ENVIRONMENTAL POLLUTION 266 (2020) 115279

https://doi.org/10.1016/J.ENVPOL.2020.115279

A multi-pollutant methodology to locate a single air quality monitoring station in small and medium-size urban areas

Marta Doval Miñarro[a,*], Daniel Bañón[a,b], José A. Egea[b], Isabel Costa-Gómez[a], Antonia Baeza Caracena[c]

[a] Technical University of Cartagena (UPCT), Department of Chemical and Environmental Engineering, Paseo Alfonso XIII, 52; E-30203 Cartagena (Murcia) SPAIN.
[b] Center for Edaphology and Applied Biology of Segura (CEBAS-CSIC). Campus Universitario de Espinardo, 30100, Murcia, SPAIN
[c] University of Murcia (UMU), Department of Chemical Engineering, Faculty of Chemistry, 30100, Murcia, SPAIN
Corresponding author: Tel.: +34 968325552; email address: marta.doval@upct.es

Abstract

Air quality management is underpinned by continuous measurements of concentrations of target air pollutants in monitoring stations. Although many approaches for optimizing the number and location of air quality monitoring stations are described in the literature, these are usually focused on dense networks. However, there are small and medium-size urban areas that only require one monitoring station but also suffer from severe air pollution. Given that target pollutants are usually measured at the same sampling points; it is necessary to develop a methodology to determine the optimal location of the single station. In this paper, such a methodology is proposed based on maximizing an objective function, that balances between different pollutants measured in the network. The methodology is applied to a set of data available for the city of Cartagena, in southeast Spain. A sensitivity analysis reveals that 2 small areas of the studied city account for 80% of the optimal potential locations, which makes them ideal candidates for setting up the monitoring station. The methodology is easy to implement, robust and supports the decision-making process regarding the siting of fixed sampling sites.

Capsule

This paper describes an easy-to-use methodology to find out the optimum location for setting up a single air quality monitoring station, especially useful for small and medium-size urban areas.

Keywords**:** air quality, monitoring stations, optimization, small urban areas, medium-size urban areas

1. Introduction

Monitoring pollutant concentrations in air is pivotal to properly manage air quality. European Directive 2008/50/EC (European Commission, 2008) on ambient air quality and cleaner air for Europe establishes the minimum number of sampling points for fixed measurements of concentrations of sulfur dioxide (SO_2), nitrogen dioxide (NO_2) and nitrogen oxides (NO_x), lead, benzene and carbon monoxide (CO) in ambient air as a function of the population of an agglomeration or zone in its Annex V, amended by Commission Directive 2015/1480. Similar requirements for fixed measurements of ozone can be found in Annex IX of European Directive 2008/50/EC, also modified by Commission Directive 2015/1480. In Table 1, these numbers are gathered. Requirements are different if maximum concentrations measured are above the upper assessment threshold, or if they are in between the upper and lower assessment thresholds, except for ozone where the number of fixed monitoring points depends on the type of agglomeration or zone.

Table 1.- Minimum number of sampling points for fixed measurements of SO_2, NO_2 and NO_x, lead, benzene, CO and O_3 to assess compliance with limit values for the protection of human health (European Commission, 2008).

Population of agglomeration or zone (thousands)	Fixed measurements of SO_2, NO, NOx, lead, benzene and CO		Fixed measurements of O_3	
	If maximum concentrations exceed the upper assessment threshold	If maximum concentrations are between the upper and lower assessment thresholds	Agglomeration	Other zones
0-249	1	1		1
250-499	2	1	1	2
500-749	2	1	1	2
750-999	3	1	1	2
1000-1499	4	2	3	3
1500-1999	5	2	3	4
2000-2749	6	3	4	5
2750-3749	7	3	5	6
3750-4749	8	3	One additional station per 2 million inhabitants	
4750-5999	9	4		
≥ 6000	10	4		

A few instructions are also given regarding the macroscale selection of monitoring sites. Thus, for protection of human health, sampling points *shall be sited in such a way as to provide data on the following*:

— the areas within zones and agglomerations where the highest concentrations occur to which the population is likely to be directly or indirectly exposed for a period which is significant in relation to the averaging period of the limit value(s),

— levels in other areas within the zones and agglomerations which are representative of the exposure of the general population.

Achieving both of these requirements is fairly difficult for zones or agglomerations where just a single sampling point is required (González Ferradás et al., 2010). In these cases, the final location is eventually selected most of the times based on the judgement of the managers of the air quality monitoring network of that area. Something similar happens in denser networks, which are not usually designed as a comprehensive whole. A number of works in the literature have aimed at redistributing existing stations in order to optimize the information and the cost (Ainslie et al., 2009; Al-adwani et al., 2015; Andó et al., 1999; Pires et al., 2008a, 2008b; Wu et al., 2011). In some cases, a lower number of stations than the existing ones is required to properly characterize an area without any redundancies. In other works, however, the coverage area needs to be increased by installing one or more stations. In other studies, the dimension of the network is appropriate but a relocation of the stations is needed.

Different approaches have been used to select fixed monitoring sites. As a starting point, most of them imply knowing the spatial distribution of pollutant concentrations over the study area. These can be attained, for instance, by means of the data provided by an already available dense air monitoring network (Ainslie et al., 2009; Andó et al., 1999; Orlowski et al., 2017; Pahlavani et al., 2017; Pope and Wu, 2014; Wu et al., 2011). Moreover, passive sampling campaigns (González Ferradás et al., 2010; Lozano et al., 2010; Orlowski et al., 2017; Soares et al., 2018) and modelling have commonly been used. Both of these strategies allow for describing the study area with a high spatial resolution at a low cost.

Among the most used air dispersion models stand out the use of the multiple cell approach (Al-adwani et al., 2015; Elkamel et al., 2008; Zoroufchi Benis et al., 2016), the WRF-CALPUFF (Hao and Xie, 2018), the WRF-CMAQ (Araki et al., 2015) or Gaussian plume models to estimate pollutant concentrations coming from point sources (Chen et al., 2006; Corti and Senatore, 2008; Mazzeo and Venegas, 2008; Mofarrah and Husain, 2010). Other methodologies take advantage of the correlation between some atmospheric properties such us turbidity and pollutant concentrations (Sarigiannis and Saisana, 2008).

Once this information is available, the spatial resolution of the grid can be increased by kriging or other geostatistical methods that interpolate concentrations. At this point, the selection of the monitoring sites can be done under different approaches. Some studies have directly selected sites that fulfil the Directive's macroscale siting requirements (González Ferradás et al., 2010; Lozano et al., 2010); whereas others have used optimization algorithms (Chen et al., 2006; Mofarrah and Husain, 2010; Sarigiannis and Saisana, 2008).

