



Monitoring of NO₂ air pollution in the port of Livorno and spatialization of data

Chiara Collaveri¹ · Bianca Patrizia Andreini¹ · Elisa Bini¹ · Luca Borin² · Fiammetta Dini¹ · Stefano Fortunato¹ · Camilla Grossi² · Marina Rosato¹ · Bianca Sivieri¹

Received: 3 October 2023 / Accepted: 6 February 2024 / Published online: 26 February 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

In this work, an estimate of the spatial distribution of NO₂ is presented starting from data measured with passive samplers in 10 locations in the municipality of Livorno in Tuscany (Italy). The data from the passive samplers were integrated with measurement campaigns carried out within the port of Livorno and with data from the fixed stations. The Municipality of Livorno is subject to pressures deriving from the emissions of the port, the heavy industry and the demographic activity (traffic and heating) of a municipality with middle/high population density (270 inhabitants/km²) for Tuscany Region. Despite the many and varied pressures, the only exceedances of the air quality limit values in the last decade concerned the annual average of nitrogen oxides in the urban traffic station. However, the port makes an important contribution in terms of emissions to nitrogen oxides, therefore the main objective of this work is to represent the NO₂ levels in the urban area of Livorno, highlighting the various contributions. To do this, Ordinary Kriging was carried out on the measured values after removing the local trend through the use of a beta index, a method reported in the literature for Belgian network and also applied in Italy for spatial representativeness. In this work we also tried to best represent the contribution of traffic as in our data set there is a urban traffic station, with the highest NO₂ levels in Tuscany, lower only than those in the regional capital Firenze. With a very simplified method, that can be improved in future works, it was possible to estimate the effects of the port on the city in comparison with the other sources, treating the background levels separately and stratifying the levels of road traffic based on the flows of the main roads.

Keywords Kriging · Emissions · Air quality · Port · Nitrogen oxides

Introduction

Port activities have a major impact on the surrounding urban areas. In terms of emissions, naval traffic makes a non-negligible contribution to municipal emissions both for the main pollutants (NO_x, SO_x, PM and NMVOC) and for greenhouse gases (CO₂). In addition to maneuvers within the port area, substantial contributions derive from the housing of ships at the docks in most cases with auxiliary engines running.

In this work, an estimate of the spatial distribution of NO₂ is presented starting from data measured with passive samplers in 10 locations in the municipality of Livorno in Tuscany (Italy). The measurements were carried out as part of the Aernostrum project on air quality in ports and at the port-city interface. Aernostrum is an Interreg project that involved several Mediterranean ports to estimate the impacts of port activity and identify possible mitigation measures [Aernostrum Interreg Maritime n.d.]. The data collected with the passive samplers were integrated with those from the fixed air quality stations and mobile measurements carried out within the port, as part of the project. The data was processed to produce a map of the spatial distribution of NO₂ in a domain within the municipality, to represent the contribution of the port to the interface with the city. For each site a β index (Janssen et al. 2008), (Janssen et al. 2012), (Piersanti et al. 2013) was calculated; this index is used to remove at each point the trend due to

✉ Chiara Collaveri
c.collaveri@arpat.toscana.it

¹ Regional Center for Air Quality, ARPAT, Via Giovanni Marradi 114, Leghorn, Italy

² Earth Science Department, Pisa University, Via Santa Maria 53, Pisa, Italy

the local contributions of the pressure sources in order to carry out a Ordinary Kriging on the detrended data and spatialize the measured data. This work was focused on defining the appropriate factors to build the beta index for the domain and context studied. In particular, specific first guess coefficients have been developed for traffic, for port and industrial emissions and also for urban areas. In particular, according to the optimization results, it was evaluated how to manage the traffic contribution; the optimization is the process that linearizes the relationship between the beta index and the measured levels by acting on the first guess coefficients, and comparison between first guess coefficients and optimized coefficients also gives important information.

Materials and methods

Beta index

The beta index is a parameter that can be calculated at each point of the domain and associated with each monitoring site. It represents the characteristics of a site by summarizing the information on land use and local emission pressures in a single value.

The beta index was proposed and used in Belgium for the evaluation of the spatial representativeness of the stations (Janssen et al. 2012), and for the interpolation of air quality

data (Janssen et al. 2008). It is calculated in a buffer around the measurement site according to the formula:

$$\beta = \log \left[1 + \frac{\sum_i a_i \cdot n_{RCLi}}{\sum_i n_{RCLi}} \right]$$

n_{RCLi} number of pixels (or area) in the i -th Corine Land Cover class.

a_i coefficient related to emissions for the i -th class, used to weigh the contribution of a particular class of land use on the concentrations of pollutants in the air.

The radius of the vicinity buffer is a free parameter in the methodology, but for Janssen et al. 2008, it was adopted a radius of 2 km. The radius chosen can depend also on the type of pollutant considered and on the variability of land use in the area of the study. When the buffer is chosen too small, it becomes difficult to have more than one of the CLC classes included in the buffer (essential for a good optimization), and in this way it is also considered a very local contribution. When the buffer is too large, the site-specific character of the CLC class distribution disappears and the distribution evolves to a general spectrum with very little discriminating power. In this study as better explained in the next section a buffer radius of 1 km was chosen.

The logarithm function, given by the methodology, is related to the fact that the beta index is a synthetic index of

Table 1 Sampling periods

Sampling periods			
Site	Method	Sampling start	Sampling End
Calambrone	Passive samplers	03/08/2021	07/09/2021
Via Enriques			
Bengasi (check sampling)*		30/11/2021	24/12/2021
Bastia			
Provincia		03/03/2022	17/03/2022
Regina	Air quality stations Tedyne analyzers	17/03/2022	01/04/2022
ARPAT			
Gemini			
Via Toti			
LI-Cappiello		01/01/2021	31/12/2022
LI-La Pira	Mobile monitoring		
LI-Carducci			
LI-ENI Stagno			
Bengasi MM			
	Teledyne analyzer	10/08/2021	10/09/2021
		11/09/2021	12/10/2021
		11/01/2022	01/03/2022
		02/03/2022	27/04/2022

pressures on air quality for which a linear relationship with air quality data is assumed.

The beta index is then specific for each pollutant of interest. In order to establish a relationship between the polluting potential and the land use classes, the CLC classes are grouped into 11 categories and then combined with the EMEP sectors of the activities that produce emissions into the atmosphere, same approach used in Janssen et al. (2008). In the Tuscany Region the beta index was previously applied to estimate the representativeness of the regional network stations, limited to PM10 background stations (Regione Toscana Rappresentatività 2015), as indicated in the methodology proposed at national level by ENEA (Piersanti et al. 2013). Coefficients optimization is made with the R software [R Core Team 2022] and in particular the optim routine, using the L-BFGS-B methodology (Byrd et al. 1995). This methodology allows you to define appropriate variability constraints (maximum and minimum value) for each of the independent variables of the optimization process. The constraint imposed for this study is that the coefficients,

representing an emission pressure associated with a given land use, are all positive.

NO₂ measure

The measurements with passive samplers were made with Radiello® Rad66 for NO_x and SO_x, with 15 days sampling and according to the UNI EN16339 method. The sampling period was repeated four times during the year covering different seasons and for this work the average of all sampling was used as an estimate of the annual average in each location. The NO_x data from the fixed and mobile stations are measured by Teledyne API T200, Teledyne API 200A or Thermo Electron 42 analysers, in compliance with the technical standard UNI EN 14211 and D.Lgs 155/2010, which in Italy is the transposition of the European Directive for air quality. The sampling periods are shown in Table 1, while the sampling sites are shown in Fig. 1.

