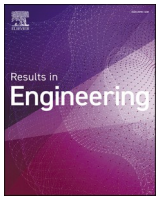
Results in Engineering 17 (2023) 100937

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering



Comparison of spatial interpolation techniques for innovative air quality monitoring systems

Nicoletta Lotrecchiano [a,b], Diego Barletta [a], Massimo Poletto [a], Daniele Sofia [a,b,*]

[a] Dipartimento di Ingegneria Industriale, Universitá Degli Studi di Salerno, Via Giovanni Paolo II 132, 84084, Fisciano, SA, Italy
[b] Sense Square Srl, Corso Garibaldi 33, Salerno, SA, 84123, Italy

ARTICLE INFO

Keywords:
PM10
Spatial expansion
Pollutant
Spatial model
Air quality

ABSTRACT

Different integration methods were tested to integrate data from a dynamic road network (ROM) in which pollution measurement sensors were mounted over delivery vans. Two methods were purposely developed, the isoelliptical expansion - ISOE - method accounting for the wind convective transport of pollutants and the modified isoelliptical expansion - MISOE - method in which, furthermore, local specific deviation of the pollution are estimated from historical sequences of pollution levels. The results obtained by these methods were compared with the well-known inverse distance weighted - IDW - method, which is only based on the distance from the interpolation sources. The comparison of the errors between the estimated values and the available measures reveals that the MISOE model provides more accurate estimated values with a low associated error. The ISOE model is more complicate than the IDW but provides better estimations in windy days. The maps of the local adjusting coefficients estimated month by months are able to identify critical areas to address in local environmental policy decisions.

1. Introduction

Improving air quality is the challenge of recent years to which institutions are dedicating many efforts and resources to implement suitable strategies (30,31],[1]). Great concern to the effects of air quality makes it necessary and useful to know the values of pollutants concentrations with a high spatial resolution [2]. In fact, traditional air quality measurement systems, are located in a few points on the territory and cannot provide complete information of the environmental condition. Therefore, to complement these systems, new monitoring networks, based on IoT technology, have been installed in recent years, implemented by individuals or groups of citizens in order to increase the available data spatial resolution [3,4]. Further data comes from dynamic road networks with high temporal resolution (ROM) that have been implemented in recent years in Italy [5]. ROM networks are an innovative technology developed to monitor the air quality in a dynamic on-road way [6]. This measuring device technology is designed to be located on a moving vehicle which, by moving along the area considered, allows the pollutants measurements [7]. Given the variety of air quality information sources, the creation of maps showing the pollutants concentrations to describe the area in a timely manner is necessary to continuously show the pollution levels, allowing to appreciate the studied parameter surface trend [29]. In these maps the territory is divided in cells each covering a portion of it. Each cell is associated to a value of the measured or estimated quantity to be represented and is characterized by the geographic coordinates associated to its centroid. The necessity to estimate represented values derives from the fact that the measured quantities deriving from ROM and fixed stations may be available only for a limited set of cells; and it may also happen that these are not uniformly distributed on the cell grid. To overcome this problem, it is possible to estimate missing data over the grid by using the available ones by using geographic interpolation methods [8]. These methods are special adaptations of more general spatial interpolation techniques that are procedures to evaluate the values assumed by a certain quantity on a continuous domain (surface, volume) starting from the known values (calculated or measured) in some points. The interpolation procedure can be carried out with a global approach, that makes use of all the other known values in the domain of interest to estimate interpolated values, or with a local approach, considering only the measure values in the neighbourhoods of the region where the variable is estimated [9]. Understandably, the reliability of any interpolating method will increase with the number of available data on the map and with their uniform

* Corresponding author. Dipartimento di Ingegneria Industriale, Universitá degli Studi di Salerno, Via Giovanni Paolo II 132, 84084, Fisciano, SA, Italy.
E-mail address: dsofia@unisa.it (D. Sofia).

https://doi.org/10.1016/j.rineng.2023.100937
Received 17 November 2022; Received in revised form 26 January 2023; Accepted 1 February 2023
Available online 4 February 2023

spatial distribution.

The Gaussian plume model is often used in dispersion models of variable complexity, starting from the simplest case, in which a single source is accounted for, and going up to more complex cases in which the presence of many sources is accounted for. The spatial pollutants in these models are affected by the influence of the parameters from a physical-meteorological point of view [10,11]. In fact, weather influence strongly the pollutants dispersion and the emissions related to the domestic heating [12,13]. The multisource plume model is among the most widely used models for the mathematical modeling of pollutants concentrations in urban areas. The multi-source model is then developed by superimposing the individual plumes from all the emissions considered. The multi-source Gaussian plume model adapted to interpolation of data, considers each single measured value as a pollution source dispersing into the environment according to a Gaussian plume distribution [14]. Sources are idealized as being large point sources discharging at different heights above ground level (Calder 1977). If the source physical properties are considered, such as height, flow rate and the variables associated with the emission such as gas velocity, temperature, etc., it is possible to combine all the sources, averaging data of the individual stacks according to a multi-component system mixing rules, in a single superstack. This type of model was used by Ref. [15] to predict SO_2 concentrations in an urban and industrial context in Kuwait, providing a better prediction in pollutant concentrations than traditional models [15]. More recent models, to evaluate the pollutants concentrations on the ground level, are based on the pollutant concentration values measured by air quality monitoring systems and tends to reconstruct and locate the pollution source, and are useful to describe the results of accidental contaminant releases [16–19].

Among the possible approaches that can be adopted to integrate missing data on the map there are the deterministic methods that calculate the missing values as a function of the values occurring in the neighboring points using a rule based on the system physics. These methods include the inverse of distances (Inverse Distance Weighted – IDW), the polynomial interpolations (Spline) and the polygons of influence (Nearest Neighbour Analysis – NNR). The general idea behind the use of polynomial interpolations is that each interval between data points, represents the graph with a simple function obtained by connecting data with lines. This function could be linear, quadratic or cubic according to data. The NNR method exploits the proximity between points and the influence between them to calculate missing values at points of interest. The IDW method is based on the concept that the correlation between data on the map is inversely related to the distance between points. It will be described in greater detail in the following. In general, these methods do not consider the statistical properties of the measured points, but assign a unique value to the function of interest, considering the variations of physical parameters in the domain. This implies sufficient knowledge of the surface to be modeled. With these model it is not possible to obtain error estimates on the results obtained.

The missing values can be estimated also by using a geostatistical model. The geostatistical methods (kriging) are based on the measurement of the values spatial autocorrelation, regardless of the possible physical meaning, and include an evaluation of the forecast error [20]. Geostatistics, in fact, studies the variables behavior, theirself and mutual spatial correlations, their structure, extracts the rules according to coherent models and uses them to carry out the operations that are required to solve specific problems. The methods of Geostatistics are applicable in all fields of applied sciences in which the phenomena of study are characterized by the spatiality. The Kriging methodology is well known in literature but there are very few works that implemented it for large datasets as those treated in this case, due the calculative burden increasing more than linearly with the dataset size. Ordinary kriging models are also employed to determine the spatial-temporal variety of annual and seasonal PM2.5 and PM10 concentrations [21].