A variety of approaches can be found in the literature obtained from the combination of a single or multipollutant approach with a single or multi-objective optimization. For the most relevant works found in the literature regarding air quality network design, Table 2 shows for each of them, if it is single or multipollutant-oriented, how the spatial information on pollutant concentration has been obtained, and which techniques have been used to select the final location of sampling sites, with indication of the nature of the selecting criteria (single or multi-objective). With a few exceptions, however, most of them deal with two or more stations, which gives the possibility and flexibility of monitoring both traffic-oriented and background locations.

Table 2.- Different approaches for the selection of air pollution monitoring sites

Reference	**Single (S) or multi-pollutant (M)**	**Spatial distribution of concentrations**	**Single (S) or multi-objective (M)**	**Design/optimization technique**
(Andó et al., 1999)	S	Existing air monitoring station	S	Linear and neural models / selection of candidate points with maximum levels of pollutants
(Corti and Senatore, 2008)	M	Mobile analytical lab/ ISC3 dispersion model	S	Selection of sites with higher pollutant concentrations
(Baldauf et al., 2001)	M	Air dispersion model	M	Risk-based approach
(Kanaroglou et al., 2005)	S	Passive samplers / land-use regression model	M	Location-Allocation model
(Chen et al., 2006)	S	ISC3 dispersion model	M	Modified bounded implicit enumeration algorithm/constraint arrangement method
(Sarigiannis and Saisana, 2008)	S	Atmospheric turbidity as a surrogate for air pollution	M	Maximum gain of the information function with the minimum overlap of stations
(Mazzeo and Venegas, 2008)	S	DAUMOD model for diffuse sources / ISCST3 model for point sources	S	Rank of preselected sites according to number of exceedances

(Elkamel et al., 2008)	M	Air dispersion model (multiple cell approach)	M	Heuristic optimization algorithm
(Ainslie et al., 2009)	S	Existing air monitoring network	S	Entropy-based Bayesian algorithm
(González Ferradás et al., 2010)	S	Diffuse sampling + kriging	S	Selection of sites with concentrations representative of population exposure
(Mofarrah and Husain, 2010)	M	ISC3 dispersion model	M	Fuzzy analytical hierarchy process with triangular fuzzy numbers / Sphere of influence
(Lozano et al., 2010)	S	Passive samplers + spatial interpolation	M	Selection of candidate points that fulfil EC Directive micro and macro scale requirements
(Wu et al., 2011)	S	Existing monitoring network + geostatistical estimation methods to interpolate	M	Stochastic optimization method
(Pope and Wu, 2014)	S	Existing monitoring network	M	Multi-indicator approach coupled to GIS-based model
(Al-adwani et al., 2015)	M	Air dispersion model (multiple cell approach)	M	Sphere of influence / neural network
(Araki et al., 2015)	S	Chemical transport model (WRF/CMAQ) + kriging	S	Genetic algorithm + simulated annealing
(Zoroufchi Benis et al., 2016)	M	Air dispersion model (multiple cell approach)	M	Sphere of influence / Ant colony optimization

				algorithm and genetic algorithm
(Orlowski et al., 2017)	M	Existing air monitoring network + passive samplers	M	PROMETHEE selection tool
(Pahlavani et al., 2017)	M	Existing monitoring network	S	Shannon information index / correlation maps / fuzzy overlay process
(Hao and Xie, 2018)	M	Air dispersion model (WRF/CALPUFF)	M	Genetic algorithm
(Soares et al., 2018)	S	Existing air monitoring network + diffusive sampling + modelling	S	Chemical transport model + cluster analysis
(Alsahli and Al-Harbi, 2018)	M	No use of concentrations but surrogate parameters	M	GIS suitability analysis
(Kazemi-Beydokhti et al., 2019)	M	Existing air monitoring network	M	Analytical Network Process (ANP) method + fuzzy quantifier-guided ordered weighted averaging (OWA)

From the approaches in Table 2, methodologies that only focus on a single pollutant cannot be applied to solve the present problem (locating a single station in a small or medium-size urban area), as it should be acknowledged the need of taking into account all pollutants measured in the Air Quality Monitoring Network (AQMN). From the multi-pollutant approaches, there are works that aim at establishing a rank of the best locations based on different criteria. These could be applied to the present case. However, in some of the approaches the data needed to apply the methodology can be difficult to be accessed and, if available, its uncertainty is high. For instance, in Baldauf et al. (2001), an inhalation rate per time unit representative of each class of population subgroups (children, elderly, healthy adults and adults with previous diseases) is needed, as well as the average exposure time to the pollutant, the average body weight and the number of individuals belonging to each population group for all the grids that the city is divided into. These data are not available straightaway, making this methodology very complex to apply. Another methodology that could be applied is described in Orlowski et al. (2017). They start from the data of an existing AQMN and they rank the existing monitoring sites as a function of the correlation among stations data. The less correlated stations are on top of the rank, whereas the more correlated ones are at the bottom. To do this, they use monthly average concentrations of the target pollutants for 3 years, which represents a huge amount of data, which can only be available if there is an existing network or by modeling. Other works (Al-adwani et al., 2015; Elkamel et al., 2008; Mofarrah and Husain, 2010; Zoroufchi Benis et al., 2016) propose a similar methodology but they obtain the time series needed to calculate the correlation coefficients from modeling. Air pollution models can provide very useful information on air quality from the knowledge of the emissions and the atmospheric processes that control pollutant dispersion, transport, deposition and conversion, together with the meteorology and topography of the studied area. However, they must be run by experts due to their complexity, their results can have large uncertainties and they require very demanding computational requirements. Thus, these techniques cannot always be applied, as big data sets are needed, and these are difficult to obtain without an existing network or the use of modeling.

Another limitation of other methodologies (Kazemi-Beydokhti et al., 2019), as admitted by the authors, is that experts' opinions are involved in the selection of weights using linguistic variables whose values can be rather ambiguous (e.g., "at least one", "few", "some", "half", "many", "most"), so the process could be highly influenced by these opinions. Mofarrah and Hussain (2010) try to overcome the possible uncertainty of the weight selection based on experts' opinion by using triangular fuzzy numbers, but they are still based on the experts' opinion scale.

A general reader may also wonder why not using a general method for locating a facility, such as the location-allocation method (LAM). In fact, these models have not often been used in selecting optimum sites for air quality monitoring stations. An exception is the work by Kanaroglou (2005), that was strongly criticized by Kumar (2009). In his own words "*the goal of an optimal network for air pollution monitoring should be to capture the best representation of air pollution exposure with the available sample size (or the minimum one) rather than optimizing geographic access or attendance at the monitoring stations*" and "*the LAMs are not designed to solve spatial sampling problems*".