Calculating the composite uncertainty with the maximum uncertainties for each step of the method provided by the EN

Fig. 1 Sampling sites

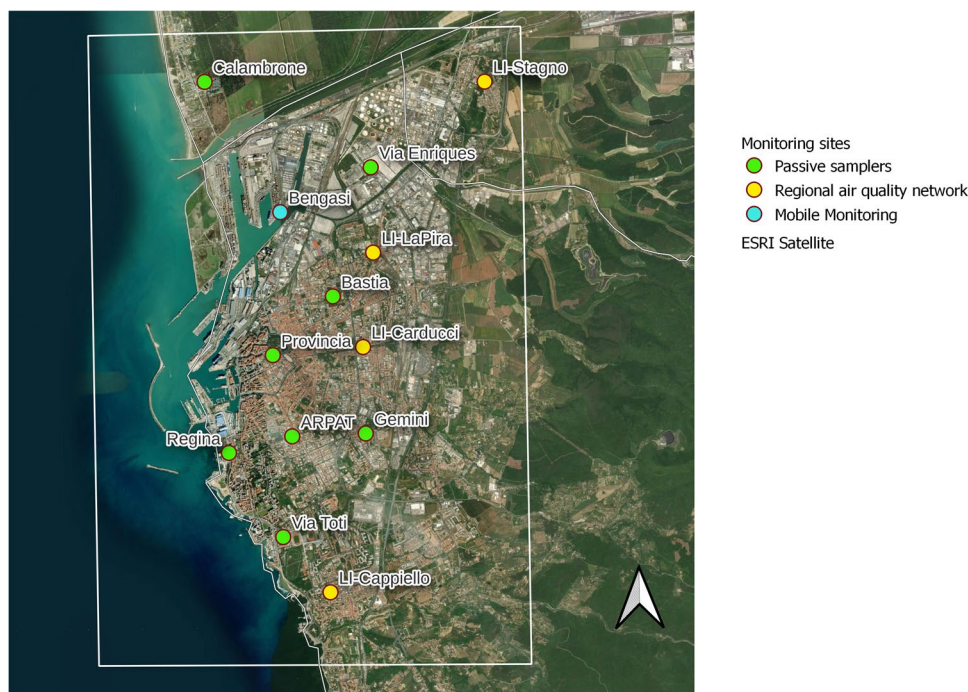


Table 2 Comparison between passive samplers and continuous sampling

Site/method	03.08.21– 07.09.21 µg/m ³	30.11.21– 14.12.21 µg/m ³	03.03.22– 17.03.22 µg/m ³	17.03.22– 01.04.22 µg/m ³
Li-La Pira Passive sampler	14	25	16	18
Li-La Pira Teledyne (same periods)	14	26	18	19
Bengasi* Passive sampler	25	-	-	-
Bengasi Teledyne	29	-	28	32

16339:13, we obtain for the results of the passive samplers an expanded uncertainty of about 30%. Continuous analyzers used for monitoring in this study comply with UNI EN 14211:2012 as indicated by Legislative Decree 155/2010, and thus have expanded uncertainty of less than 15 percent. To verify the comparability between the two methods, passive sampling was performed in parallel with Teledyne at the LI-La Pira site and, limited to the August campaign, at the Bengasi site. The results are shown in Table 2.

Representativeness on an annual basis is ensured by coverage of different seasonal periods as required by Legislative Decree 155/2010 which, for indicative campaigns, prescribes that the measurement be carried out on variable days of the week and distributed throughout the year for a total of eight weeks. In addition, to verify representativeness on an annual basis, the results of the radiello on Li La Pira were compared with those of the regional LI-La Pira station, which has full annual coverage. The results are shown in Table 3.

To verify that sampling periods are representative also from a meteorological point of view the wind regime was examined, and wind roses for each of the periods in Table 3 are shown in Fig. 2. The city of Livorno is subject to breeze regimes and the

most frequent winds are those on the northeast/southwest axis that are also the directions of strongest winds.

The wind profile results similar in all periods.

The domain is a rectangle of dimensions $10 \times 6,7$ km containing the urban area of the Municipality of Livorno, the port area and the industrial area of Livorno Nord, extended to the municipality of Collesalveti; the domain also contains the Calambrone area, a fraction of the Municipality of Pisa, which is a natural area and constitutes an impact zone for both the industrial area and the port. All cartographic processing were carried out with QGIS (3.26 Firenze) (QGIS Project n.d.). For the Ordinary Kriging geoprocessing was made using the plugin Smart Map (Pereira et al. 2022).

The emission data, classified according to SNAP97 codes, are extracted from the IRSE Regional Inventory of Emission Sources of the Tuscany Region for the year 2017 (IRSE Database 2017).

Additional traffic data were downloaded from Open Street Maps [Open Street Maps Project (n.d.)], available with Open Database License, for the study domain.

Data of Land Use are from Corine Land Cover 2018 (Corine Land Cover n.d.).

Table 3 Comparison on annual mean basis

	Annual means [$\mu\text{g}/\text{m}^3$]
Li-La Pira Passive sampler	18
Li-La Pira Teledyne (same periods)	19
Li- La Pira annual mean 2022	17
Li- La Pira annual mean 2021	16
Li- La Pira annual mean from 08.21 to 07.22	17

Results and discussion

On the 13 points considered, the distribution of land use over a 1 km buffer was examined. The radius, smaller than that used for PM10, is motivated both by the lower spatial representativeness that can be expected for NO₂ and by the dense measurement grid with a pitch of approximately 1 km.

Before choosing the buffer radius, several possibilities were explored (Fig. 3a, b, c).

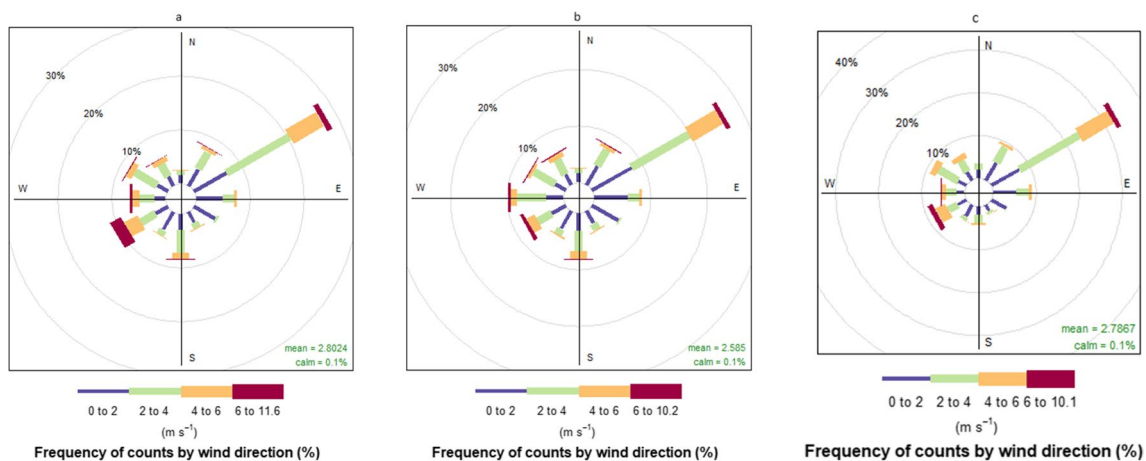


Fig. 2 Wind rose comparison in 2021 (a), 2022 (b) and in the sampling periods (c)