Another possibility to estimate air quality in the missing cells, is to use measured data in the neighbour cells with purposely designed algorithms [22] that takes into account the meteorological parameters such as the wind dependence. In the present study, the results of the applications of few simple methods is assessed by comparing, in each cell where available data are present, the values predicted by the algorithm using the neighbour data and the experimental values. Different criteria will be adopted and compared. In some of them the influence of the neighbour cells on the prediction of the local concentration is based exclusively on the distance between the centre of the influencing cell and the centre of the tested cell, in others the direction and the speed of the wind are taken into consideration in the calculation procedure. The methods proposed involves various innovative approaches, in fact this study use air quality data provided by an extremely innovative technology really and currently measured in an Italian city. Moreover, the model is applied to a very large area. Finally, the developed method for the spatial estimation of missing data was compared with other techniques to evaluate the best solution.

Information deriving from air quality monitoring systems are used in applications to support smart cities in various sectors [23]. In detail, the proposed model will provide detailed information on air quality with a very high resolution, in this way it will be possible to analyse pollution at the local level, highlighting any critical issues in to propose solutions for pollution mitigation. Starting from the monitoring network located in any territory, the developed model is able to expand these values and thus obtain indications on the pollution levels to be able to implement targeted and concrete actions for the reduction of pollution on a local scale.

2. Materials and methods

2.1. Air quality data

The data used in this study includes all measurements made by the ROM network in Milan between January and June 2020 [24]. The ROM network was made of 53 measuring devices located on courier vans. The monitoring devices allowed the measurements of gas and suspended particulate matter. In this study the analysis will be limited to measures relative to PM10. The design of a monitoring network involves a preliminary phase of defining the installation positions. These positions can be defined with a modeling approach, using operational street pollution models [6].

The starting point is the subdivision of the ground surface in cells with a shape able to cover the whole ground surface over which the pollution is monitored by means of fixed or mobile sensors (e.g. Triangles, squares, or hexagons). Raw data associated with the time and position from which they are taken are used to calculate calculating hourly averages of the pollution within the cell $c_{i,j}$ for the cell i and the time j. Given a time j, and considering the random position of the ROMs the available data will not be able to provide complete sets of values $c_{i,j}$. Hence the need to apply spatial interpolation methods.

2.2. Methods definition

Three methods for spatial expansion of data were compared. The first method considered is called of isoelliptical expansion (ISOE) was firstly developed by Ref. [22] and assumes that a local variation of the concentration of pollutants with respect to the city average affects the neighborhood with changes in the local concentration in the same direction, and that the points on the ground affected by the same concentration changes are located along elliptic iso-concentration curves in which the major axis has the same direction of wind and that the ellipse elongation increases with the wind intensity. In this work, a modification to the ISOE method is proposed, to account for the local specific changes by considering the history of the pollution of each cell in a one-month long period (MISOE method). The effectiveness of these two models is tested against another frequently adopted method, which is the Inverse Distance Weighted model (IDW), in which the distance

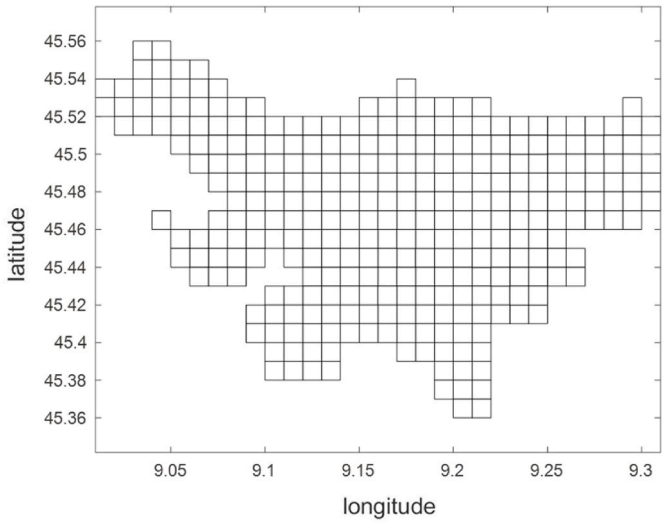


Fig. 1. Map of the area considered divided into a square grid.

between point is the only parameter affecting correlation, irrespective of the present atmospheric conditions. In the following, more details of the methods tested are reported.

For the description of the methods, it is assumed that the ground surface is dived in equal shaped cells (i.e. squares, hexagons, triangles) able to cover the whole surface and that a single value if each pollutant concentration is associated to the cell, ideally applied at the cell centroid (Fig. 1).

2.3. ISOE model

The first step of the ISOE method is the calculation of $C_{i,j}$, the variation of local concentration at time j, $c_{i,j}$, with respect to the background level at the same time j, $c_{bg,j}$, that is

$$C_{i,j} = c_{i,j} - c_{bg,j} \tag{1}$$

As above mentioned, the underlying idea is that any local pollution source determining a positive variation from the background level or any local pollution sink determining a negative variation from the background level will have the same effect on the neighborhood along closed iso effect lines that have the shape of ellipses, located on ground so that one of the two focal points is centered on the source, and that the major axis is oriented along the wind direction. It is also hypothesized that the ellipse elongation, that is the difference between the two axes, increases with the wind intensity. The ellipse axes are related to each other by the following system of equations:

$$\begin{cases} a_j^2 + b_j^2 = 2r_{0j}^2 \end{cases} \tag{2}$$

$$\left\{ a_j^2 = b_j^2 + \gamma r_{0j}^2 v_j^2 \right. \tag{3}$$

were a_j and b_j are the semi-major and semi-minor axes length at time j, v_j is the wind speed normalized with respect to the historical maximum monthly measured value. Time j, r_{0j} is the distance at which the decay along the ellipse is obtained for a wind speed equal to zero and γ is an adjustable model parameter relating the wind speed and the ellipse elongation. The parameter r_{0j} is associated to a decay factor α_j by means of a decay coefficient d_c:

$$\alpha_j = \exp\left(r_{0j} d_c\right) \tag{4}$$

Provided a second point that is placed at a distance r from the pollution source in a direction making an angle θ_j with the wind direction at time j, according to the ellipse equation in polar coordinates, it is:

$$r = \frac{a_j\left(1-\lambda_j^2\right)}{1-\lambda_j \cos\theta_j} \tag{5}$$

Where the ellipse elongation, λ, is defined as follows:

$$\lambda_j = \frac{\sqrt{a_j^2 - b_j^2}}{a_j} \tag{6}$$

Combining (2) and (3), it yields:

$$b_j^2 = a_j^2 \frac{2-\gamma v_j^2}{2+\gamma v_j^2} \tag{7}$$

accordingly, equation (6) becomes:

$$\lambda_j^2 = \frac{2\gamma v_j^2}{2+\gamma v_j^2} \tag{8}$$

which allows to calculate a_j from equation (6):

$$a_j = \frac{r\left(1-\lambda_j \cos\theta_j\right)}{1-\lambda_j^2} \tag{9}$$

As a result, using equations (2) and (7) it is possible to calculate r_{0j} to be used in equation (4) to derive the local decay factor α_j