An interesting approach by Alsahli and Al-Harbi (2018) consists of using a GIS system to establish the optimum sites for the AQMN. They produce 4 layers based on the following criteria: population density, spatial proximity to main roads, spatial proximity to industries and spatial proximity to high-traffic areas. Each layer is ranked and weighted, and finally they are spatially overlaid to produce a final layer that illustrates the optimum locations. The methodology is validated by comparing the optimal sites retrieved with the concentrations of pollutants measured by existing AQM stations close to the optimum sites. However, one can argue why using surrogate measurements of air pollution concentrations such as population density or distance to main emission sources instead of actual measurements of air pollution. It is necessary to remember that concentrations in air of pollutants are the result of many complex processes and, that not all pollutants are emitted, but some are produced through chemical transformations from those emitted, and their maximum concentrations may occur far from the emission sources, such as ozone.

120 In summary, although there are many approaches in the literature to tackle this problem, most of
121 them are complex to apply and do not pay attention to simple but critical details such as validation
122 of model predictions, adequate selection of weights and/or a global sensitivity analysis.

123 Urban areas in countries of the Organization for Economic Cooperation and development
124 (OECD) are classified as large metropolitan areas if they have a population of 1.5 million or more;
125 metropolitan areas if their population is between 500 000 and 1.5 million; medium-size urban
126 areas if their population is between 200 000 and 500 000; and, small urban areas if their population
127 is between 50 000 and 200 000 (OECD, 2018). In Europe, the percentage of population that lives
128 in small or medium-size urban areas can be up to 100% in some countries such as Luxembourg
129 (Table 3).

130 Table 3.- Percentage of population in small and medium-size urban areas in 2014 in several
131 European countries (OECD, 2018).

Country	Small urban areas (%)	Medium-size urban areas (%)	Total (%)
Sweden	25.3	5.0	30.3
France	14.6	23.0	37.6
Germany	9.6	30.0	39.6
Italy	19.4	21.0	40.4
United Kingdom	13.9	31.0	44.9
Spain	20.9	26.0	46.9
The Netherlands	21.5	29.0	50.5
Czech Republic	46.0	10.0	56
Luxembourg	0.0	100	100

132

133 Although it is acknowledged that larger metropolitan areas are prone to have worse air quality, it
134 is also important to monitor the air quality of small and medium-size urban areas to ensure the
135 necessary protection of their inhabitants along with that of larger metropolitan areas. In most
136 cases, only one monitoring station is required in these areas. In this paper, a multi-pollutant single-
137 objective methodology is proposed to select the best location to set up a single monitoring station

in small or medium-size urban areas. The methodology is proved to be simple and robust, in contrast to other approaches found in the literature.

2. Methodology

The methodology developed in this paper is based on the fact that it is common that different pollutants share the same monitoring sites; that is, there is not a network exclusive for each pollutant, but these are usually measured at the very same sites. This means that the selection of sampling points should not be assessed based on individual pollutants but taking into account all of them. This makes even more sense in the particular case where only a single monitoring station is required in an area or agglomeration: all pollutants will be measured in the only station of the area.

As mentioned before, it is fairly difficult to comply with the Directive macroscale criteria for a single pollutant with a single monitoring station (measuring at the same time both maximum and representative concentrations). This becomes even more complicated if more pollutants come into play. Thus, the approach of this paper to select the monitoring site tries to take into account all the pollutants measured in the network through the use of an objective function in the form of:

$$f(x,y) = W_1 \frac{[Poll_1]_{x,y}}{[Poll_{1,avg}]} + W_2 \frac{[Poll_2]_{x,y}}{[Poll_{2,avg}]} + \cdots + W_n \frac{[Poll_n]_{x,y}}{[Poll_{n,avg}]} \qquad (1)$$

Where W_1, W_2, and W_n are coefficients that range from 0 to 1 and weigh the relative importance of each measured pollutant, $\left[Poll_j\right]_{x,y}$ is the concentration of pollutant $j = \{1,2,\dots,n\}$, being n the total number of pollutants, at the location (x, y), and $\left[Poll_{j,avg}\right]$ is the average concentration of pollutant j in all measured locations.

Given that this methodology has been developed for locating a single station, the optimization criterion is to select the location where $f(x, y)$ is maximized. Regarding the value of coefficients W_1, W_2 and W_n, they can be assigned according to different criteria. For instance, violations of

163 the limit values for one or several pollutants are good arguments for assigning high values of W
164 to those pollutants.
165 Another criterion may be based on the variability of pollutant concentrations over the study area.
166 A high spatial concentration variability may be a good reason to assign a high value of W to a
167 particular pollutant. On the contrary, low spatial variability would mean that it does not really
168 matter where the pollutant is measured as its concentrations do not significantly change over the
169 study area.
170 In Section 3 the proposed methodology is applied to a medium-size urban area (Cartagena, Spain)
171 and both the aforementioned criteria are applied. Two values of $f(x,y)$ are calculated in each
172 sampled location. Subsequently, kriging is performed to obtain estimates of the objective
173 function (equation 1) in unsampled sites and the optimum locations according to each index are
174 selected. Other criteria to assign weights to each pollutant may also be valid and, in order to
175 study the sensitivity of the methodology to changes in the values of W_1 to W_n, a sensitivity
176 analysis is also performed in Section 3.3 for the case study.

177 ## 2.1.Kriging interpolation

178 Kriging is an interpolation method introduced by Daniel Krige in the 1950s and formalised later
179 by Matheron (Matheron, 1963). The term kriging includes different least-squares methods that
180 provide best linear unbiased predictions (BLUP) (Oliver and Webster, 2014). Kriging provides
181 interpolation values without any bias and minimizing variance, which makes the method one of
182 the most used in different applications like environmental studies.

183 Kriging formulation is based on the idea of "random processes", whose principles can be
184 synthetized as follows (Oliver and Webster, 2014):

185 • The value of a property, say z (i.e., a pollutant concentration), at any (consider two
186 dimensional) location $\mathbf{d} = (x, y)$, and denoted by $z(\mathbf{d})$, is a realization of a random variable
187 $Z(\mathbf{d})$.
188 • The set of the infinite random values at all locations, is a random process, and is also
189 denoted $Z(\mathbf{d})$.

- The random variable is spatially correlated at some scale.

Pollutant concentrations in an area can be seen as spatial random variables. When collecting data, usually one single measurement per location under the same conditions is available. Thus, in order to be able to do compute statistics from these realizations, the assumption of stationarity must be considered: there is the same degree of variation among locations. This means that the random process can be represented by the model

$$Z(\mathbf{d}) = \mu + \epsilon(\mathbf{d}) \tag{2}$$

where μ is the mean of the process and $\epsilon(\mathbf{d})$ is a random quantity with a mean of zero and a covariance $C(\mathbf{h})$, with $\mathbf{h}$ representing the distance between observations. Under some assumptions, the covariance is usually replaced by the semivariance, $\gamma(\mathbf{h})$, which is the basis to analyse spatial dependence using the semivariogram analysis. There are different models for the variograms, being the most popular ones the Spherical, Exponential, Gaussian or Matern among others (see Appendix of Oliver and Webster, 2014). The selection of the appropriate model as well as the values of the associated parameters is usually carried out by cross-validation techniques based on the experimental observations.