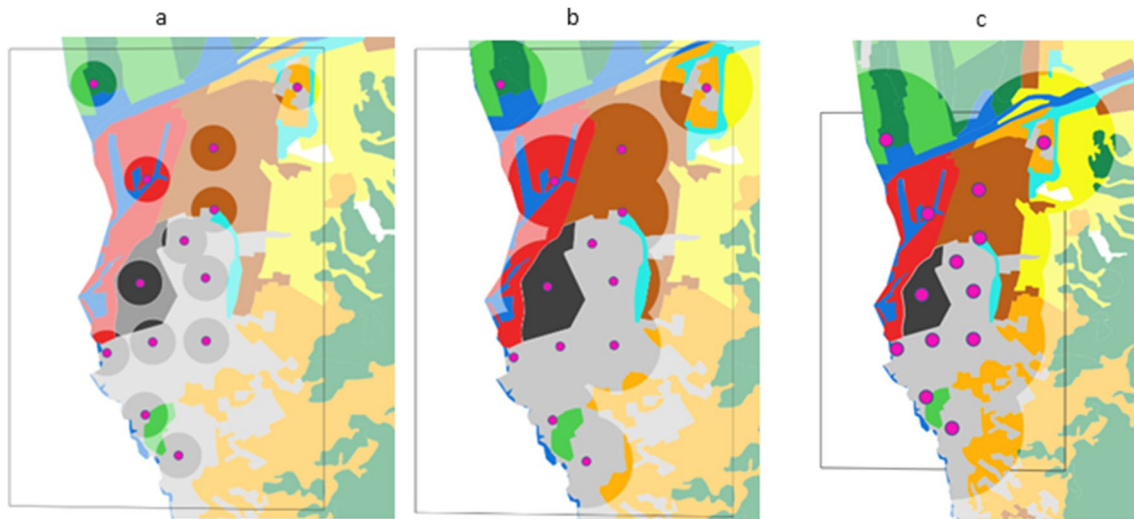


Fig. 3 Buffer radius 500 m (a), 1000 m (b), 2000 m (c)

RIO classes are highly aggregated so land use in many buffers with a radius of 500 m would be of a single type, thus not allowing a good optimization of the relationship between beta and NO_2 value. A 2-km buffer on the chosen domain, on the contrary, does not seem suitable to adequately represent a contribution to beta that could differentiate local specificities.

To choose the study domain, the distribution of land use in the municipality of Livorno and neighboring municipalities were examined (Fig. 4a, b and Fig. 5). The study domain was selected to focus on the urban area, excluding all the countryside around the city but including all the industrial

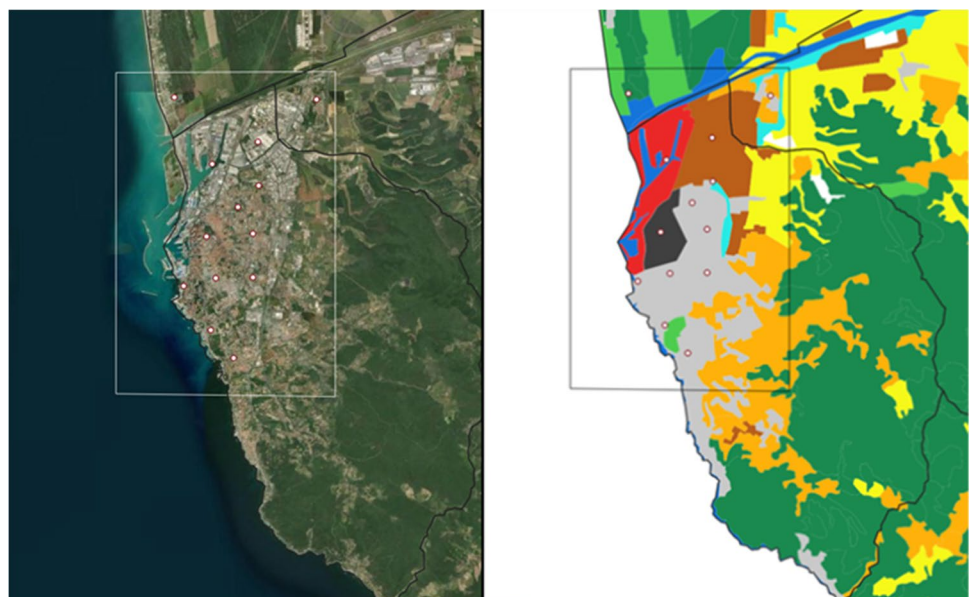
areas further north, in the neighboring municipalities of Collesalveti and Pisa.

The entire domain area includes two land use classes that do not fall into any of the buffers: class 142, urban green, with a percentage of 0.5% and class 132, landfills, with a percentage of 0.3%.

For the purposes of this paper for beta calculation, class 142, where present, is assimilated to RIO CL2 (discontinuous urban fabric) and class 132 is assimilated to RIO CL3 (Industrial or commercial units).

Four sites (Calambrone, Cappelio, Regina and via Toti) have a smaller buffer area because they are less than 1 km

Fig. 4 Satellite view and Land use view of the Municipality and domain of study



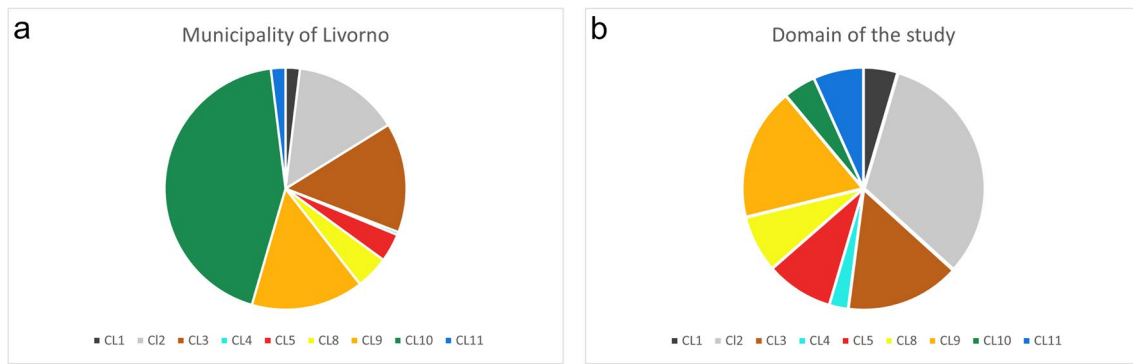


Fig. 5 Comparison of Land use distribution in the municipality of Livorno (a) and in the Domain of the study (b)

from the sea, and the marine body of water is not considered. In a first approximation the marine body is not considered because it is assumed that the sea does not bring an emission contribution on the city. Actually, the sea routes and the roadstead could have a non-negligible contribution but the emission data available at the moment are very uncertain. This issue could be further investigated in the future.

RIO Land use category for each site in a buffer radius of 1 km are shown in Fig. 6. Continuous urban areas are found exclusively in the city center and are in high percentage only in the Provincia site (63%). Continuous urban areas are also present in the buffers of the sites ARPAT (24%) and Bastia (16%) and in lower percentages in Regina (8%) and Carducci (6%). Apart from the Bengasi site (which is purely port and industrial and where the

urban share between continuous and discontinuous does not reach 1%), all the sites have a significant share of discontinuous urban land use.

The ENI Stagno site, located in the municipality of Collesalveti in front of the refinery, is characterized by a rather low percentage of urbanization (11%, the lowest after Bengasi) and by the presence of a high percentage of industrial land (14%) surrounded by agricultural areas (permanent crops 30%) and fields (arable land 25%). Road infrastructures has maximum coverage at this site. Land use with road infrastructures is also present in Bastia (2.5%), Carducci (10%), Gemini (1.5%) and La Pira (2.5%), but further investigations have been carried out for this type of land use. The industrial character is prevalent in the sites of Enriques (91%) and La Pira (63%), but is also present in the sites in the northern area of Livorno,