$$r_{0j}^2 = a_j^2 \frac{2}{2+\gamma v_j^2} \tag{10}$$

Considering that the ground is divided in cells which are characterized by a different pollution concentration value, in principle all the cells may affect the concentration of the other neighbour cells. In particular, it is reasonable to consider that the effect could decrease with increasing distance from the cell according to a mathematical law for a weighing factor. Consequently, it can be assumed that the effect of the concentration in single cells non-adjacent to the one considered could have only a second order effect. For this reason, it was decided not to consider non-adjacent cells in the interpolation method adopted for the search of missing concentration values. Consequently, the estimated concentrations $CE_{i,j}$ in each target cell i at time j, are obtained as a linear combination of the neighbour cells of the target cell. The set of neighbour cells can be defined as follows: $i \in I$ where I is the set of cells; $K^{(i)}$ is the set of cell belonging to the neighborhood of the target i t. c.; x_i and y_i are the Cartesian coordinates of the i target cells centroids. Furthermore;

$$K^{(i)} = \{k \in I: |x_k - x_i| \le L_{xcell} \bigwedge |y_k - y_i| \le L_{ycell} \bigwedge k \ne i\} \tag{11}$$

where L_{xcell} and L_{ycell} are the characteristic projections in the x and y direction of the cell size. In this case study, square cells parallel to the coordinate axes are considered and, therefore, $L_{xcell} = L_{ycell} = L_{cell}$ is the length of the cell side. $C_{k,j}^{(i)}$ are the observed concentration at time j in the cell k belonging to $K^{(i)}$. The estimated concentrations in the target cell $CE_{i,j}$ were calculated from a weighted combination of the observed concentration $C_{k,j}^{(i)}$.

$$CE_{i,j} = \sum_{k \in K} C_{k,j}^{(i)} w_{k,j}^{(i)} \tag{12}$$

where $w_{k,j}^{(i)}$ is the weighting factor for the concentration contribution of the k cell to the i cell at time j calculated as follows:

$$w_{k,j}^{(i)} = \frac{\alpha_{k,j}^{(i)}}{\sum_{k \in K} \alpha_{k,j}^{(i)}} \tag{13}$$

From equation (13) it appears that for equally distant neighbour cells

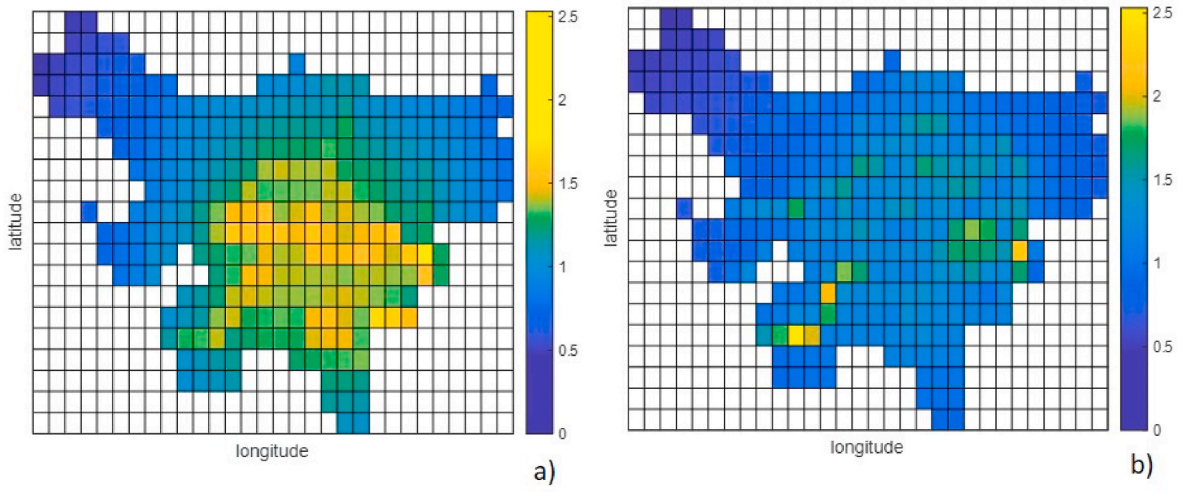


Fig. 2. Map of local monthly adjustment coefficients for PM10 for all the city cells: a) winter season, b) summer season.

from the target cells, the coefficient d_c used in equation (4) is irrelevant. In our case with a square network, the centroid of the neighbour cells along the diagonal are located a bit further from the centroid of target cells than the centroid of the cell on the sides. However, unless a strong decay coefficient, smaller than 1 km^{-1} is used, its value turned out to be scarcely affecting the final estimation value. For the calculations used in this paper, d_c was set to 1.57 km^{-1}. Finally, the estimated concentrations $ce_{i,j}$ are calculated from the inverse of equation (1):

$$ce_{i,j} = CE_{i,j} + c_{bg,j} \tag{14}$$

2.4. MISOE model

The MISOE model was obtained by modifying the previously developed ISOE model by introducing the local adjustment coefficient A_i to account for local variations of the measure which tend to keep the local measured value above or below the average and do not contribute to changing the neighbour values. In fact, each cell is representative of different urban conditions and has a typical pollution level. The use of local adjustment coefficient includes in the model the cell peculiarity. Since the availability of cell data depends on the specific day, the idea is that it might be possible to estimate these local adjustment coefficients considering a long enough time to have statistically representative values sequences but short enough to be still specific of the season considered. In order to calculate this coefficient, for each pollutant, it was considered the ratio between the local time monthly average, $M_{i,m}$, of the measured values $c_{i,j}$ values occurring in the cell i in that month, m, and the monthly average $M_{T,m}$ of all the $c_{i,j}$ values for the pollutant in all the cells within the town. Namely, for each pollutant, for the position in the cell i, the month coefficient is calculated:

$$A_{i,m} = \frac{M_{i,m}}{M_{T,m}} \tag{15}$$

The descriptive coefficient of each cell $A_{i,m}$ is estimated month by month from the historical information available for each cell on a seasonal basis considering a number N_m of months. The global coefficient for each cell is obtained from the monthly coefficients calculated as follows:

$$A_{i,s} = \sum_m \frac{A_{i,m}}{N_m} \tag{16}$$

For the winter season, the months of November, December and January were considered and, therefore $N_m = 3$ while for the summer season, the months of April, May and June were considered (also $N_m = 3$).

As a result, considering the local adjustment coefficient, the estimated concentration calculated by Eq. (3) is modified as follows:

$$ce_{i,j} = A_i * \left(\sum_{k \in K} \frac{C_{k,j}^{(i)} + c_{bg,j}}{A_k} w_{k,j}^{(i)} \right) \tag{17}$$

2.5. Inverse distance weighted model

The inverse distance weighted model (IDW) is a widely used method which is based on the assumption (realistic for many phenomena) that the properties in a given point are more similar to those of nearby points, compared to those of more distant points [25]. Everything is related to everything else, but near values have a larger effect. For practical reasons a finite number N of cells in a finite neighborhood is considered in the calculation, according to Tobler's first law.

$$ce_{i,j} = \frac{\sum_{l=1}^{N} c_l r_{i,l}^{-2}}{\sum_{l=1}^{N} r_{i,l}^{-2}} \tag{18}$$

In this case the weights that will determine the value of the unknown position are inversely proportional to the square of the distance $r_{i,l}$ of the l -th point considered from the target cell i. The larger the distance, the less is the effect that points distant from the point where the concentration is unknown will have during the interpolation process. Although the assumptions underlying the IDW method are simple, in applying this method it is good to consider that:

- The power and the number of source points/radius to be interpolated is arbitrary.
- Estimates of uncertainty and possible errors are not carried out.
- The surfaces obtained may contain numerous peaks or holes (bulls' eyes) around the known data. The spatial arrangement of the samples is not taken into consideration: isolated points and clusters have the same importance.