As stated above, one of the strongest points of kriging is that it provides an estimation of both the mean and the variance of the predictions in the selected locations (commonly called grid). This makes the method especially suitable to do inference about the predictions and even to design optimization algorithms which make use of its statistical power (Egea et al., 2009; Huang et al., 2006; Jones et al., 1998; Ur Rehman et al., 2014; Villemonteix et al., 2008). There are no explicit expressions of the kriging estimations for the mean and the variance. The mathematical procedure to obtain them is out of the scope of this work and the reader is referred to specific textbooks like e.g. Olea (Olea, 1999).

Kriging has been employed in the present study because it has been widely used in air pollution research. In particular, ordinary kriging was selected due to its robustness and the fact that no trend was observed in the data. Apart from the above-mentioned strong points of the method,

Jerrett et al., (2005) highlight that kriging can better deal with much of the erroneous local variability produced with other interpolations given the intrinsic structure of the kriging model. A recent review by Xie et al., (2017) provides a list of references where kriging has been used in air pollution studies to assess single and multi-pollutant concentrations. This review also describes other spatial interpolation approaches that could be used in the proposed methodology, which is not restricted to the use of any interpolant.

In this work, *automap* R package (Hiemstra et al., 2009) was used to obtain the most appropriate semivariance models using cross-validation and its optimal parameter estimates. With this information, kriging maps were made with Surfer 16 software (Golden Software Inc., Golden, Colorado (US)).

2.2.Experimental set-up

The procedure described above was applied to the data obtained in a one-week measurement campaign conducted in Cartagena in July 2009. Although the data set was not recently obtained, it is still fit for testing the proposed methodology.

Cartagena is a medium-sized coastal city (211,996 inhabitants) from the south-east of Spain (INE, 2018), which has a Mediterranean climate with warm summers, mild winters and a dry season in summer (rainfall rarely exceeds 300 mm per year). As regards the wind pattern, it is quite homogenous throughout the year prevailing those winds from the northeast (offshore breezes) or from the southeast (onshore breezes) (AEMET, 2011).

Emissions are mainly due to road traffic, with a fleet of 679 vehicles per 1,000 inhabitants, with diesel being the most used fuel. Other significant pollutant sources are the port, a military arsenal and a nearby industrial estate (DGT, 2018).

Ambient air quality information was collected using Radiello® diffusive samplers (Cocheo et al., 1996). The concentration of benzene, NO_2 and ozone were measured in 12 sampling points spread throughout the city. In order to estimate personal exposure to those species and to evaluate their

magnitude and spatial distribution, the samplers were sited 1.8 m above the pavement and 3 m away from the centerline of the closest road complying with the requirements of European regulations (European Commission, 2008).

Passive sampling meets the requirements of the EC Directive for benzene and other volatile organic compounds in ambient air (CEN, 2005) and has been validated in laboratory experiments to control the effects of temperature, air humidity and wind speed (Ballesta et al., 2005).

Once in the laboratory, the adsorbent cartridges were desorbed and analyzed. The benzene samples were desorbed from the activated carbon using carbon disulfide and the extracts analyzed by GC-FID. The NO_2 molecules were adsorbed in a microporous polyethylene cartridge impregnated with triethanolamine (TEA). Under these conditions, NO_2 is quantitatively transformed to nitrite. The presence of TEA generates an alkaline pH that prevents the oxidation of nitrites to nitrates. Finally, the nitrite ions are quantified by ion chromatography. The adsorbing cartridge for sampling ozone is covered by 4,4'-dipyridylethylene that reacts with ozone (ozonolysis) to give 4-pyridylaldehyde. This aldehyde is condensed with 3-methyl-2-benzothiazolinone hydrazone to yield the corresponding azide, yellow colored, which is determined by visible spectrophotometry at 430 nm (Merck, 2019).

After determining the mass of benzene, NO_2, and ozone in their respective cartridges, the concentrations (C), in $\mu g \cdot m^{-3}$, were obtained through equation (3) using the corresponding sampling rates (SR) and the sampling time (t = 7 days = 10 080 min) (Cocheo et al., 1996). This methodology has been extensively tested and validated according to the protocols from the European Committee for Standardization (CEN).

$$C(\mu g \cdot m^{-3}) = \frac{mass\ (\mu g)}{SR\ (mL \cdot min^{-1}) \cdot t(\text{min})} \cdot 10^6 \quad (3)$$

3. Results and discussion

3.1. Concentrations of benzene, NO_2, and O_3 in Cartagena

Table 4 gathers the measured concentrations of NO_2, O_3, and benzene during the field campaign. From these, isoconcentration contour maps were obtained for the study area using kriging as a geostatistical technique to interpolate pollutant concentrations. Kriging interpolations were overlaid on the map of the city in Figure 1 to actually observe the locations of maximum concentrations of each pollutant.

Table 4. Concentrations of benzene, NO_2, and O_3 measured in each sampling point and their coordinates.

Sampling point	X	Y	Benzene ($\mu g \cdot m^{-3}$)	NO_2 ($\mu g \cdot m^{-3}$)	O_3 ($\mu g \cdot m^{-3}$)
CT-1	-0.9743391	37.6073911	1.6	106	25.4
CT-2	-0.9905928	37.6070843	0.73	53.6	10.1
CT-3	-0.9935480	37.6112485	2.1	151	9.31
CT-4	-0.9938839	37.6167716	0.72	86.6	25.1
CT-5	-1.0024809	37.6163332	0.57	49.5	24.2
CT-6	-0.9864958	37.6101527	0.96	57.2	30.1
CT-7	-0.9993913	37.6269410	0.83	55.7	55.5
CT-8	-0.9812570	37.6062953	1.1	135	4.01
CT-9	-0.9745406	37.6121690	0.10	196	0.386
CT-10	-0.9796451	37.6073911	0.35	51.0	58.8
CT-11	-0.9777645	37.6041913	0.96	43.8	14.1
CT-12	-0.9820630	37.6023941	0.28	97.9	17.0

It should be noted that two of the three locations with maximum concentrations of benzene and NO_2 take place at the same sites (CT-3 and CT-8). Both sites are located next to two busy roads, so it is likely that these higher concentrations may be caused mainly by traffic emissions. CT-1 registers the second highest concentration of benzene, which may be due to the fact that this point is also close to a main road and located less than 75-m apart from a petrol station.

