to noise reduction programs of two steel plants, better detecting and predicting noise sources to take predictive actions on them.

Funding

This research was funded by the University of the Basque Country UPV/EHU, grant number UZTATU HBP PIE i3lab 24-24, and by the European Commission project SILENCE, Research Program of the Research Funds for Coal and Steel-Prj. No.: 101112516. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Sokhi RS, Singh V, Querol X, Finardi S, Targino AC, Andrade MF, et al. A global observational analysis to understand changes in air quality during exceptionally low anthropogenic emission conditions. *Environment International*. 2021; 157: 106818. Available from: https://doi.org/10.1016/j.envint.2021.106818.

[2] Alberdi E, Álvarez I, Hernández H, Oyarbide-Zubillaga A, Goti A. Analysis of the air quality of the Basque Autonomous Community using spatial interpolation. *Sustainability*. 2020; 12(10): 4164. Available from: https://doi.org/10.3390/su12104164.

[3] Gangoiti G, De Blas M, Gómez M, Rodríguez-García A, Torre-Pascual E, García-Ruiz E, et al. Impact of the COVID-19 lockdown in a European regional monitoring network (Spain): Are we free from pollution episodes? *International Journal of Environmental Research and Public Health*. 2021; 18(21): 11042. Available from: https://doi.org/10.3390/ijerph182111042.

[4] Barrero MA, Orza JA, Cabello M, Cantón L. Categorisation of air quality monitoring stations by evaluation of PM_{10} variability. *Science of the Total Environment*. 2015; 524-525: 225-236. Available from: https://doi.org/10.1016/j.scitotenv.2015.03.138.

[5] De Jalón G, Burgess P, Yuste C, Moreno G, Graves A, Palma J, et al. Dry deposition of air pollutants on trees at regional scale: A case study in the Basque Country. *Agricultural and Forest Meteorology*. 2019; 278: 107648. Available from: https://doi.org/10.1016/j.agrformet.2019.107648.

[6] Querol X, Alastuey A, Pandolfi M, Reche C, Pérez N, Minguillón M, et al. 2001-2012 trends on air quality in Spain. *Science of the Total Environment*. 2014; 490: 957-969. Available from: https://doi.org/10.1016/j.scitotenv.2014.05.074.

[7] Borge R, Réquia W, Yagüe C, Jhun I, Koutrakis P. Impact of weather changes on air quality and related mortality in Spain over a 25 year period [1993-2017]. *Environment International*. 2019; 133: 105272. Available from: https://doi.org/10.1016/j.envint.2019.105272.

[8] Gangoiti G, Albizuri A, Alonso L, Navazo M, Matabuena M, Valdenebro V, et al. Sub-continental transport mechanisms and pathways during two ozone episodes in northern Spain. *Atmospheric Chemistry and Physics*. 2006; 6(6): 1469-1484. Available from: https://doi.org/10.5194/acp-6-1469-2006.

[9] Santa-Marina L, Barroeta Z, Irizar A, Álvarez J, Abad E, Muñoz-Arnanz J, et al. Characterization of PCDD/F and dl-PCB levels in air in Gipuzkoa (Basque Country, Spain). *Environmental Research*. 2023; 228: 115901. Available from: https://doi.org/10.1016/j.envres.2023.115901.

[10] Zhao Z, Wu J, Cai F, Zhang S, Wang YG. A statistical learning framework for spatial-temporal feature selection and application to air quality index forecasting. *Ecological Indicators*. 2022; 144: 109416. Available from: https://doi.org/10.1016/j.ecolind.2022.109416.

[11] Alahamade W, Lake I, Reeves CE, de La Iglesia B. Evaluation of multivariate time series clustering for imputation of air pollution data. *Geoscientific Instrumentation, Methods and Data Systems*. 2021; 10(2): 265-285. Available from: https://doi.org/10.5194/gi-10-265-2021.

[12] Chen PC, Lin YT. Exposure assessment of $PM_{2.5}$ using smart spatial interpolation on regulatory air quality stations with clustering of densely-deployed microsensors. *Environmental Pollution*. 2022; 292: 118401. Available from: https://doi.org/10.1016/j.envpol.2021.118401.

[13] Aguilar FJ, Agüera F, Aguilar MA, Carvajal F. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogrammetric Engineering and Remote Sensing*. 2005; 71(7): 805-816. Available from: https://doi.org/10.14358/PERS.71.7.805.

[14] Chen C, Li Y, Zhao N, Guo B, Mou N. Least squares compactly supported radial basis function for digital terrain model interpolation from airborne LiDAR point clouds. *Remote Sensing*. 2018; 10(4): 587. Available from: https://doi.org/10.3390/rs10040587.

[15] Rusu C, Rusu V. Radial basis functions versus geostatistics in spatial interpolations. In: Bramer M. (ed.) *IFIP International Federation for Information Processing*. Boston, MA: Springer; 2006. p.119-128.

[16] Bezdek JC, Ehrlich R, Full W. FCM: the fuzzy c-means clustering algorithm. *Computers and Geosciences*. 1984; 10(2-3): 191-203. Available from: https://doi.org/10.1016/0098-3004(84)90020-7.

[17] Rustam Z, Kamalia A, Hidayat R, Subroto F, Suryansyah SA. Comparison of fuzzy c-means, fuzzy kernel c-means, and fuzzy kernel robust c-means to classify thalassemia data. *International Journal on Advanced Science, Engineering and Information Technology*. 2019; 9(4): 1205-1210. Available from: https://doi.org/10.18517/ijaseit.9.4.9580.

[18] Xie XL, Beni G. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991; 13(8): 841-846.

[19] Open Data Euskadi Portal. *Basque Government Open Data Portal*. 2023. Available from: https://opendata.euskadi.eus/inicio/ [Accessed 23rd December 2023].