The power used in Equation (18) can assume different values depending from the real-life case [26] and according to that the value used in this work is 2. The IDW method is widely used for the pollutants spatialization [27] also for the assessment of human exposure to pollution [28].

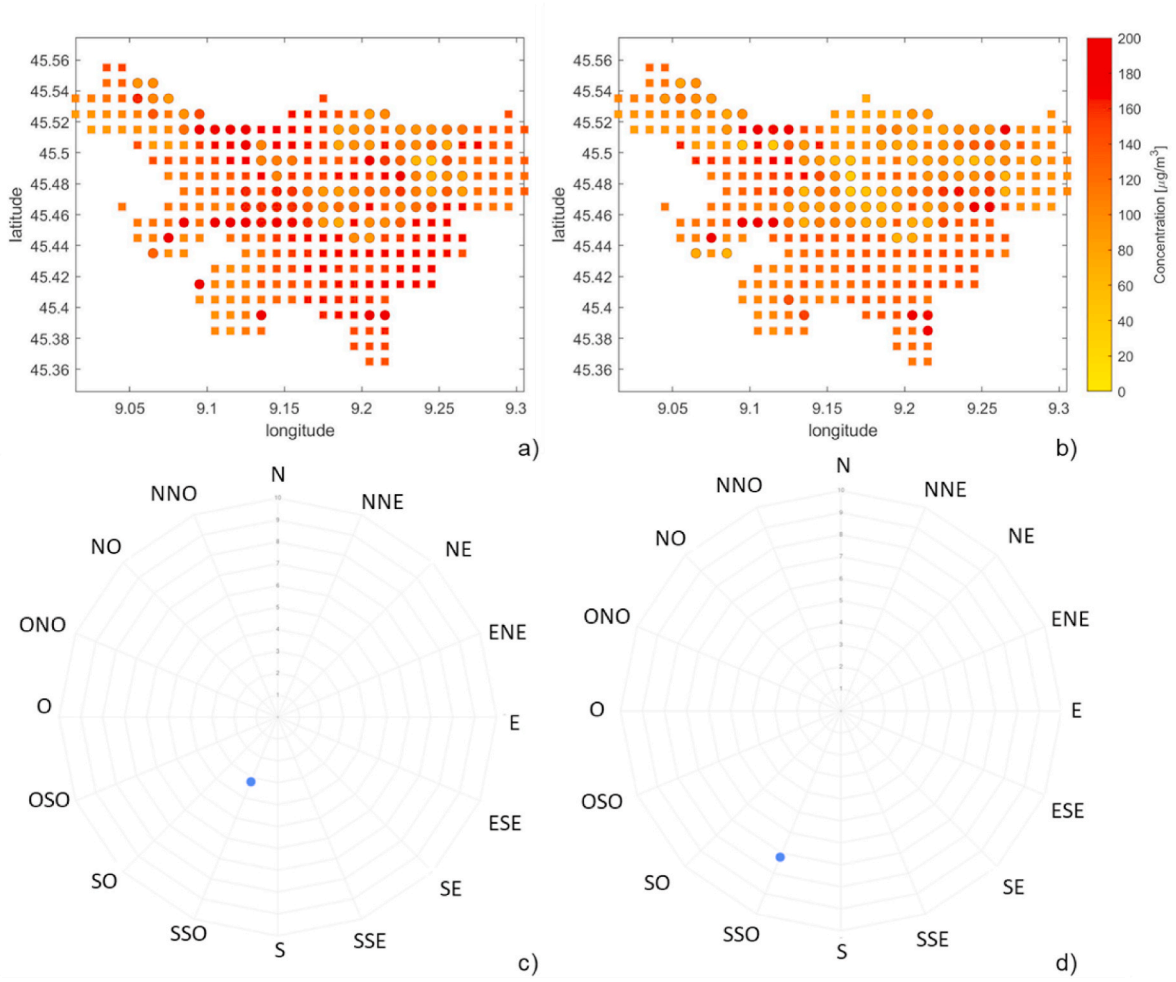


Fig. 3. PM10 daily average concentrations obtained in Milan by the ROM network during a) 13th (wind of 3.2 km/h from 202.5°) and b) 17th (wind of 6.5 km/h from 247.5°) January 2020. Panels c) and d) report the wind intensities and directions. Circles represent measured values while triangles represent estimated values.

2.6. Error

All the proposed spatial expansion models were compared by calculating the normalized mean square error (NRMSE) estimating the values of the pollutant concentration, $ce_{i,j}$ in all the N cell were a measured value $c_{i,j}$ was available. Finite periods including T time values j values, it is:

$$NRMSE = \sqrt{\sum_{j=1}^{T}\sum_{i=1}^{N} n_{i,j}\left(\frac{ce_{i,j}-c_{i,j}}{c_{i,j}}\right)^2 \Big/ \sum_{j=1}^{T}\sum_{i=1}^{N} n_{i,j}} \quad (19)$$

where $n_{i,j}$ assumes value 1 if the measured value $c_{i,j}$ exists, else it is 0.

3. Results and discussion

One of the results obtained from the data analysis using the MISOE model is the local adjustment coefficient on the area considered. The coefficients were reported directly on the map divided in square cells by means of a chromatic scale for two of the seasons considered (winter and summer in the specific case) for PM10. The data considered do not include values for the weekdays of Saturday and Sunday. In fact, the ROM network does not work on weekends and on other holidays, when the vehicles do not travel (like the 1st of January).

As it can be seen in Fig. 2, in the winter season, the coefficients relating to PM10 provide higher values in the city centre and decrease moving towards the periphery of the town. This indicates the presence of a background pollution level that could be represented by domestic heating which, going towards the city centre, is also affected by the presence of other activities. The city centre, in the specific case of Milan, is affected by heavy daily traffic during working days and therefore favourable to the production and accumulation of pollutants. Nevertheless, the presence of some green areas located in the city centre allows to have lower pollution levels and consequently lower adjustment coefficients in some cells. Analysing the map of the adjustment coefficients in the summer season (Fig. 2), it can be seen that, differently from the winter season, there is no prevailing trend in the spatial distribution of the coefficient values. These assume lower values than in winter for the majority of the cells, and few hot spots can be identified. These spots correspond to ground cells, such as the Navigli, characterized by intense summer activities which correspond to high vehicular traffic intensity determining the consequent increase in the local pollution levels. In general, the orography of Milan, similarly to the whole river Po Valley, does not favour the dispersion of pollutants and, thus, the tendency of airborne pollutants to stagnate in the atmosphere explains the high measured values of concentration.