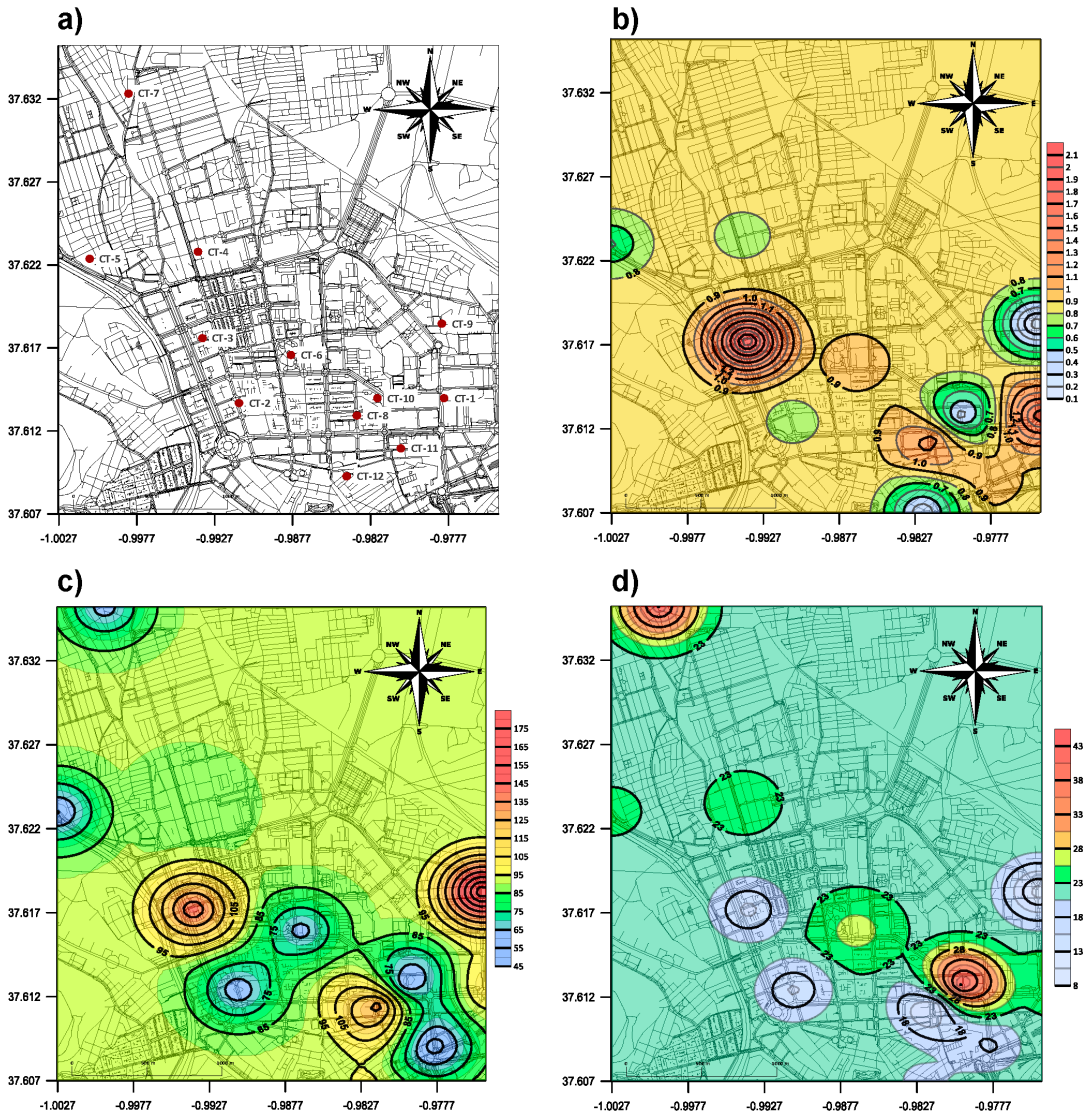


Figure 1. Location of diffusive samplers (a) and concentrations of pollutants in Cartagena obtained by kriging interpolation for b) benzene, c) NO_2 and d) O_3. Concentrations in $\mu g \cdot m^{-3}$ at 293 K and 101.3 kPa.

CT-9 shows the highest concentration of NO_2 and the lowest of benzene, which is indicative of an emission source of NO_2 different from traffic in this location. Regarding ozone, its minimum levels occur at locations with the highest concentrations of benzene and/or NO_2 (CT-9, CT-8 and CT-3). This behavior is not unexpected, as ozone is a secondary pollutant that is produced from photochemical reactions in which NO_2 and volatile organic compounds, such as benzene, are involved. Given the opposite dynamics of these pollutants in the air of Cartagena, it is not possible

to find a location where all three pollutant concentrations are maximum. Thus, objective function described in equation (1) was used to find the locations that maximize it (Section 3.2).

3.2. Determination of the optimum locations to set up a single monitoring station

The location of the monitoring station should not be based solely on the concentration of a single pollutant since it might not be representative of the overall air quality scenario. Equation (1) was used to calculate two indexes representing an average "pollutant" concentration in the locations where the three pollutants were measured.

The elaboration of the pollution indexes was formulated as a weighted sum of the most relevant pollutants in environmental pollution studies, i.e., benzene, NO_2, O_3, as illustrated by equation (1). Other pollutants may also be taken into account. This weighted sum was performed on normalized concentrations, that is, the concentration of each pollutant in each location was divided by the average concentration of that pollutant in all locations sampled.

Two different sets of weights were used to obtain the pollution indexes. The first one was obtained examining the concentrations measured of each pollutant in comparison to their respective limit values, what was called "violation index". The second one was produced considering that the weight of each pollutant is proportional to its variance over the considered space domain, i.e. it assures greater weights to the pollutants whose concentrations show more variability, so these may be considered more difficult to control and regarded as the most critical. The second index was called "proportional-to-variability index".

The study of the concentrations of benzene, NO_2 and O_3 in Cartagena during the field campaign revealed that the most critical pollutant was NO_2, with weekly average concentrations at some points that exceeded the hourly limit value (200 $\mu g \cdot m^{-3}$). The maximum concentrations of O_3 and benzene were 58.9 and 2.13 $\mu g \cdot m^{-3}$, both below half the objective value for ozone (120 $\mu g \cdot m^{-3}$) and half the hourly limit value of benzene (5 $\mu g \cdot m^{-3}$), respectively. For this reason, according to

the "violation index", the weights W_{NO2}, W_{O3} and W_{benz} were assigned as 0.5, 0.25 and 0.25, respectively.

In order to calculate the weights W_{NO2}, W_{O3} and W_{benz} according to the "proportional-to-variability" criterion, the standard deviation of each pollutant concentration over the studied area was calculated. Each weight was calculated as follows:

$$W_i = \frac{SD_i}{SD_{total}} \quad (4)$$

Where SD_i is the standard deviation of the normalized concentrations of pollutant i over the studied area and SD_{total} is the sum of the standard deviations of the normalized concentrations of all pollutants taken into account in the study. The set of weights derived from this calculation was 0.328, 0.269 and 0.403 for benzene, NO_2 and O_3, respectively.