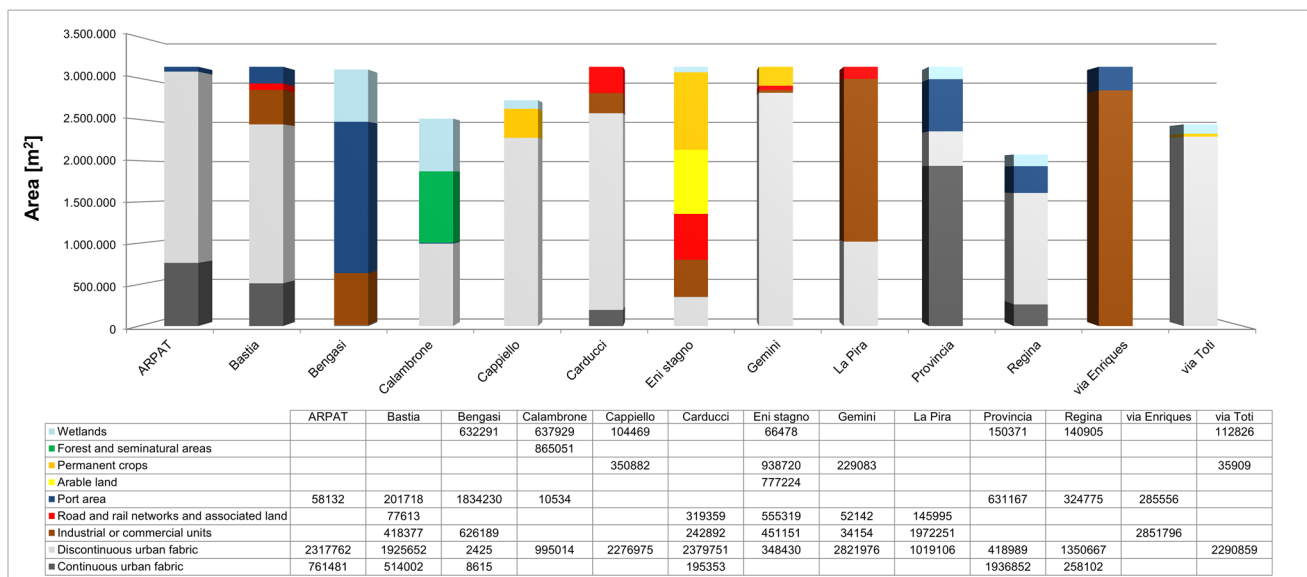


Fig. 6 RIO Land use category for each site in a buffer radius of 1 km

Bastia (13%) and Carducci (8%), in ENI-Stagno (14%) and near the port area in the Bengasi site (20%). Only Calambrone contains in its buffer forests and seminatural areas.

A first guess of coefficients is carried out according to the Table 4 of association of SNAP emission categories with Rio land use classes.

Emissions from domestic heating and urban (only diffuse) traffic emissions are associated with continuous urban fabric land use. For discontinuous fabric, to consider the lower population density, a coefficient composed of 70% of heating and 50% of urban traffic was considered. The industrial areas include macro-sectors 01, 03, 04 and 09 (excluding landfills and the burning of pruning waste, the latter attributed to agricultural areas).

The entire sector 07 of road transport is attributed to the CL4 code, with some specific modifications illustrated below.

The CL5 code includes all port activities in macro-sector 08 (080401). Maritime routes, fishing and boating

are excluded. Codes CL6 and CL7 are not present in our domain.

Agricultural activities, arable land or permanent crops, include the related emissions from macro-sector 10, off-road vehicles in agriculture and the agriculture heating sector (0203). For arable land, the entry relating to the burning of stubble is added, while for permanent crops 0907, burning of prunings and agricultural residues, is added. Wooded land and semi-natural environments are associated with a minimum amount of residential heating (1%) and sector 0807, off-road in forestry. Finally, the humid environments in the study domain represent:

- the marinas for pleasure boats to which emissions from boating (080405) have been associated;
- the body of water inside the port of Livorno, which is instead assimilated to the port land area.

In our set of measurements there is an urban traffic station, LI-Carducci, which has among the highest NO₂ levels

Table 4 First guess coefficient composition

Rio class	CLC codes	Description	SNAP sectors	Description
CL1	111	Continuous urban fabric	0202 + urban 07	Non-industrial combustion plants (residential) + Road Transport (only diffuse urban traffic)
CL2	112	Discontinuous urban fabric	0202 *0.7 + urban 07 * 0.5	(70%) Non-industrial combustion plants (residential) + (50%) Road Transport (only diffuse urban traffic)
CL3	121	Industrial or commercial units	0201 + urban 07* 0.3 + 01 + 03 + 04 + 09 (excluded 0907 and land disposal)	Non-industrial combustion plants (commercial and institutional) + (30%) Road Transport (only diffuse urban traffic) + industrial sectors
CL4	122	Road and rail networks and associated land	07	Road transports
CL5	123,124	Port area	0804 + 0805	
CL8	211,212,213	Arable land	1001 + 0203 + 0806 + 1003	Agriculture (100202 + 100105) + Non-industrial combustion plants (agricultural) + offroad vehicle in agriculture + stubble burning
CL9	221,222,223, 231,241,242, 243,244	Permanent crops, Pastures and heterogeneous agricultural areas	1001 + 0203 + 0806 + 1003	Agriculture + Non-industrial combustion plants (agricultural) + offroad vehicle in agriculture + 0907
CL10	311–313, 321–323, 331–335	Forest and seminatural areas	0807 + 0202 * 0,01	Offroad vehicle in silviculture + 1% Non-industrial combustion plants (residential)
CL11	411–412, 422–423, 511–512, 521–523	Wetlands	080405	Boating

in Tuscany, lower only than those of the FI-Gramsci urban traffic station, as in Annual report in ARPAT website (ARPAT Rapporto 2023). So the traffic contribution was deepened by extracting the roads with traffic volumes greater than 100 vehicle per hour, ($\text{trafficvol} > 100$) from the Open transport map layers falling within our domain. For each buffer around the stations, total kms of roads with $\text{trafficvol} > 100$ and the total volume of traffic from these roads were calculated (Table 5).

The two indicators are almost equivalent except for Carducci and Gemini, for which the value increases significantly if the traffic volume is considered, and for Eni Stagno for which, on the contrary, the value decreases (Fig. 7).

At this point a factor dependent on the distance of the monitoring site from the road was added.

The length/distance indicator was chosen, having already made a selection on the volume of roads, as only those with volumes greater than 100 vehicles per hour are considered. However, both these indicators seem closer to reality and also more coherent with each other (Fig. 8).

This indicator could be used in two alternative ways:

1. As a proxy to replace the CL4 variable with a surface proportional to that of the CL4 in the Carducci buffer, i.e.:

$$\text{CL4_new}(i) \Rightarrow \text{CL4}(c) * [\text{L}(i)/\text{d}(i)]/[\text{L}(c)/\text{d}(c)]$$

CL4_new: new area assignment for the i-th site

L(i)/d(i) length over distance indicator calculated for the i-th site

L(c)/d(c) length over distance indicator calculated for Carducci

CL4(c) RIO CL4 class area in the Carducci station buffer

To keep the total area of the buffer unchanged, the difference between CL4 and CL4_new for each buffer is redistributed (or scaled if negative) in the other RIO land use classes in a proportional way to their presence in the area.

2. After interpolation of background values only, as a proxy to attribute a further contribution to NO₂ levels to buffers containing selected roads

In this work the procedure in point 1 was used to calculate the beta indexes for each site with the first guess coefficients. Then the optimization was carried out with and without the Carducci station. Finally, based on the results, the method was chosen.

The coefficients thus attributed, and normalized with respect to the maximum contribution, give a value of the beta index ranging from a minimum of 0.024 for the Calambrone site to a maximum of 0.26 for the Bengasi site.

