A PRACTICAL APPROACH FOR THE OPTIMIZATION OF AIR MONITORING NETWORKS USING SATELLITE DATA

Lucimara Bragagnolo
Doutoranda do Programa de Pós-Graduação em Engenharia Civil e Ambiental, Universidade de Passo Fundo - UPF
BR 285 Km 292,7 | Campus I, Bairro São José - São José, Passo Fundo - RS, 99052-900
lucimarabragagnolo@hotmail.com

Vinicius Luiz Pacheco
Doutorando do Programa de Pós-Graduação em Engenharia Civil e Ambiental, Universidade de Passo Fundo - UPF
BR 285 Km 292,7 | Campus I, Bairro São José - São José, Passo Fundo - RS, 99052-900
vinimanfroipacheco@gmail.com

Moacir Kripka
Docente do Programa de Pós-Graduação em Engenharia Civil e Ambiental, Universidade de Passo Fundo - UPF
BR 285 Km 292,7 | Campus I, Bairro São José - São José, Passo Fundo - RS, 99052-900
mkripka@upf.br

ABSTRACT

The proper location of air quality monitoring networks is essential for determining concentrations of air pollutants, enabling management, control, and pollution mitigation actions, as well as studies to be developed based on representative data. Usually, station allocation methods involve qualitative or quantitative methods that require a previous step of modeling concentration data, limiting their applicability to regions with a scarcity of this type of information. In this context, this study sought to integrate remote sensing products from the Sentinel 5P satellite that collects data on concentrations of different types of pollutants with a linear optimization method, to allocate a network of air monitoring stations. As a case study, the Metropolitan Region of São Paulo was used. The methodology has the potential to be applied in any region of the globe since the data obtained by the satellite are global in scope with daily temporal resolution.

KEYWORDS. Air quality. Monitoring stations. Linear Programming. Sentinel 5P.

Paper topics: AG&MA – PO na Agricultura e Meio Ambiente; PM – Programação Matemática

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

1. Introduction

Problems related to air pollution are recognized worldwide and are considered critical in most countries with a high degree of urbanization and industrialization. Because of this, monitoring the concentrations of different pollutants is extremely important, especially for populations located in these most critical regions [Wang et al. 2013]. In addition, decision-makers and scientists have shown a growing interest in this type of information, once chronic effects on people's health have been observed when they are continuously exposed to air pollution [Apparicio et al. 2021; Huang et al. 2018; X. Liu et al. 2016]. Thus, air monitoring networks capable of capturing representative concentrations are essential to enable studies that seek to understand the impacts of air pollutants on human health and the environment.

Usually, the choice of the monitoring stations allocation is carried out qualitatively. Decisions about the minimum number of sampling stations to meet the objectives of a study and the location of individual sampling stations to obtain readings considered to be representative are generally based on available air quality information given by other sources, meteorological data, and the dispersion model results [Noll et al. 1977]. This type of approach at the time of decision-making can lead to less precise positioning of the equipment, impairing the mapping of the actual concentrations of pollutants in the region of interest.

To overcome these limitations, some more quantitative approaches, mainly involving optimization algorithms, started to be applied in some studies [Elkamel et al. 2008; Gupta et al. 2018; Sarigiannis and Saisana 2008; Wang et al. 2013]. Optimization algorithms make it possible to identify the best areas for allocation of these stations by taking into account parameters such as pollutant concentrations and population density in their objective functions or restrictions, among others. As an example, Elkamel et al. [2008] used a heuristic optimization algorithm to identify the optimal configuration of an air monitoring network. The study applied a composite objective function taking into account values of concentrations of SO_2, NO_x, and CO and the violation of concentrations on environmental standards. The concentrations for the research were obtained from the application of a mathematical model of dispersion. In the same year, Sarigiannis and Saisana [2008] constructed a multi-objective problem to perform the allocation of monitoring systems using information based on remote sensing of the troposphere and conversion models to determine the concentrations of the pollutants of interest (NO_2 and SO_2). The parameters considered in the objective were standard violation score, land use score, the density of cultural heritage sites, population density, and cost function. Wang et al. [2013] used a genetic algorithm also to allocate monitoring stations. To elaborate the objective function, the authors took into account the concentrations observed for the study region (obtained by modeling), the population in the area, and implementation costs. In a more recent study, Gupta et al. [2018] applied the Spatial Simulated Annealing (SSA) algorithm as an optimization technique to identify the best locations to implement a monitoring network. The authors used a regression model that takes into account land use to model NO_2 concentrations to then generate results to be inserted into the optimization algorithm.