The whole MISOE model developed was applied to the PM10 values measured by the ROM network in Milan considering days with different climatic and vehicular traffic conditions. In particular, four days of January 2020 (Figs. 3 and 4) were considered as an example of winter conditions and one day of May 2020 as an example of summer conditions (Fig. 5). The maps obtained report the measured values as circles while the estimated values are represented by triangles. PM10 concentrations are reported on a red scale if they are distributed in a high

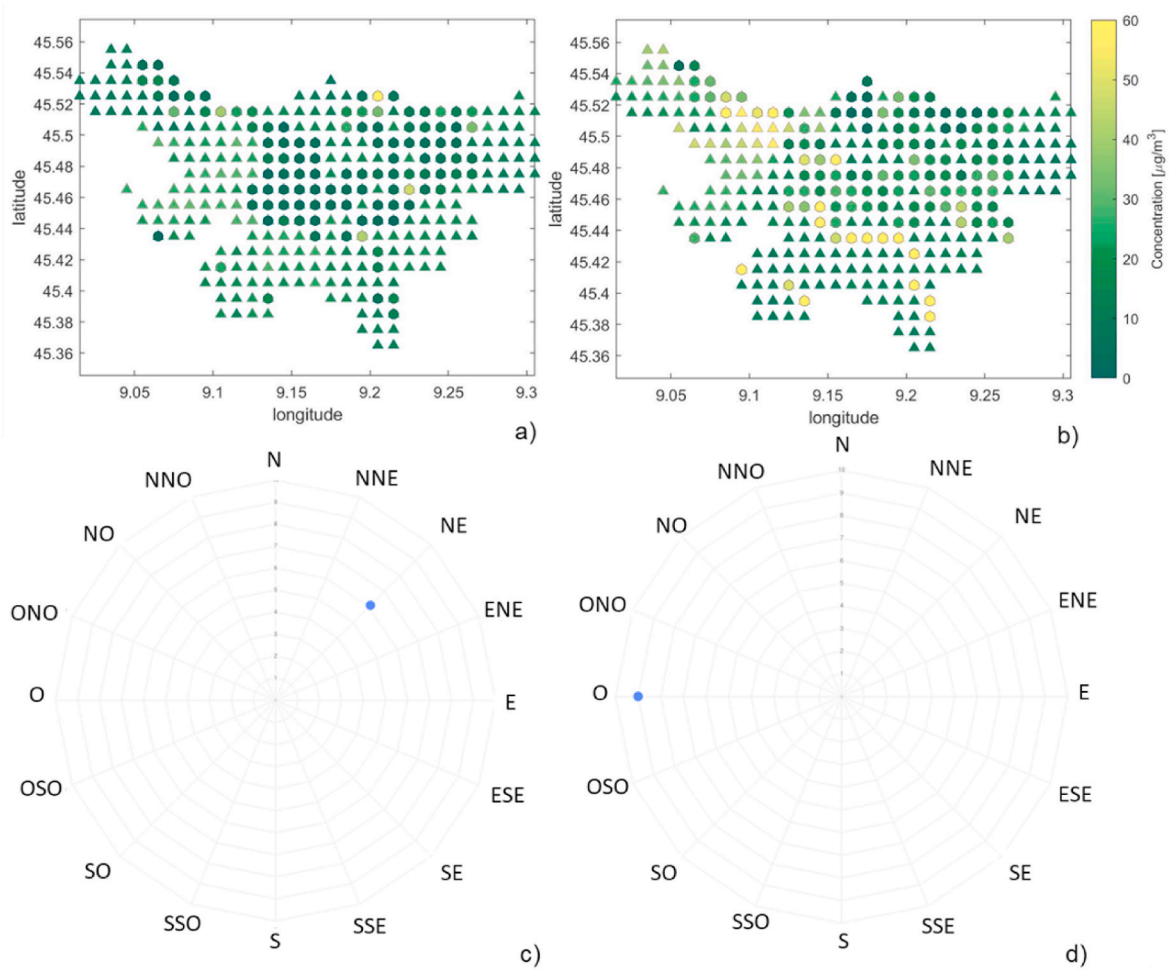


Fig. 4. PM10 daily average concentrations obtained in Milan by the ROM network during a) 20th (wind of 6.1 km/h from 45° with rain) and b) 29th (wind of 9 km/h from 270°) January 2020. Panels c) and d) report the wind intensities and directions. Circles represent measured values while triangles represent estimated values.

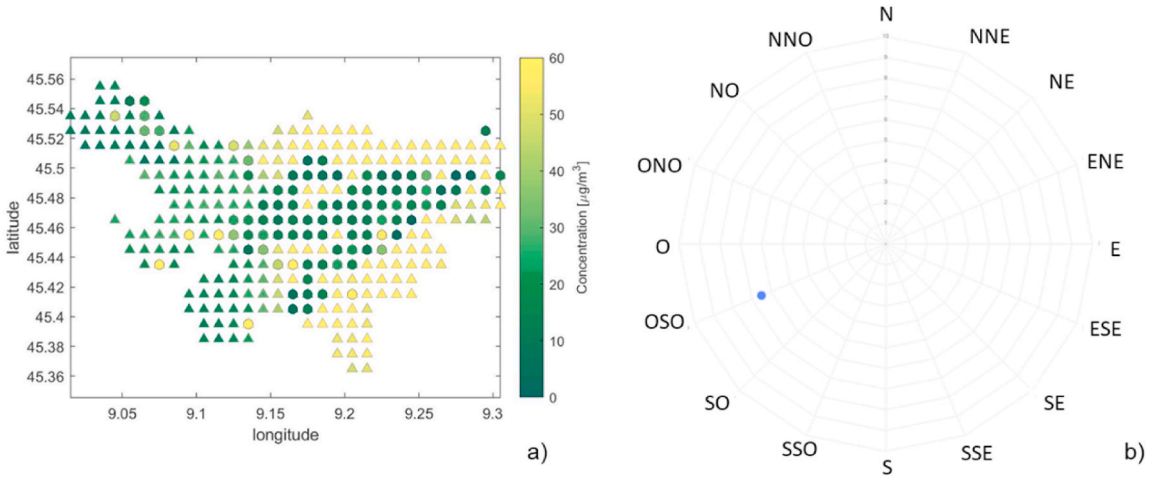


Fig. 5. PM10 daily average concentrations obtained in Milan by the ROM network during May 28, 2020 (wind of 7.2 km/h from 202.5°). Panel b) reports the wind intensity and direction. Circles represent measured values while triangles represent estimated values.

concentration range and on a green scale if they are distributed in a low concentration range. Data maps shown in Fig. 3 correspond to two days characterized by a high vehicular traffic which considerably contributed to the measured PM10 values. The map reported in Fig. 3a shows a rather homogeneous distribution of concentrations, which can be explained with the low wind speed (32.4 km/h from 202.5°) favouring the pollutants stagnation and accumulation. In the map of Fig. 3b, instead, the distribution of concentration values is more affected by the influence of a slightly more intense wind (6.5 km/h from 247.5°) which contributes a little to the dispersion of pollutants while maintaining

Table 1
NRMSE for the model analyzed.

NRMSE				
	Number of measured data	IDW	ISOE	MISOE
January	2265	0.45	0.4	0.17
May	2223	0.4	0.35	0.12
January wind <1.5 m/s	729	0.3	0.25	0.16
January wind >1.5 m/s	1536	0.35	0.33	0.14
May wind <2.5 m/s	1182	0.4	0.42	0.20
May wind >2.5 m/s	1041	0.6	0.57	0.25

pretty high concentration values.

The PM10 values shown in Fig. 4 are low as they correspond to rainy days (Fig. 4a) and windy days with high intensity (Fig. 4b) with moderate vehicular traffic intensity. The pollution mitigation effect appears to be more significant due in the presence of rain rather than with intense wind.

During the spring/summer period considered (May), wind intensity values higher than those recorded in January and the absence of domestic heating determined lower pollutants concentrations.