Table 5 Roads data relative to the monitoring site

	ARPAT	Bastia	Bengasi	Calambrone	Cappiello	Carducci	Eni Stagno	Gemini	La Pira	Provincia	Regina	Via Enriques	Via Toti
Total length (km)	3830	7585	3490	0	1845	7548	17080	9767	7642	6654	2404	6009	4653
Total Volume (vehicles)	118216	226453	100578		2000	457242	263448	392925	98645	164557	16031	99523	89058
Minimum distance (m)	323	90	200	-	290	12	280	280	62	20	16	580	138
Length/distance (m)	11,9	84,2	17,4		6,4	629	61	34,8	123	333	150	10,3	33,7
Volume/distance	366	2516	503		6,9	38103	941	1403	1591	8228	1002	172	645

Fig. 7 Roads data

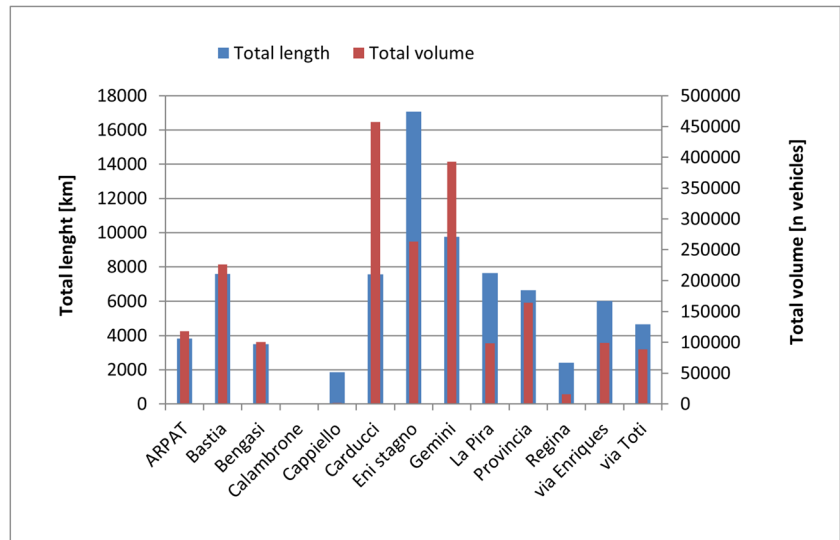


Fig. 8 Roads data relative to the monitoring sites distance

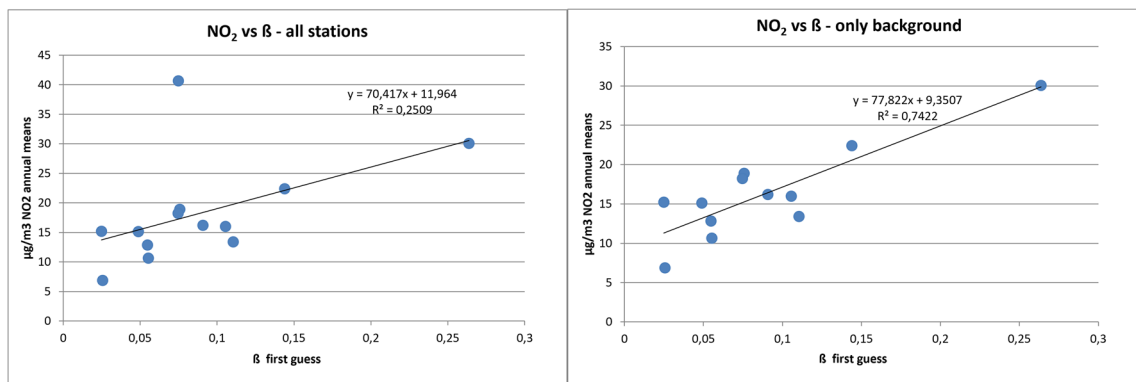
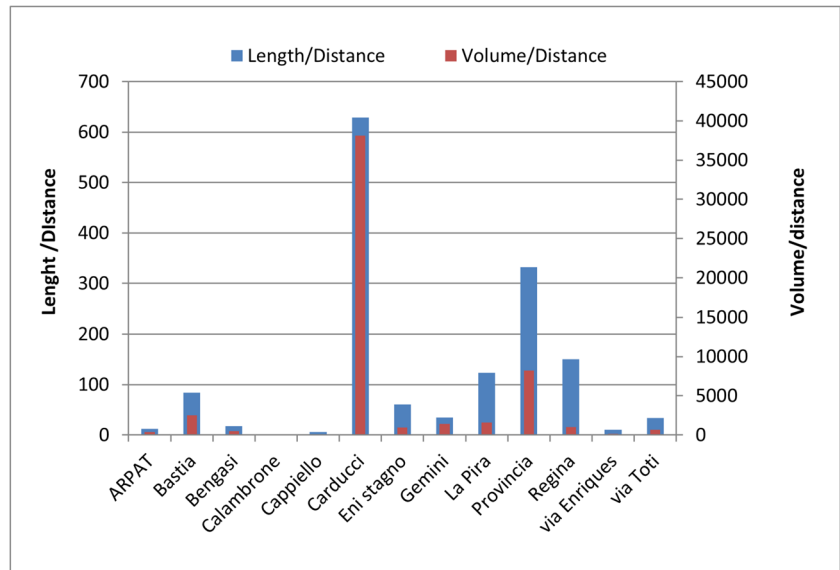


Fig. 9 Relationship between beta and NO₂ before optimization with all stations and with only background

With the first guess coefficients, the relationship between beta and measured NO₂ values gives an R² of 0.25 if all stations are considered and an R² of 0.74 if the only traffic station is excluded (Fig. 9).

After the optimization, considering all the stations the R² goes from 0.25 to 0.74, while with only the background stations it goes from 0.74 to 0.90. Furthermore, by including the traffic station, an important difference is observed between the first guess coefficients and the optimized ones; in fact, to linearize the traffic value, the routine emphasizes the coefficient relating to traffic (Fig. 10).

For this reason, the traffic station was excluded from the optimization and the contribution of the main roads added a posteriori in a 10 m buffer, as a proportional contribution to the NO₂ measured in Carducci based on the relative flows.

The beta values were used to remove the trend from the measured data. For detrending a constant concentration of 70 µg/m³ was taken as reference. Values of NO₂ calculated on the basis of beta for all the points of measure were subtracted from this value and then added to measured values. Results are shown in Fig. 11.

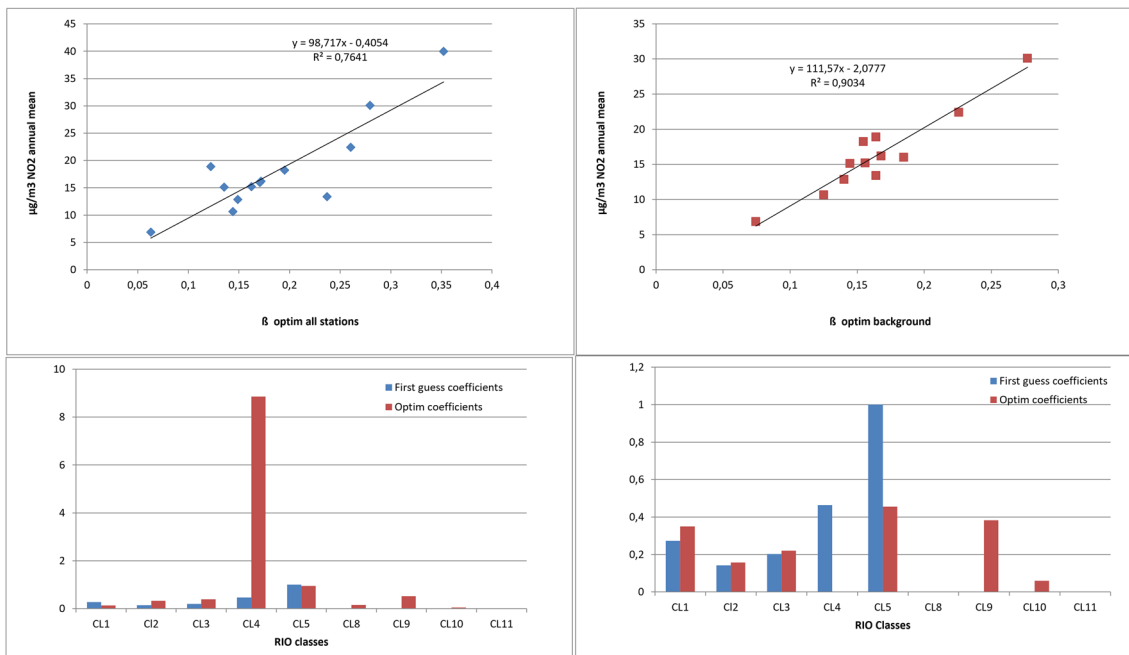
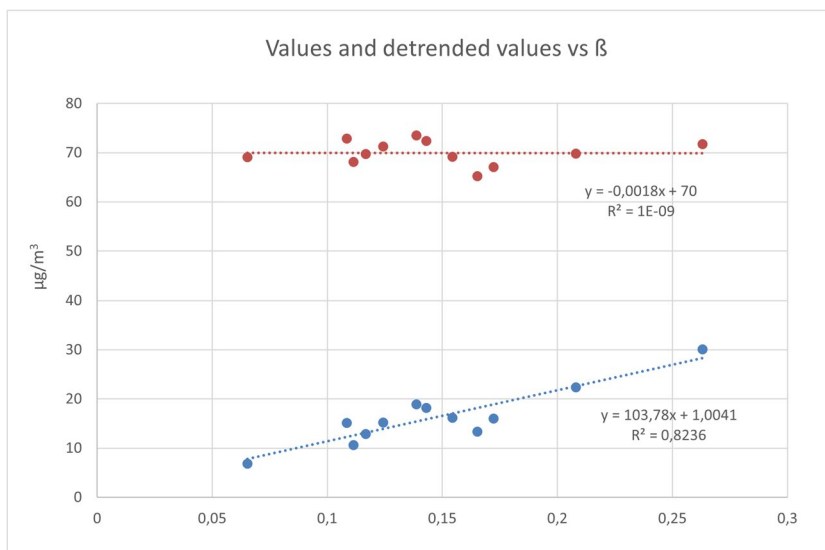