Considering the potential of methodologies that use optimization algorithms to solve the allocation of air pollution monitoring stations' problems, their applicability may be limited to regions that have information on available air pollutant concentrations, as they are necessary as input parameters in these algorithms. As pointed out in the studies aforementioned, concentration data are usually obtained through existing data and/or mathematical modeling. With the development of new remote sensing technologies, sensors and methods began to enable satellite mapping of pollutant concentrations on a global scale, such as the Sentinel-5P satellite [ESA, 2017], launched in 2017 and capable of monitoring ozone, methane, formaldehyde, aerosol, carbon monoxide, NO_2 and SO_2 in the tropospheric layer of the atmosphere, with a daily temporal resolution.

Data from the Sentinel-5P satellite have been widely applied in research related to atmospheric sciences [Jeong and Hong 2021; J. Liu 2021; Mehmood et al. 2021; Park et al. 2021]. Verified the good agreement of this information with data obtained at ground level [Verhoelst et al. 2021], it is possible to identify the potential of using this product in solving problems related to optimizing the allocation of air monitoring networks, thus eliminating the need to produce

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

secondary data through mathematical modeling, or limiting the applicability in certain areas because of the lack of data, for example.

Given this context, this study aims to present an optimization methodology for the allocation of air pollution monitoring stations using information from the Sentinel-5P satellite and the open-source language Python. As far as we know, this is the first study that seeks to aggregate such free tools and products to solve the problem here proposed.

2. Methodology

2.2. Study area

To apply the optimization methodology, the Metropolitan Region of São Paulo (RMSP) was selected as the study area, being the largest metropolitan region in Brazil and encompassing about 21.5 million inhabitants [IBGE, 2018]. The region covers 39 municipalities in the state of São Paulo, with a high degree of urbanization. The location of the RMSP is illustrated in Fig. **1**. This area was selected for methodological application because it is a highly urbanized and industrialized region, where the highest levels of concentrations of air pollutants are usually found. In addition, the RMSP has a well-established air quality monitoring network [CETESB 2013], which can be used to verify the validity of the location results obtained through the methodology proposed in this study.



Fig. 1 Location of the Metropolitan Region of São Paulo. Source: Carvalho e Freitas [2006]

2.3. Data collection

Two main sources of information were used to collect data on pollutant concentrations and information about population density in the study area. Such sources are detailed below.

2.3.1. Pollutant concentrations

Spatial information on the concentration of different pollutants for the study area was collected from the Sentinel-5P satellite image database [ESA, 2017]. The identification at a global level of different compounds is given by means of an ultraviolet, visible, near, and short infrared spectrometer called Tropomi. The sensor takes measurements every second, covering an area approximately 2,600 km wide and 7 km long at a resolution of 7 x 7 km. The captured light is separated into different wavelengths using grid spectrometers and then measured with four different detectors for the respective spectral bands. The UV spectrometer has a spectral range of 270-320 nm, the visible light spectrometer has a range of 310-500 nm, the NIR has a range of 675-775 nm and the SWIR has a range of 2305-2385 nm. Tropomi is capable of monitoring ozone, methane, formaldehyde, aerosol, carbon monoxide, NO_2, and SO_2 in the atmosphere.

For this work, information was collected for 4 pollutants: carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and sulfur dioxide (SO_2). These compounds were selected because they

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

usually represent the largest proportions of air pollution in urban centers, and also because they are usually the most monitored compounds in air quality monitoring stations. To take into account fluctuations throughout the week and one month, daily images were obtained for May/2021. Concentrations are given in $\mathrm{mol/m^2}$.

2.3.1.1. Image processing

The images were obtained in NetCDF (.nc) format and were processed to allow them to be imported into the geographic information software (GIS) QGIS. To do so, we used the 'HARP Python Python library [HARP Developers 2015], which has a set of tools for ingesting, processing, and comparing satellite data. For each of the images obtained, the command 'harpconvert' was used, which performs the conversion of raw files from Sentinel 5P to georeferenced files with 2.5 km x 2.5 km resolution that are compatible with the GIS used to develop this research.

After the conversion process, the images were imported into QGIS, cut to the study area, and, from them, matrix maps referring to the mean values observed for each pollutant were generated. The average value of each of the pollutants was obtained using the raster calculator of the GIS. As the objective is to identify the areas with the highest concentrations of a given pollutant, it was decided to use the average value of observations over time, and not the maximum value, since maximum values may not represent the places that have higher concentrations most of the time.