Finally, a kriging interpolation of the two indexes for the area of Cartagena city was performed in order to obtain the optimum location of the monitoring station according to both situations. As it can be seen in Figure 1, there is one clear optimum location to set up the monitoring station according to the "violation index". This location (-0.9934682, 37.611330) is close to sampling point CT-3 (less than 15 meters apart) and the value of its index is 1.181403. Regarding the "proportional-to-variability index", the second and third locations with the highest values of the index were two points located very close to CT-3 (less than 15 meters apart), which makes them ideal places to set up the monitoring station. On the other hand, a location close to CT-7 had the maximum value of the "proportional-to-variability index" and would therefore constitute an optimum location for setting the monitoring station according to this criterion. In order to choose among the two best sampling locations, a good criterion would be selecting the area representative of more person-hours of exposure. In this regard, the selected site would be CT-3, which is a typical urban area, whereas CT-7 can be classified as suburban.

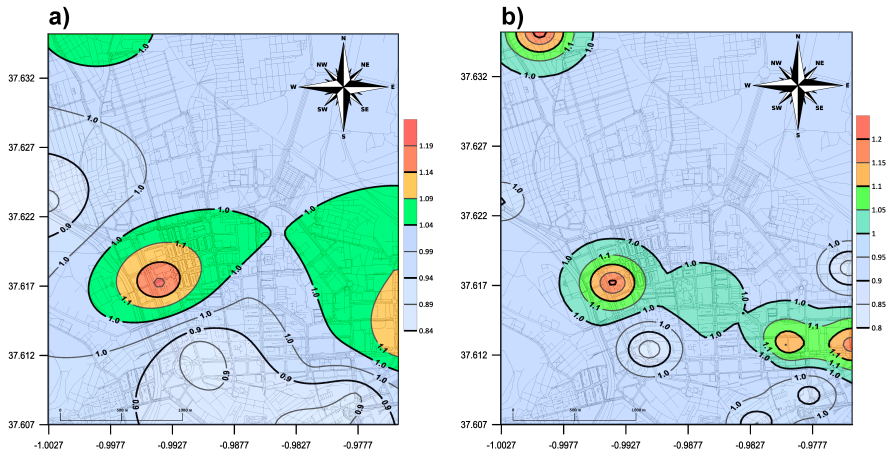


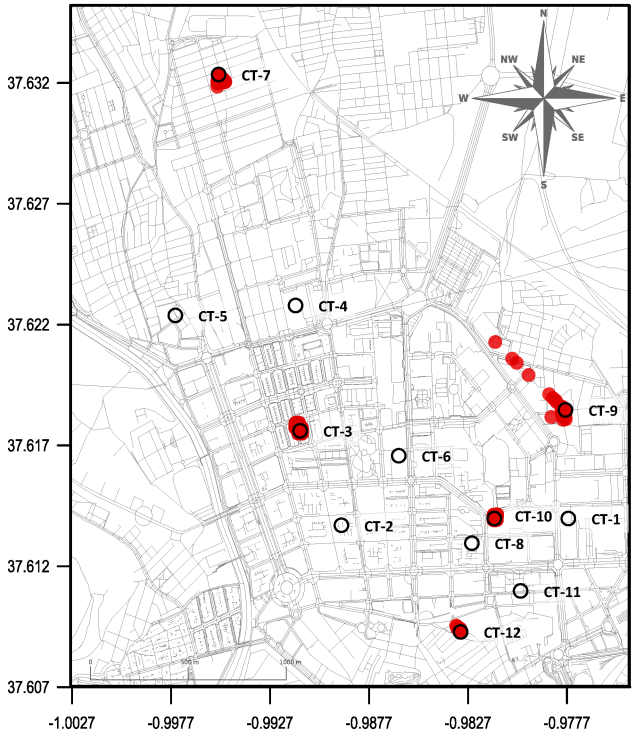
Figure 2. Kriging interpolation of the "violation index" (a) and the "proportional-to-variability index" (b).

3.3. Sensitivity analysis of weight selection

Several weight combinations different from those proposed above could be chosen for the calculation of the proposed index values. These different weight combinations would give rise to different index values whose maximums could return different optimal areas to locate the air quality monitoring station.

In order to test the effect of the choice of weights on the index values, and thus, the robustness of the methodology, many kriging interpolations were carried out with all possible combinations of three weights with a resolution of two decimals (5151 combinations).

The results manifest five clear confined candidate areas (Figure 3). Out of 5151 total points, 2680 points (~52%) corresponded to an area next to CT-3, which was the area predicted by the "violation index"; while 1459 (~28%) to an area next to CT-7, which was the area predicted by the "proportional-to-variability index". From the remaining points, 502 (9.7%) belong to a region next to CT-9, 497 (9.6%) to a zone next to CT-10 and 13 (0.25%) points are nearby CT-12. This means that the areas next to CT-3 and CT-7 account for 80% of the optimal locations and thus, are ideal spots to set up the monitoring station.



356

357 Figure 3. Maximum values of 5151 kriging interpolations covering all combinations of weights.
358 Five plausible areas to locate a monitoring station appear.

359 Average weights ($W_{NO_2,av}$, $W_{O_3,av}$ and $W_{benz,av}$) in the five candidate areas are displayed in
360 Table 5. When an area has a high weight in one pollutant is indicative of a high concentration of
361 that pollutant in that area. For instance, when high importance is given to NO_2, its coefficient
362 W_{NO2} will be high compared to the coefficients of the other two pollutants. This means that
363 locations with a high concentration of NO_2 will retrieve high values of the index, as it is the case
364 of the area close to CT-9.

365 In CT-3 the proportion of benzene/NO_2/O_3 weights is approximately distributed as 50/30/20,
366 whereas in area CT-7 this distribution is approximately 25/25/50. Thus, it can be concluded that
367 in the two best candidate areas the relative weights of each pollutant are more or less balanced,
368 with benzene being the pollutant that dominates in CT-3 area, and ozone the one that dominates
369 in CT-7 area. Regarding the 3 remaining minority areas, locations close to CT-9 are NO_2

pollution-prone, locations close to CT-10 are ozone-prone, and locations next to CT-12 have a distribution of weights of roughly 50/20/30.

Table 5. Average weights of benzene, NO_2 and O_3 in the five areas obtained in the sensitivity analysis (standard deviation in parentheses).