Fig. 10 Optimization results and coefficients comparison with complete dataset and only background

Fig. 11 Values and detrended values vs β



Normalization with respect to the variance is not considered here because for each point of passive samplers there are only four measure and variances do not show a trend. Detrended values are interpolated using the Smart Map Plugin (Pereira et al. 2022) in QGis with variogram settings suggested by the application, that is a maximum distance of 5064.575 m and a lag of 1695.941 m. The model selected is *linear to sill* with the sill value of $6.726 \mu\text{g}/\text{m}^3$, very similar to the sample variance. The plugin also provides the Moran index of the set of data which varies from -1 to 1 and is a measure of spatial correlation. A value of 0 would indicate a spatial pattern not unlike a random phenomenon. For the detrended values, the value of the Moran index is

0.136, indicating that most of the variance is removed by detrending data, so most of the variance in our measures seems to be explained by the beta index.

The resulting map is shown in Fig. 12 and the standard deviation estimated is $2.76 \mu\text{g}/\text{m}^3$.

On the same grid of the interpolation the β index is calculated. Beta index is here considered as a property of each point of the domain, influenced by the surroundings in a circular buffer of 1 km. The detail of calculation is better explained in Fig. 13.

The shift of the area form one point to the other is given by a 12% variation. For example, see points highlighted in figure. Passing from the left to the right the yellow area

Fig. 12 Detrended values spatial interpolation results

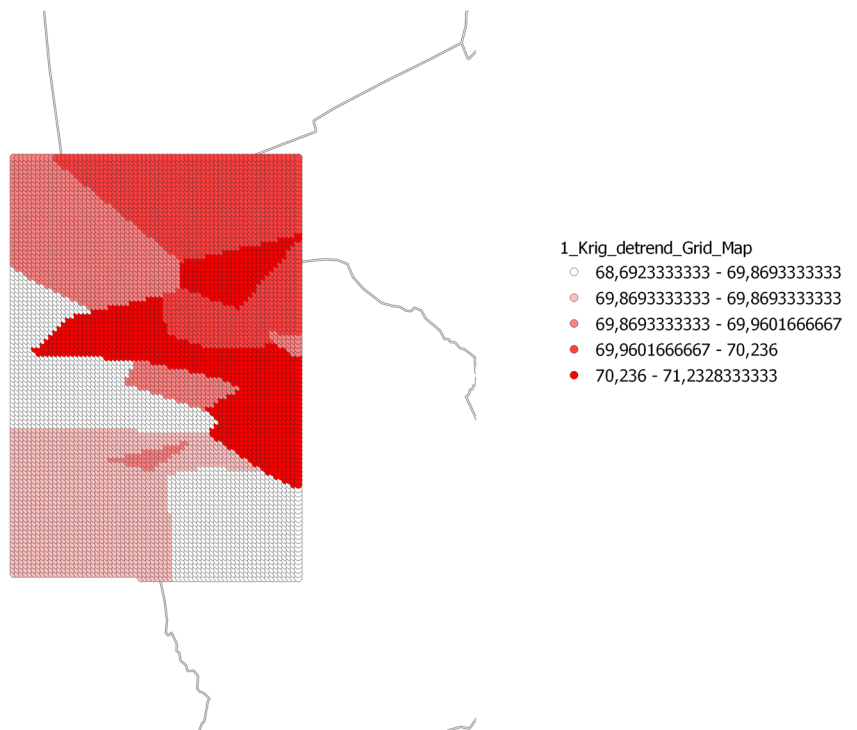
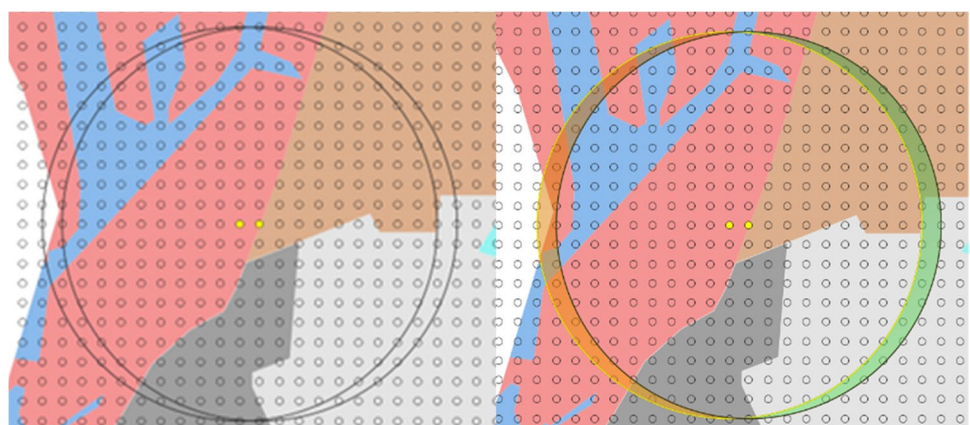


Fig. 13 Calculation of β on a grid of 100 m x 100 m



(6%) is lost whereas the green area (6%) is gained. In terms of land use the change for the case in the example is given in Fig. 14.

In terms of beta index the variation in this case is from 0,6870 for the left point to 0,6601 for the right point as the

effect of substituting harbor areas with urban discontinuous and industrial areas.

Beta distribution on the grid is shown in Fig. 15.