In the representative map of 28/05, reported in Fig. 5, it can be noticed how the dispersion direction of PM10 is aligned with the wind direction (SSW). The medium-high intensity wind speed and the high vehicular traffic intensity determine PM10 daily averages between 50 and 60 $\mu g/m^3$ in some cells; these values are above the alarm threshold defined by D. Lgs. 155/2010.

The NRMSE results reported in Table 1 allow a direct comparison of the performance of the examined models. In fact, inspection of the Table clearly indicates that the modified isoellipse model (MISOE) has the lowest NRMSE values. In general, the errors increase as the wind intensity increases because the greater dispersion affects the estimates that are less precise. At low wind intensity values, the errors associated with the IDW and the ISOE models are similar and are slightly larger than the error associated with the MISOE model. It appears that the MISOE model allows to have a better estimate of the missing data as it contains the information of the historical pollution of the cell.

Each measured data was estimated using the proposed models (ISOE, MISOE and, IDW models). In particular, the PM10 measured and estimated concentrations trend obtained with the three models proposed for a cell for the entire month of January was reported. The analysis of Fig. 6 shows how the estimated values with all the models follow the trend of the measured values. Furthermore, the estimated values obtained with the MISOE model are closest to the measured values in line with the low NRMSE value shown in Table 1. Although all three proposed models consistently estimate PM10 values, the MISOE model provides the values that are closest to those measured.

Given the high spatial resolution of the system considered, it was possible to analyse the pollutants trends, in particular of PM10, for some particularly interesting cells highlighted in Fig. 7. Four cells representative of the territory have been identified. Cell A corresponds to the city centre, affected by all those activities mainly related to tourism and vehicular traffic due to public transport. Cell B corresponds to the industrial area, where a large number of manufacturing activities exist. The other considered cells are representative of the peripheral area located in the North West of the city (C), and of a Southern area characterized by an intense vehicular traffic (D).

From the maps obtained on a daily basis, reported in Fig. 8, it can be seen how the local adjustment coefficients vary according to the conditions. This makes possible some interpretation on the link between causes and effects. For instance, on some days, higher pollution concentration located in the northern part of the city can be noticed. In particular, the map in Fig. 8 reports the highest values of the local adjustment coefficients and a higher pollutant concentration area localized in a specific cell. This occurrence could be due to local conditions increasing the pollution levels with respect to the rest of the city. Considering the cell corresponding to the city centre (cell A), as reported in Fig. 7, it can be seen that it exhibits medium-high concentration values throughout the month.

The same behaviour is observed for the cell corresponding to the industrial area (cell B in Fig. 7). With regards to the cell located on the north-western periphery (cell C in Fig. 7), the local adjustment coefficients assume medium-low values throughout the month of January. The area with high vehicular traffic intensity (cell D in Fig. 7) exhibits high coefficients which indicate the high activity of the area especially on Friday. On a weekly basis it is observed that during Fridays the cell is always greater than the other cells; this highlights both the flow of cars

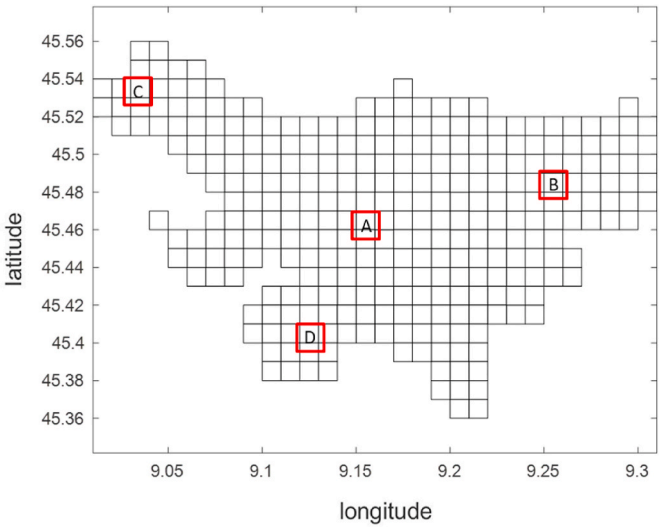


Fig. 7. Location of cells representative of different city areas.

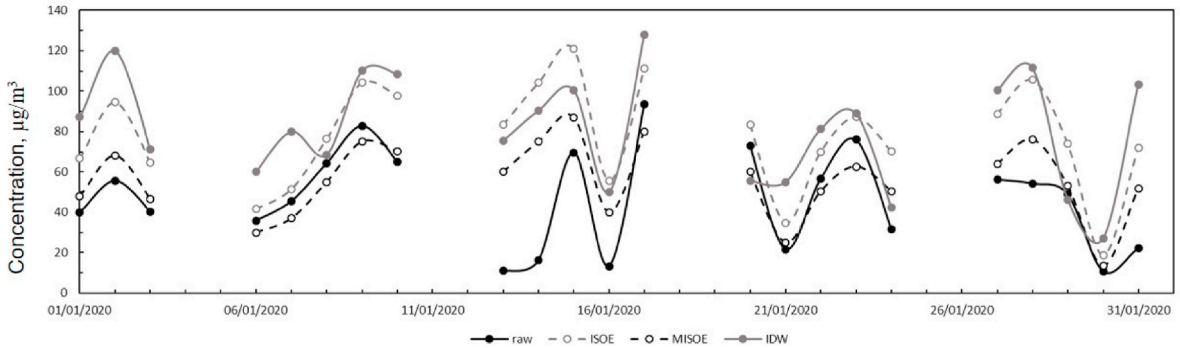


Fig. 6. PM10 concentrations measured (solid black line) and estimated with the ISOE (dashed grey line), MISOE (black dashed line) and IDW (grey solid line) models for a cell during January 2020.