Area	Benzene	NO_2	O_3	Comment
CT-3	0.49 (0.21)	0.34 (0.21)	0.18 (0.11)	≈50/30/20
CT-7	0.22 (0.13)	0.25 (0.16)	0.52 (0.1)	≈ 25/25/50
CT-9	0.06 (0.04)	0.77 (0.11)	0.17 (0.11)	NO_2-prone
CT-10	0.09 (0.06)	0.12 (0.09)	0.79 (0.08)	Ozone-prone
CT-12	0.47 (0.11)	0.17 (0.13)	0.36 (0.02)	≈50/20/30

The methodology presented here provides one clear optimal site for locating the monitoring station; however, it is based on one-week average data. It would be interesting to compare the results obtained in other campaigns carried out at different seasons of the year. This would provide further evidence for the selected optimal site (if the distribution pattern of pollutant concentrations is steady throughout the year) or would retrieve new optimal sites. In this context, González Ferradás et al. (2010) found out that the distribution pattern of pollutants over a city remained more or less steady along the year if the emission sources also did, although the values of concentrations varied due to different dispersion conditions. If this applies to Cartagena, other measurement campaigns would validate the results obtained here.

It is also important to mention that in this work particulate matter (PM) was not measured nor taken into account in the objective function. PM is one of the most alarming pollutants nowadays so, if possible, it is recommended to include it as an extra term in the objective function when this methodology is applied in future.

4. Conclusions

In this paper, a multi-pollutant methodology is described to support decision-making processes regarding the location of a single air quality monitoring station in small and medium-size urban

areas. The methodology is based on the fact that all measured pollutants have to be taken into account in an objective function, (equation 1), which is maximized. Each measured pollutant is assigned a relative weight depending on different criteria. Two criteria have been proposed in this work, namely, a "violation" criterion, where a pollutant is given a high weight when its concentrations exceed periodically its limit values; and a "proportional-to-variability" criterion, where the highest weight is given to the pollutant whose concentrations show a greater variability over the studied area.

In order to ensure that the methodology does not provide as many different locations as combinations of weights, a sensitivity analysis was performed trying 5151 different combinations. The methodology has turned out to be very robust as only 5 candidate locations are eventually retrieved. From these, two of them account for the 80% of the potential locations, being the ideal places to locate the single monitoring site. These two locations coincide with the best locations obtained when applying both of the proposed criteria. The methodology is easy to apply and can be used with as many pollutants as desired, which facilitates and supports the management of the air quality in small and medium-size urban areas only requiring one monitoring station. The recent development of low-cost sensors may help further validate the proposed methodology through sampling campaigns in different seasons of the year.

Acknowledgements

The sampling campaigns used to test the methodology were funded by the Consejería de Agricultura y Agua of the Comunidad Autónoma de la Región de Murcia.

References

AEMET, 2011. Iberian Climate Atlas [WWW Document]. URL http://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/Atlas-climatologico/Atlas.pdf (accessed 1.3.19).

Ainslie, B., Reuten, C., Steyn, D.G., Le, N.D., Zidek, J. V, 2009. Application of an entropy-based Bayesian optimization technique to the redesign of an existing monitoring network for single air pollutants. J. Environ. Manage. 90, 2715–2729. https://doi.org/10.1016/j.jenvman.2009.02.016

Al-adwani, S., Elkamel, A., Duever, T.A., Yetilmezsoy, K., Abdul-wahab, S.A., 2015. A Surrogate-Based Optimization Methodology for the Optimal Design of an Air Quality Monitoring Network. Can. J. Chem. Eng. 93, 1176–1187. https://doi.org/10.1002/cjce.22205

Alsahli, M.M., Al-Harbi, M., 2018. Allocating optimum sites for air quality monitoring stations using GIS suitability analysis. Urban Clim. 24, 875–886. https://doi.org/10.1016/j.uclim.2017.11.001

Andó, B., Cammarata, G., Fichera, A., Graziani, S., Pitrone, N., 1999. A Procedure for the Optimization of Air Quality Monitoring Networks. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 29, 157–163.

Araki, S., Iwahashi, K., Shimadera, H., Yamamoto, K., Kondo, A., 2015. Optimization of air monitoring networks using chemical transport model and search algorithm. Atmos. Environ. 122, 22–30. https://doi.org/10.1016/j.atmosenv.2015.09.030

Baldauf, R., Lane, D., Marote, G., 2001. Ambient Air Quality Monitoring Network Design for Assessing Human Health Impacts from Exposures to Airborne Contaminants. Environ. Monit. Assess. 66, 63–76. https://doi.org/10.1023/A:1026428214799

Ballesta, P.P., Field, R.A., Saeger, E. De, 2005. Population Exposure to Air Pollutants in Europe (People), in: Communicating European Research. Springer Netherlands, Dordrecht, pp. 211–217. https://doi.org/10.1007/1-4020-5358-4_37

CEN, 2005. EN 14662-5 Ambient air quality. Standard method for measurement of benzene concentrations. Diffusive sampling followed by solvent desorption and gas chromatography.

Chen, Ching H., Liu, W.L., Chen, Chia Hsing, 2006. Development of a multiple objective planning theory and system for sustainable air quality monitoring networks. Sci. Total Environ. 354, 1–19. https://doi.org/10.1016/j.scitotenv.2005.08.018

Cocheo, V., Boaretto, C., Sacco, P., 1996. High uptake rate radial diffusive sampler suitable for both solvent and thermal desorption. Am. Ind. Hyg. Assoc. J. 57, 897–904.

https://doi.org/10.1080/15428119691014404

Corti, A., Senatore, A., 2008. Project of an Air Quality Monitoring Network for Industrial Site. Environ. Model. Assess. 65, 109–117.

DGT, 2018. Informes personalizados - Parque de vehículos [WWW Document]. URL https://sedeapl.dgt.gob.es/WEB_IEST_CONSULTA/informePersonalizadoCaptcha.faces (accessed 1.3.19).

Egea, J.A., Vazquez, E., Banga, J.R., Martí, R., 2009. Improved scatter search for the global optimization of computationally expensive dynamic models. J. Glob. Optim. 43, 175–190. https://doi.org/10.1007/s10898-007-9172-y

Elkamel, A., Fatehifar, E., Taheri, M., Al-rashidi, M.S., Lohi, A., 2008. A heuristic optimization approach for Air Quality Monitoring Network design with the simultaneous consideration of multiple pollutants. J. Environ. Manage. 88, 507–516. https://doi.org/10.1016/j.jenvman.2007.03.029

European Commission, 2008. Directive 2008/50/EC on ambient air quality and cleaner air for Europe. Off. J. Eur. Communities 152, 1–43.