Once β is calculated for each point of the grid, the reordering of interpolated values was made (Fig. 16). A

Fig. 14 Comparison of Land Use distribution for two next points on a grid of 100 m × 100 m

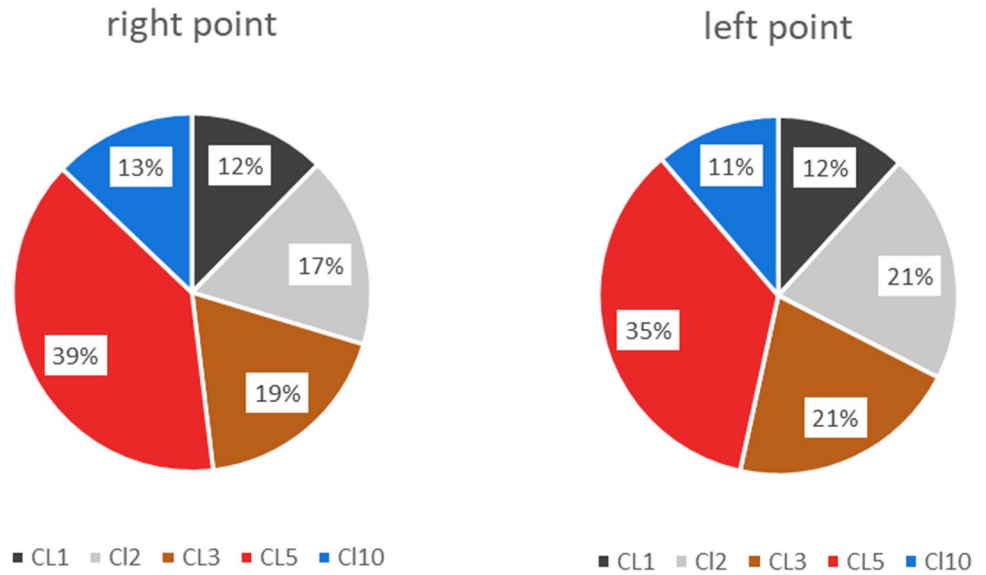
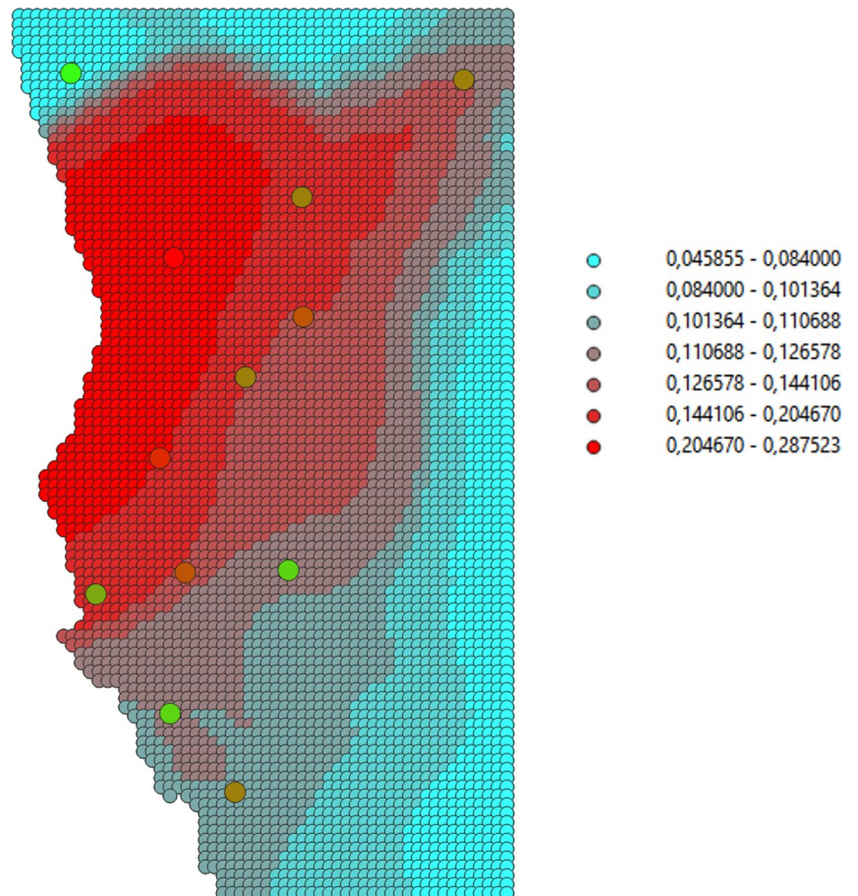


Fig. 15 Beta distribution on the domain



comparison was also made between NO₂ calculated through the optimized linear relationship between beta and measured concentrations, and NO₂ retrended (Fig. 16). The two estimates are very similar and this fact seems to confirm that almost all the variance is removed by the β. In Fig. 17 results from both methods are shown for measurement points.

Map in Fig. 18 shows the background levels estimated for NO₂ in the city of Livorno, highlighting the contribution

of the port, the port-city interface and peripheric and background values.

Roads, with a traffic volume over 100 vehicles per hour, are added to complete the representation assuming that street levels can be estimated a posteriori for each arch on the base of their relative volume and data measured in the traffic station Carducci taken as reference.