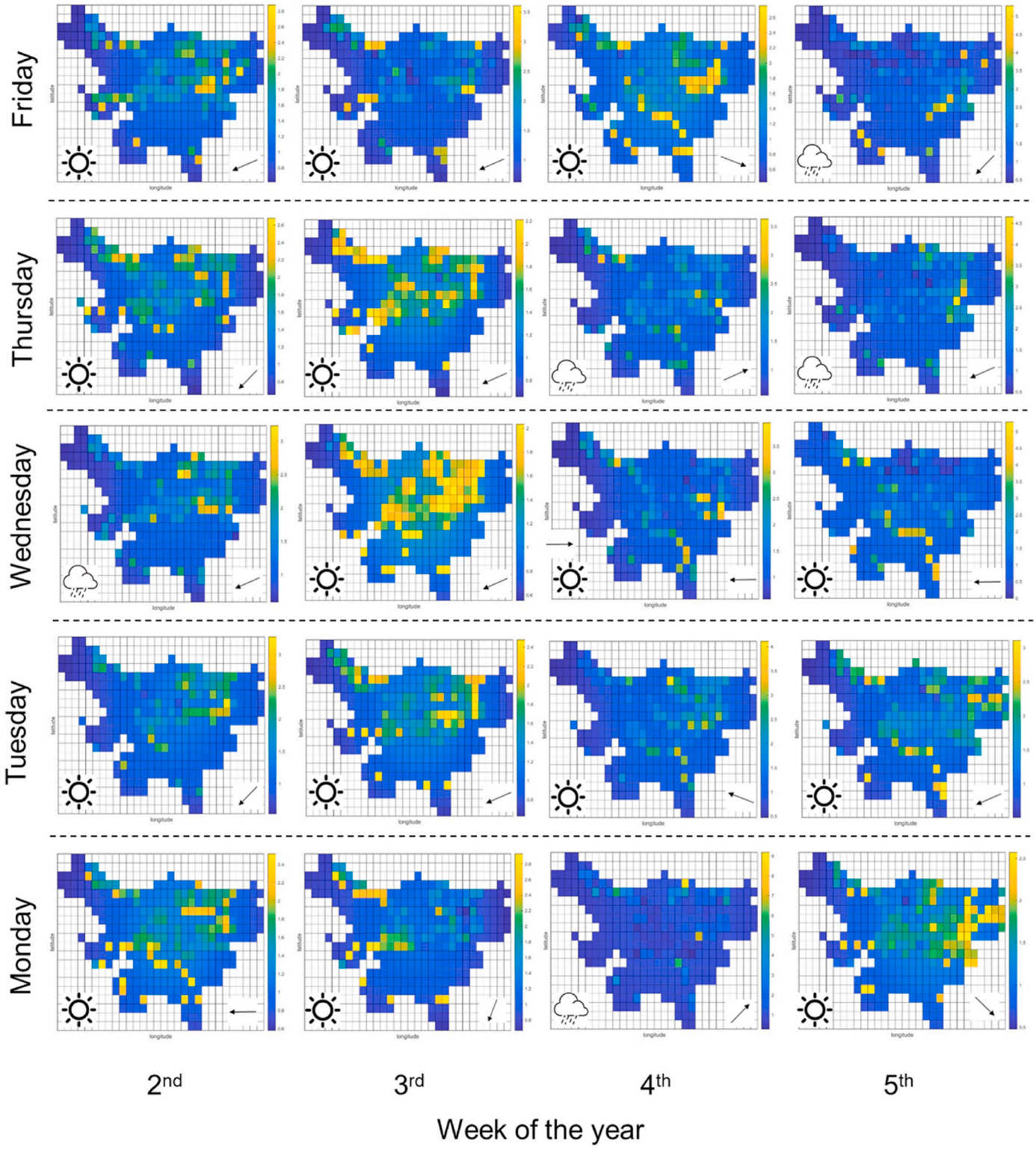


Fig. 8. Map of the local adjustment coefficients calculated on a daily basis for the 2nd to 5th week of the 2020.

leaving the city for the weekend and greater circulation due to the afternoon-evening activities that characterize the area.

Fig. 9 reports the trends of the local adjustment coefficients calculated on a monthly basis between January and May 2020 for the four cells highlighted in Fig. 7. In January the lowest coefficient is obtained for the NW periphery while the highest are in the city centre and in the high intensity traffic zone. In May, the coefficient of the high intensity traffic zone increases and the other coefficients are on average lower values respect to January. It can be noticed that in all cases a minimum is observed in the month of March. This result is in agreement with the results of the analysis of the pollution trend during the spring 2020 lockdown [22]. In fact, in the period between January and March there is an average 70% decrease of the local adjustment coefficients. Moreover, it appears that the coefficients in May are still lower than those in

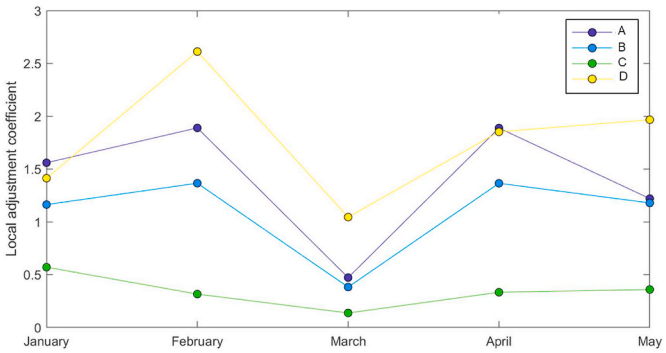


Fig. 9. Local adjustment coefficients for the A-D squares considered in Fig. 7 for January–May 2020.

January, with the exception of the area D with high vehicular traffic intensity.

4. Conclusion

It was possible to integrate data from a dynamic road network (ROM) in which pollution measurement sensors were mounted over delivery vans driving in Milan. Different integration methods were tested. Purposely developed ISOE and MISOE models were compared with the IDW method reported in the literature. In windy days, the ISOE model, that accounts for the effect of wind, provides lower error values than the IDW method, in which only distances are accounted for. The MISOE model accounting also for local specific deviations of pollution source provides more accurate estimated values with a low associated error. In fact, in this method, the introduction of the local adjusting coefficients is able to take into account the local characteristics of the pollution on the basis of historical trends. And This allowed the improvement of the performance of the ISOE model; in fact, the MISOE model provided estimated values that are much more similar to those effectively measured. These analysis features maximize the potentialities of pollution measurement networks characterized by high space and time resolution. Furthermore, the analysis of the distribution of the local adjusting coefficients could be useful to highlight the most critical areas with respect to pollution level and may be of help to policy makers for the implementation of pollution mitigation policies.

Credit author statement

Daniele Sofia: Conceptualization: Ideas; formulation or evolution of overarching research goals and aims, Methodology: Development or design of methodology; creation of models, Writing - Review & Editing Preparation: creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre-or postpublication stages, Supervision Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team, Nicoletta Lotrecchiano: Data Curation: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse, Writing - Original Draft: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation), Diego Barletta: Visualization Preparation, creation and/or presentation of the published work, specifically visualization/data presentation, Massimo Poletto: Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection, Methodology: Development or design of methodology; creation of models.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] A. Aslam, M. Ibrahim, A. Mahmood, M. Mubashir, H.F.K. Sipra, I. Shahid, S. Ramzan, M.T. Latif, M.Y. Tahir, P.L. Show, Mitigation of particulate matters and integrated approach for carbon monoxide remediation in an urban environment, J. Environ. Chem. Eng. 9 (4) (2021), 105546, https://doi.org/10.1016/j.jece.2021.105546.

[2] S. Gurram, A.L. Stuart, R.A. Pinjari, Agent-based modeling to estimate exposures to urban air pollution from transportation: exposure disparities and impacts of high-resolution data, Comput. Environ. Urban Syst. 75 (2019) 22–34, https://doi.org/10.1016/j.compenvurbsys.2019.01.002.

[3] W. Jiao, G.S.W. Hagler, R.W. Williams, R.N. Sharpe, L. Weinstock, J. Rice, Field assessment of the village green project: an autonomous community air quality monitoring system, Environ. Sci. Technol. 49 (10) (2015) 6085–6092, https://doi.org/10.1021/acs.est.5b01245.

[4] B. Sirmacek, R. Vinuesa, Remote sensing and AI for building climate adaptation applications, Results in Engineering 15 (2022), 00524, https://doi.org/10.1016/j.rineng.2022.100524.

[5] D. Suriano, A porTable air quality monitoring unit and a modular, flexible tool for on-field evaluation and calibration of low-cost gas sensors, Hardware 9 (2021), E00198, https://doi.org/10.1016/j.ohx.2021.e00198.

[6] D. Sofia, N. Lotrecchiano, A. Giuliano, D. Barletta, M. Poletto, Optimization of number and location of sampling points of an air quality monitoring network in an urban contest, Chemical Engineering Transactions 74 (2019) 277–282, https://doi.org/10.3303/CET1974047.