The matrix maps with the mean values were then exported in .TIF format and imported into the Python environment with the help of the rasterio library [Gillies and others 2014], which makes it possible to manipulate georeferenced information.

2.3.1.2. Aggregation of pollutant maps

Due to differences in emission sources, diffusion and transformation rates, and other factors, each pollutant usually has differences in its spatial distribution, the number and location of monitoring stations would therefore be specific for each pollutant [Modak and Lohani 1985]. However, a monitoring stations network requires the ability to capture the spatial and temporal variability of different pollutants. If an optimization project is driven one pollutant by one pollutant, this would lead to a large computational cost and an increase in the complexity of the optimization process. Because of this, this study took into account a "synthetic evaluation concentration", as proposed by Zheng et al. [2011]. This approach groups concentrations of different pollutants, simplifying the optimization process and reflecting the synthetic characteristics of pollutants on a regional scale. Therefore, we used Jordan's formula in probability theory to calculate the synthetic evaluation concentrations ignoring the mutual influences between pollutants. [Zheng et al. 2011]. The equation is described in the form:

$$\mathrm{C_i(s)} = 1 - \left[\left(1 - \mathrm{C_i(x_1)}\right)\left(1 - \mathrm{C_i(x_2)}\right)\left(1 - \mathrm{C_i(x_3)}\right) \ldots \left(1 - \mathrm{C_i(x_n)}\right)\right] \quad (1)$$

where i is the grid number; $\mathrm{C_i(s)}$ is the concentration of synthetic assessment of multi-pollutants in the grid i; $\mathrm{C_i(x_n)}$ is the average normalized pollutant concentration $\mathrm{x_n}$ in grid i. The normalization of the average concentrations of each pollutant is given by:

$$\mathrm{C_i(x)} = \frac{\mathrm{x_i} - \min(\mathrm{x})}{\max(\mathrm{x}) - \min(\mathrm{x})} \quad (2)$$

The final map with the synthetic evaluation concentrations was then submitted to the optimization algorithm that is detailed in Section 2.3.

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

2.3.2. Population density information

The population density information for the metropolitan region of São Paulo was obtained in vector format (.shp) through DataGEO – São Paulo Environmental System [Estado de São Paulo 2019], for the year 2019. The vector file was converted to raster (matrix data) considering a spatial resolution of 2.5 km x 2.5 km. Population density is related to each municipality that is present in the study area, that is, for each municipality, a single mean density value is associated. However, if there is more detailed information (different density values for a single municipality), the method allows these to be inserted, giving even more accuracy to the results. Population density values were also normalized, as were concentration values, according to Equation 2.

2.4. Optimization model

2.4.1. Objective function

The objective function takes into account two factors: synthetic pollutant concentrations, calculated by the Jordan formula (Equation 2), and population density. The function that seeks to maximize concentrations (regions of greater pollution) is given by:

$$\text{Max } C = \sum_{j=1}^{n} C_j \times X_j \quad (3)$$

where n are the possible candidates for the allocation of a station, j is the possible location of the pixel, C_j is the concentration calculated by the Jordan formula in the position j, and X_j is the binary value (0 or 1) to be optimized, and which indicates the allocation (1) or not (0) of a station.

The social factor, represented by population density, was considered since areas with a greater number of people should be prioritized to carry out the monitoring of pollutants. The general objective of a monitoring network is precisely to protect human health from the deterioration of air quality. The objective function can then be represented by:

$$\text{Max } P = \sum_{j=1}^{n} P_j \quad (4)$$

where P_j refers to the value of population density in the pixel j.

The two functions can then be joined into a single composite objective function, in the form:

$$\text{Max } F = \sum_{j=1}^{n} C_j \times P_j \times X_j \quad (5)$$

2.4.2. Restrictions

Ideally, more monitoring stations can better characterize the spatial variability of air quality in a region. However, the number of monitoring stations is always limited by the available budget. Thus, as a restriction, the number of monitoring stations possible to be implemented was adopted:

$$\sum_{j=1}^{n} X_j \leq B \quad (6)$$

where B is the total number of monitoring stations planned to be installed.