Fig. 16 NO₂ Retrended maps

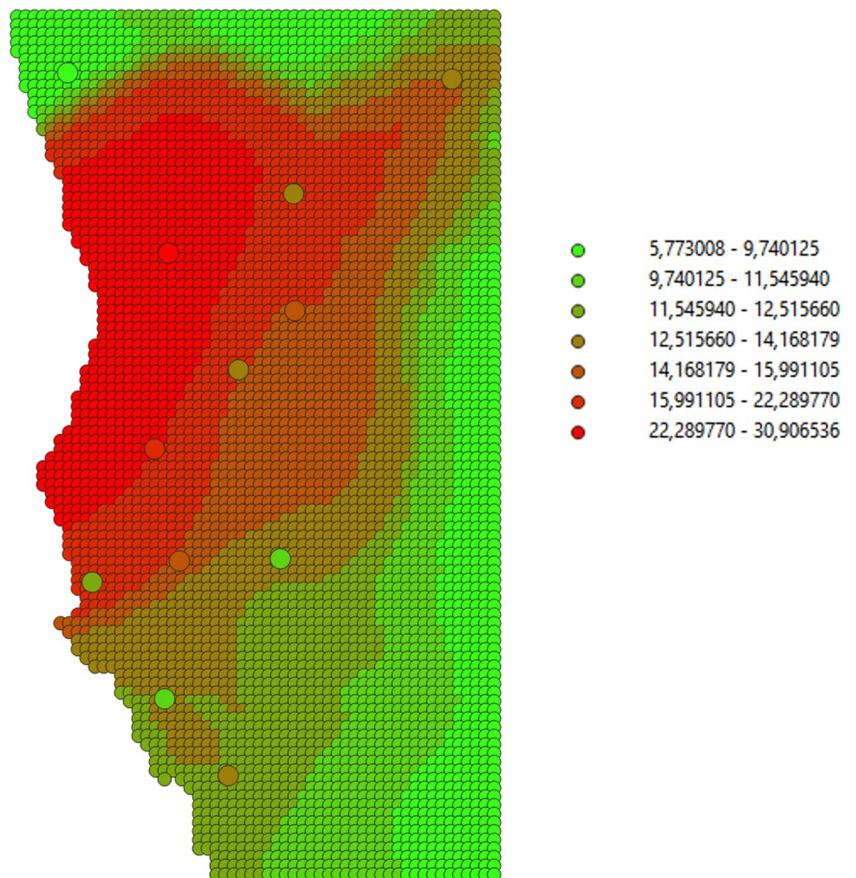


Fig. 17 Comparison between the two methods of estimating NO₂

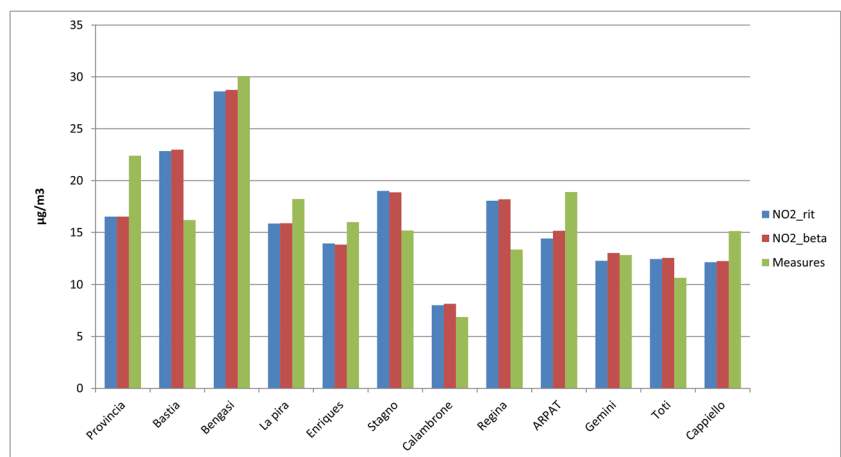
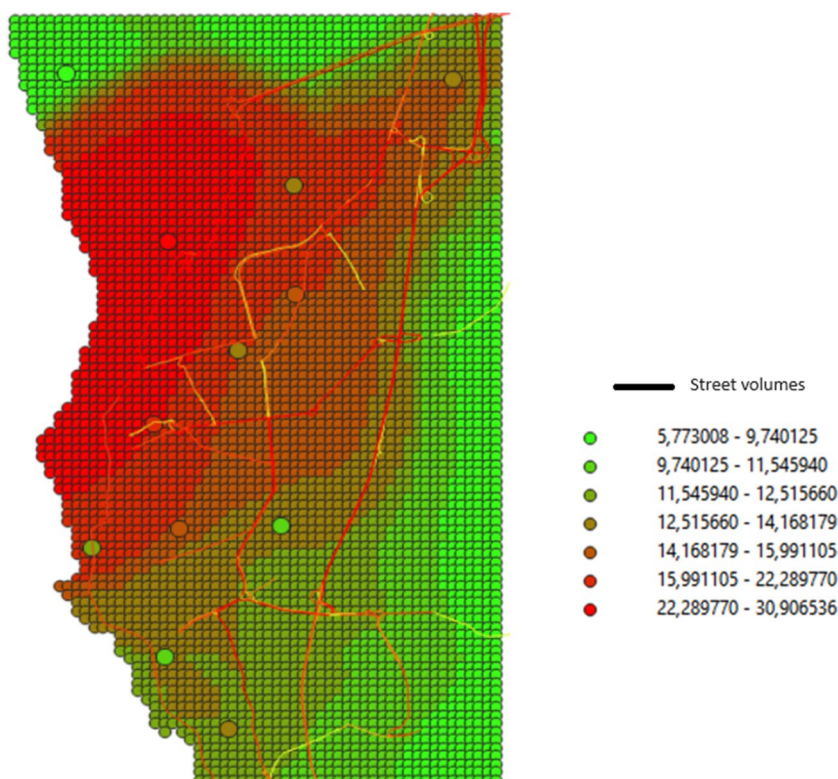


Fig. 18 Results for data spatialization: the different background levels and traffic along main roads



Conclusions

In this work, NO_2 measurements were carried out with passive samplers on a grid of points approximately 1 km apart from each other and with mobile campaigns representative of the calendar year. Some of these points were chosen within the port and others to represent adequately the port-city interface. By integrating the information obtained with the data from the monitoring stations, which represent the urban background and traffic in the city of Livorno, the spatialization of the data was carried out.

After examining various possibilities, it was decided to interpolate only the background data and, once the trend of the local contribution was removed through the use of the beta index, to treat the contribution of road traffic (limited to selected roads) as a linear source capable of influencing a buffer of a few meters around the road edge.

In the process of detrending, interpolating and retrending data, the β index is treated as a property of each point on a grid of $100 \text{ m} \times 100 \text{ m}$. The calculation of β through his linear relationship with NO_2 and the retrended map are very similar. This is probably due to the fact that detrended values vary very slightly in the domain and all the variability of data seem to be explained by beta. Even with few points, but on a well-distributed grid (1 km), for a relatively limited domain area, the β index presents good linearity with the values of the urban background even in a context where there are numerous important sources of

emissions (port and heavy industry, urban), while road traffic remains the most complicated data to manage. Excluding the traffic station and treating the “on-the road” contribution apart could be a good compromise and it is probably the reason for the good relationship between β and the measured data also before the optimization process.

Acknowledgements This work was supported by the central area laboratory, ARPAT for the chemical analyses

Author contributions All authors contributed to the study conception and design. The first draft of the manuscript was written by Chiara Collaveri and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Monitoring campaign were done for the Interreg Aernostrum project that was financed by the European Regional Development Fund (ERDF).

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets generated during and/or analysed during the current study are available as follow: all monitoring data are available on the ARPAT website (<https://www.arpat.toscana.it/notizie/arpat-news/2022/099-22/progetto-aer-nostrum-i-monitoraggi-di-arpat-per-la-qualita-dell-aria-nei-porti-di-livorno-e-portoferraio?searchterm=aernostruor>), in the Aernostrum final report, and will be available on the MO.NI.CA platform which is one product of the project. As regards to the results of the data spatialization, it can be made available on request. For all other external data used, sources are cited unambiguously in the references.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

- Aernostrum Interreg Maritime project. <https://interreg-maritime.eu/web/aer-nostrum> (Access on 01.09.2023)
- ARPAT Rapporto (2023) Relazione annuale sullo stato della qualità dell'aria, 2022. <https://www.arpat.toscana.it/documentazione/catalogo-pubblicazioni-arpat/relazione-annuale-sullo-stato-della-qualita-dellaria-in-toscana-anno-2022>. Accessed 12.01.2024
- Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput* 16:1190–1208
- Corine Land Cover The Copernicus land monitoring products and services are made available through this website on a principle of full, open and free access, as established by the Commission Delegated Regulation (EU) No 1159/2013 of 12 July 2013. (vector). <https://doi.org/10.2909/71c95a07-e296-44fc-b22b-415f42acfd0>
- IRSE Inventario regionale sorgenti di emissione Regione Toscana (2017) <https://www.regione.toscana.it/-/inventario-regionale-sulle-sorgenti-di-emissione-in-aria-ambiente-irse>. Accessed 12.01.2024
- Janssen S, Dumont G, Fierens F (2008) Clemens Mensink. *Atmos Environ* 42(20):4884–4903
- Janssen Stijn, Dumont Gerwin, Fierens Frans, Deutsch Felix, Maiheu Bino, Celis David, Trimpeneers Elke, Mensink Clemens (2012) Land use to characterize spatial representativeness of air quality monitoring stations and its relevance for model validation. *Atmos Environ* 59:492–500. <https://doi.org/10.1016/j.atmosenv.2012.05.028>. (ISSN 1352-2310)
- Open Street Map (n.d.) Project OpenStreetMap® is open data, licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF). <https://www.openstreetmap.org/copyright/en>. Accessed 01.09.2023
- Pereira GW, Valente DSM, Queiroz DMd, Coelho ALdF, Costa MM, Grift T (2022) Smart-Map: An Open-Source QGIS Plugin for Digital Mapping Using Machine Learning Techniques and Ordinary Kriging. *Agronomy* 12:1350. <https://doi.org/10.3390/agronomy12061350>
- Piersanti A, Ciancarella L, Cremona G, Righini G, Vitali L (2013) Rappresentatività spaziale di misure di qualità dell'aria. Valutazione di un metodo di stima basato su fattori oggettivi. Rapporto Tecnico RT/2013/1/ENEA, ENEA. https://www.researchgate.net/publication/280090841_Rappresentativita_spaziale_di_misure_di_qualita_dellaria-Valutazione_di_un_metodo_di_stima_basato_su_fattori_oggettivi/link/55a7d94108ae815a0420b816/download?tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19. Accessed 13.02.2024
- QGIS Development Team QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.org> (Access on 01.09.2023)
- R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>. Accessed 01.09.2023
- Regione Toscana Rappresentatività (2015) Spaziale delle stazioni della rete di monitoraggio della qualità dell'aria. https://www.regione.toscana.it/documents/10180/14975509/ARPAT_Lamma_Rappresentativita.pdf/5fdf6337-e9d3-4790-b041-9be51ab1f26f (Access on 01.09.2023)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.