[7] T. Salthammer, C. Fauck, A. Omelan, S. Wientzek, E. Uhde, Time and spatially resolved tracking of the air quality in local public transport, Sci. Rep. 12 (2022) 3262, https://doi.org/10.1038/s41598-022-07290-5.

[8] J. Ma, Y. Ding, J.C.P. Cheng, F. Jiang, Z. Wan, A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5, J. Clean. Prod. 237 (2019), https://doi.org/10.1016/j.jclepro.2019.117729.

[9] S. Shao, F. Qin, M. Xu, Q. Liu, Y. Han, Z. Xu, Temporal and spatial variation of refractive index structure coefficient over South China sea, Results in Engineering 9 (2021), 100191, https://doi.org/10.1016/j.rineng.2020.100191.

[10] C. Tee, E.Y.K. Ng, G. Xu, Analysis of transport methodologies for pollutant dispersion modelling in urban environments, J. Environ. Chem. Eng. 8 (4) (2020), 103937, https://doi.org/10.1016/j.jece.2020.103937, 2020.

[11] F. Yousefian, S.S. Faridi, F. Azimi, M. Aghaei, M. Shamsipour, K. Yaghmaeian, M. S. Hassanvand, Temporal variations of ambient air pollutants and meteorological influences on their concentrations in Tehran during 2012–2017, Sci. Rep. 10 (2020) 292, https://doi.org/10.1038/s41598-019-56578-6.

[12] R.R. Tan, K.B. Aviso, A linear program for optimizing enhanced weathering networks, Results in Engineering 3 (2019), 100028, https://doi.org/10.1016/j.rineng.2019.100028.

[13] M. Kuosa, P. Kiviranta, H. Sarvelainen, E. Tuliniemi, T. Korpela, K. Tallinen, H. K. Koponen, Optimisation of district heating production by utilising the storage capacity of a district heating network on the basis of weather forecasts, Results in Engineering 13 (2022), 100318, https://doi.org/10.1016/j.rineng.2021.100318.

[14] S. Gustafson, K.O. Kortanek, J.R. Sweigart, Numerical optimization techniques in air quality modeling: objective interpolation formulas for the spatial distribution of pollutant concentration, J. Appl. Meteorol. Climatol. 16 (12) (1977) 1243–1255, https://doi.org/10.1175/1520-0450(1977)016<1243:NOTIAQ>2.0.CO;2.

[15] N.A. Ahmad, W. Bouhamra, Development of a new long term multiple-source plume model: application on some industrial and residential areas in Kuwait, Environ. Prog. 12 (1993) 123–127, https://doi.org/10.1002/ep.670120209.

[16] G. Turbelin, S.K. Singh, J.-P. Issartel, Reconstructing source terms from atmospheric concentration measurements: optimality analysis of an inversion technique, Journal of Advanced Modelling Earth Systems 6 (2014) 1244–1255, https://doi.org/10.1002/2014MS000385.

[17] Y.O. Khaniabadi, P. Sicard, A.M. Taiwo, A. De Marco, S. Esmaeili, R. Rashidi, Modeling of particulate matter dispersion from a cement plant: upwind-downwind case study, J. Environ. Chem. Eng. 6 (2) (2018) 3104–3110, https://doi.org/10.1016/j.jece.2018.04.022.

[18] N. Lotrecchiano, D. Sofia, A. Giuliano, D. Barletta, M. Poletto, Pollution dispersion from a fire using a Gaussian plume model, International Journal of Safety and Security Engineering 10 (2020) 431–439, https://doi.org/10.18280/ijsse.100401.

[19] M. Othman, M.T. Latif, H.H.A. Hamid, R. Uning, T. Khumsaeng, W. Phairuang, Z. Daud, J. Idris, N.M. Sofwan, S.C.C. Lung, Spatial–temporal variability and health impact of particulate matter during a 2019–2020 biomass burning event in Southeast Asia, Sci. Rep. 12 (2022) 7630, https://doi.org/10.1038/s41598-022-11409-z.

[20] G. Aversano, G. D'Alessio, A. Coussement, F. Contino, A. Parente, Combination of polynomial chaos and Kriging for reduced-order model of reacting flow applications, Results in Engineering 10 (2021), 100223, https://doi.org/10.1016/j.rineng.2021.100223.
[21] B. Karimi, B. Shokrinezhad, Spatial variation of ambient PM2.5 and PM10 in the industrial city of Arak, Iran: a land-use regression, Atmos. Pollut. Res. 12 (12) (2021), 101235, https://doi.org/10.1016/j.apr.2021.101235.
[22] N. Lotrecchiano, D. Sofia, A. Giuliano, D. Barletta, M. Poletto, Spatial interpolation techniques for innovative air quality monitoring systems, Chemical Engineering Transactions 86 (2021) 391–396, https://doi.org/10.3303/CET2186066.
[23] I.V. Yarmoshenko, A.D. Onishchenko, G.P. Malinovsky, A.V. Vasilyev, M. V. Zhukovsky, MODELING and justification of indoor radon prevention and remediation measures in multi-storey apartment buildings, Results in Engineering 16 (2022), 00754, https://doi.org/10.1016/j.rineng.2022.100754.
[24] N. Lotrecchiano, D. Sofia, A. Giuliano, D. Barletta, Poletto, M. Real-time on-road monitoring network of air quality, Chemical Engineering Transactions 74 (2019) 241–246, https://doi.org/10.3303/CET1974041.
[25] P.M. Bartier, C.P. Keller, Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW), Comput. Geosci. 22 (1996) 795–799, https://doi.org/10.1016/0098-3004(96)00021-0.
[26] de Mesnard Louis, Pollution models and inverse distance weighting: some critical remarks, Comput. Geosci. 52 (2013) 459–469, https://doi.org/10.1016/j.cageo.2012.11.002.
[27] J.E. Hart, J.D. Yanosky, R.C. Puett, L. Ryan, D.W. Dockery, T.J. Smith, E. Garshick, F. Laden, Spatial modeling of PM10 and NO2 in the continental United States, 1985–2000, Environ. Health Perspect. 117 (11) (2009) 1690–1696, https://doi.org/10.1289/ehp.0900840.
[28] J. Hooyberghs, M. Clemens, D. Gerwin, F. Frans, Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium, J. Environ. Monit. 8 (11) (2006) 1129–1135, https://doi.org/10.1039/B612607N.
[29] D. Hasenfratz, O. Saukh, C. Walser, C. Hueglin, M. Fierz, T. Arn, J. Beutel, L. Thiele, Deriving high-resolution urban air pollution maps using mobile sensor nodes, Pervasive Mob. Comput. 16 (2014), https://doi.org/10.1016/j.pmcj.2014.11.008.
[30] N. Lotrecchiano, P. Trucillo, D. Barletta, M. Poletto, D. Sofia, Air pollution analysis during the lockdown on the city of milan, Processes 9 (2021) 1692, https://doi.org/10.3390/pr9101692.
[31] D. Sofia, F. Gioiella, N. Lotrecchiano, A. Giuliano, Mitigation strategies for reducing air pollution, Environ. Sci. Pollut. Res. 27 (2020) 19226–19235, https://doi.org/10.1007/s11356-020-08647-x.