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

Another restriction is related to the distance of stations. For the algorithm to avoid concentrating the stations too close together, a minimum distance at which stations must be from each other was adopted. This distance can be chosen based on the equipment's range level. The restriction is described as:

$$d \times (X_i + X_j - 1) \leq \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (7)$$

where d is the minimum distance between stations; X_i and X_j are the binary decision variables in the positions i and j; x_i and x_j are the indices referring to the matrix line that represents the study region; and y_i and y_j they are the indexes referring to the column of the matrix that represents the study region. The second part of the inequality refers to the distance between two points, which must be greater than the established distance d.

Also, the variable to be optimized X_j must be binary, then:

$$X_j = \{0,1\} \quad (8)$$

Finally, pixels represented by different types of water resources were removed (0 values were inserted in the concentration and population density map) using GIS and a layer of land use and occupation of the study area, obtained from DataGEO - São Paulo Environmental System [Estado de São Paulo 2010]. This process was necessary to prevent stations from being allocated in these areas.

2.4.3. Implementation

The construction and resolution of the optimization model were performed in Python language, using the PuLP library [Mitchell et al. 2011]. PuLP is an open-source linear programming (LP) package that extensively uses Python syntax and integrates with several standard solvers such as GLPK, COIN-OR CLP/CBC, CPLEX, GUROBI, MOSEK, XPRESS, CHOCO, MIPCL, and SCIP. The code and data used to solve the proposed problem are available at https://github.com/bragagnololu/OptStations.

2.4.4. Optimization model specifications and variables

To implement the method, the same number of automatic monitoring stations existing in the RMSP (29) was considered as a restriction on the maximum number of stations to be allocated. This was done to compare the location between the existing networks and the network to be optimized. In addition, three scenarios were evaluated considering minimum distances between stations (parameter d of Equation 7): 2, 3, and 4 pixels, which are equivalent to approximately 5, 7.5, and 10 km, respectively. The distance of 2 pixels corresponds to the minimum distance observed between two of the monitoring stations deployed in the RMSP.

Considering the RMSP, an information matrix with 48×63 pixels, constituting 3024 binary integer variables. Pixels outside the boundaries of the RMSP, as well as those represented by water resources, were set as zero. In total, 9,141,555 constraint equations were considered. This expressive number is associated with distance restrictions, where each variable is compared to the other.

After building the linear model with its constraints, the problem was solved using the GLPK Integer Optimizer solver and applying the long step dual simplex method [Kostina 2002]. A maximum execution time of the algorithm of 16,000 seconds was defined. The entire process was performed on a workstation with a 3.6 GHz Intel Xeon processor with 32 GB of RAM.

3. Results

Fig. **2** presents the average spatial distribution of the pollutants considered in the study (CO, NO_2, O_3, SO_2) obtained through the processing of images from the Landsat 5P satellite, and

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

for the defined information period. It was possible to identify higher tropospheric O_3 concentrations, in the order of 0.11 mol/m², although no expressive differences were observed in the spatial distribution of values. The second pollutant with higher magnitude values is CO, which showed higher concentrations in the region where the city of São Paulo is located, as well as some plumes located in some of the adjacent cities, spatially in the northeast of the region. The pollutant NO_2 had similar behavior to CO, although the highest concentrations were even more restricted to the more urbanized region with higher population density (Fig. **3**a). Finally, there were no expressive values and a distribution pattern for SO_2 concentrations. The highest concentrations of CO are mainly related to the emission of the pollutant given by light vehicles. NO_x emissions, on the other hand, are correlated with the circulation of heavy vehicles and industrial processes [CETESB 2020]. Ozone, being a secondary pollutant, is not directly emitted by primary sources, and its formation is associated with concentrations of pollutants such as NO_x and volatile organic compounds, as well as atmospheric conditions.

Fig. **3** presents the population density map, in inhabitants/m² before normalization and the map of synthetic concentrations already normalized used in the optimization method and obtained from the application of Equations 1 and 2. Regarding Fig. **3**a, the highest demographic densities, in the order of 6000 inhabit/km² are found in the city of São Paulo and adjacent cities. In the case of synthetic concentrations, as the ones that most varied both in spatial distribution and magnitude were the concentrations of pollutants CO and NO_2, Fig. **3**b shows similar distribution characteristics of these two compounds.