González Ferradás, E., Miñarro, M.D., María, I., Terrés, M., José, F., Martínez, M., 2010. An approach for determining air pollution monitoring sites. Atmos. Environ. 44, 2640–2645. https://doi.org/10.1016/j.atmosenv.2010.03.044

Hao, Y., Xie, S., 2018. Optimal redistribution of an urban air quality monitoring network using atmospheric dispersion model and genetic algorithm. Atmos. Environ. 177, 222–233. https://doi.org/10.1016/j.atmosenv.2018.01.011

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, C.J.W., Heuvelink, G.B.M., 2009. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Comput. Geosci. 35, 1711–1721. https://doi.org/10.1016/J.CAGEO.2008.10.011

Huang, D., Allen, T.T., Notz, W.I., Zeng, N., 2006. Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models. J. Glob. Optim. 34, 441–466. https://doi.org/10.1007/s10898-005-2454-3

INE, 2018. Murcia: Población por municipios y sexo [WWW Document]. URL https://www.ine.es/jaxiT3/Datos.htm?t=2883 (accessed 3.1.19).

Jerrett, M., Arain, A., Kanaroglou, P. et al., 2005. A review and evaluation of intraurban air pollution exposure models. J Expo Sci Environ Epidemiol 15, 185–204. https://doi.org/10.1038/sj.jea.7500388

Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient Global Optimization of Expensive Black-Box Functions. J. Glob. Optim. 13, 455–492. https://doi.org/10.1023/A:1008306431147

Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. Atmos. Environ. 39, 2399–2409. https://doi.org/10.1016/j.atmosenv.2004.06.049

Kazemi-Beydokhti, M., Abbaspour, R.A., Kheradmandi, M., Bozorgi-Amiri, A., 2019. Determination of the physical domain for air quality monitoring stations using the ANP-OWA method in GIS. Environ. Monit. Assess. 191, 299.

Kumar, N., 2009. An optimal spatial sampling design for intra-urban population exposure assessment. Atmos. Environ. 43, 1153–1155. https://doi.org/10.1016/j.atmosenv.2008.10.055

Lozano, A., Usero, J., Vanderlinden, E., Raez, J., Contreras, J., Navarrete, B., El, H., 2010. Optimization of the design of air quality monitoring networks and its application to NO 2 and O 3 in Jaen , Spain. Microchem. J. 96, 406–411. https://doi.org/10.1016/j.microc.2010.07.002

Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58, 1246–1266. https://doi.org/10.2113/gsecongeo.58.8.1246

Mazzeo, N.A., Venegas, L.E., 2008. Design of an air-quality surveillance system for Buenos Aires City integrated by a NOxmonitoring network and atmospheric dispersion models. Environ. Model. Assess. 13, 349–356. https://doi.org/10.1007/s10666-007-9101-y

Merck, 2019. Radiello passive air sampler: Overview and applications [WWW Document]. URL https://www.sigmaaldrich.com/technical-documents/articles/analytical/radiello-air-sampler.html (accessed 1.3.19).

Mofarrah, A., Husain, T., 2010. A holistic approach for optimal design of air quality monitoring network

expansion in an urban area. Atmos. Environ. 44, 432–440. https://doi.org/10.1016/j.atmosenv.2009.07.045

OECD, 2018. Urban population by city size (indicator) [WWW Document]. https://doi.org/10.1787/b4332f92-en

Olea, R.A., 1999. Ordinary Kriging, in: Geostatistics for Engineers and Earth Scientists. Springer US, Boston, MA, pp. 39–65. https://doi.org/10.1007/978-1-4615-5001-3_4

Oliver, M.A., Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. CATENA 113, 56–69. https://doi.org/10.1016/J.CATENA.2013.09.006

Orlowski, A., Namiesnik, J., Tobiszewski, M., 2017. Assessment and Optimization of Air Monitoring Network for Smart Cities with Multicriteria Decision An ...., in: Intelligent Information and Database Systems. https://doi.org/10.1007/978-3-319-54430-4

Pahlavani, P., Sheikhian, H., Bigdeli, B., 2017. Assessment of an air pollution monitoring network to generate urban air pollution maps using Shannon information index , fuzzy overlay , and Dempster-Shafer theory , A case study : Tehran , Iran. Atmos. Environ. 167, 254–269. https://doi.org/10.1016/j.atmosenv.2017.08.039

Pires, J.C.M., Sousa, S.I.V., Pereira, M.C., Alvim-Ferraz, M.C.M., Martins, F.G., 2008a. Management of air quality monitoring using principal component and cluster analysis-Part I: SO2and PM10. Atmos. Environ. 42, 1249–1260. https://doi.org/10.1016/j.atmosenv.2007.10.044

Pires, J.C.M., Sousa, S.I.V., Pereira, M.C., Alvim-Ferraz, M.C.M., Martins, F.G., 2008b. Management of air quality monitoring using principal component and cluster analysis-Part II: CO, NO2and O3. Atmos. Environ. 42, 1261–1274. https://doi.org/10.1016/j.atmosenv.2007.10.041

Pope, R., Wu, J., 2014. A multi-objective assessment of an air quality monitoring network using environmental , economic , and social indicators and GIS-based models A multi-objective assessment of an air quality monitoring network using environmental , economic , and social ind. J. Air Waste Manage. Assoc. 64, 721–737. https://doi.org/10.1080/10962247.2014.888378

Sarigiannis, D.A., Saisana, M., 2008. Multi-objective optimization of air quality monitoring. Environ. Monit. Assess. 136, 87–99. https://doi.org/10.1007/s10661-007-9725-z

Soares, J., Makar, P.A., Aklilu, Y., Akingunola, A., 2018. The use of hierarchical clustering for the design of optimized monitoring networks. Atmos. Chem. Phys. 18, 6543–6566. https://doi.org/10.5194/acp-18-6543-2018

Ur Rehman, S., Langelaar, M., van Keulen, F., 2014. Efficient Kriging-based robust optimization of unconstrained problems. J. Comput. Sci. 5, 872–881. https://doi.org/10.1016/j.jocs.2014.04.005

Villemonteix, J., Vazquez, E., Walter, E., 2008. An informational approach to the global optimization of expensive-to-evaluate functions. J. Glob. Optim. 44, 509. https://doi.org/10.1007/s10898-008-9354-2

Wu, L., Bocquet, M., Wu, L., Bocquet, M., 2011. Optimal redistribution of the background ozone monitoring stations over France. Atmos. Environ. 45, 772–783. https://doi.org/10.1016/j.atmosenv.2010.08.038

Xie, X., Semanjski, I., Gautama, S., Tsiligianni, E., Deligiannis, N., Rajan, R.T., Pasveer, F., Philips, W., 2017. A Review of Urban Air Pollution Monitoring and Exposure Assessment Methods. ISPRS Int. J. Geo-Inf. 6, 389. https://doi.org/10.3390/ijgi6120389

Zoroufchi Benis, K., Fatehifar, E., Shafiei, S., Keivani Nahr, F., Purfarhadi, Y., 2016. Design of a sensitive air quality monitoring network using an integrated optimization approach. Stoch. Environ. Res. Risk Assess. 30, 779–793. https://doi.org/10.1007/s00477-015-1189-z