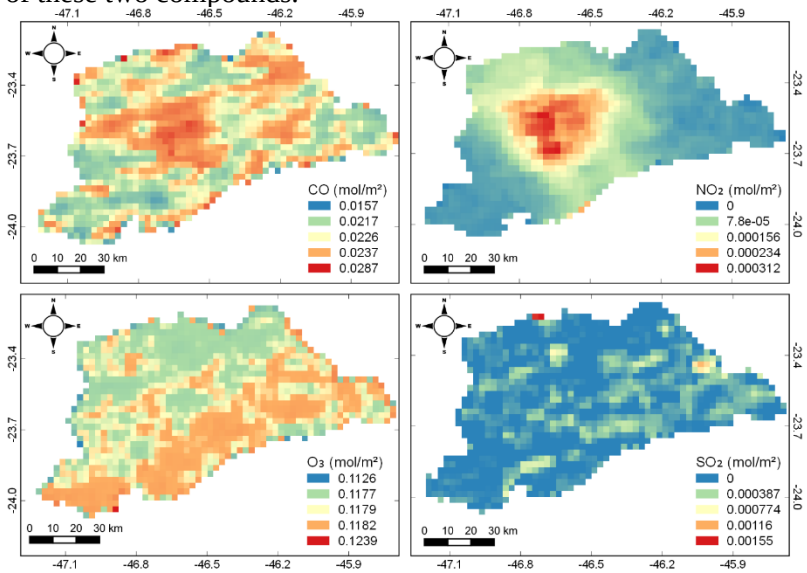


Fig. 2 Average concentrations of carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3) and sulfur dioxide (SO_2) for the metropolitan region of São Paulo, in mol/m²

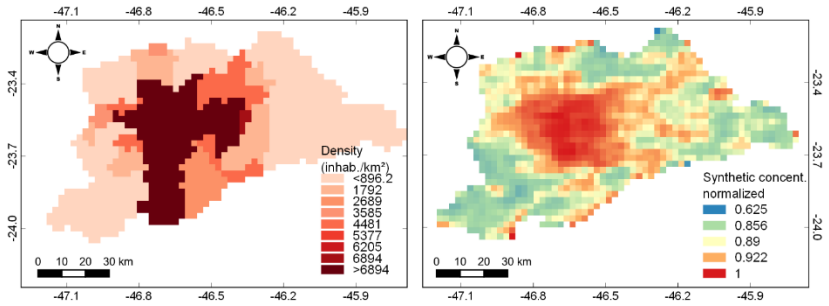


Fig. 3 Information on (a) population density for the São Paulo Metropolitan Region, and (b) normalized synthetic concentrations given by grouping the average concentrations of individual pollutants using the Jordan formula

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

The pixel-by-pixel product given between the normalized population density and the synthetic concentrations, in order to portray the objective function given by Equation 5, is represented by Fig. **4**. The white areas represent the location of water resources, where it was not considered as a potential installation site for a station. Fig. **4** also shows the location of the 29 automatic monitoring network stations operating in the RMSP, for comparison purposes with the results obtained by the optimization process. In general, such stations are located in areas with the highest concentration of pollutants, as well as in areas with higher population density, most of them located in the city of São Paulo.

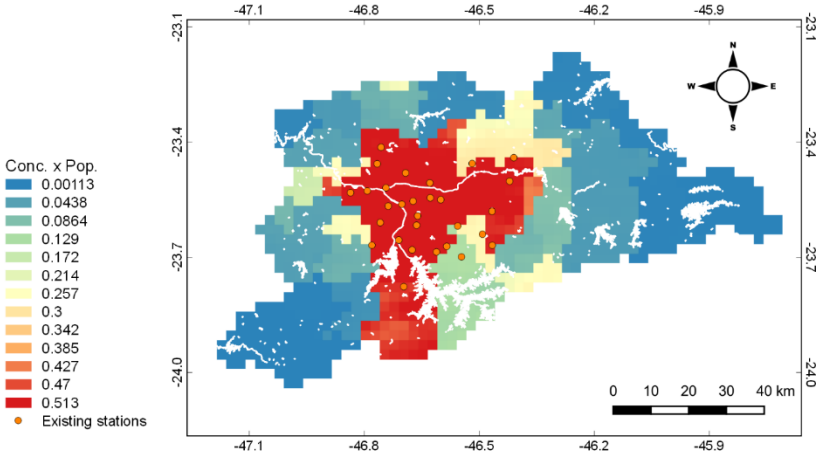


Fig. 4 Map resulting from the product between normalized synthetic concentrations and normalized population density and location of the 29 air quality monitoring stations already implemented in the Metropolitan Region of São Paulo. White areas refer to the location of water bodies that were not considered as possible locations for the implementation of a monitoring station

After carrying out the methodological procedures and application of the optimization method, the location for 29 monitoring networks was obtained, considering different minimum distances between each station: 2, 3, and 4 pixels, which represent about 5, 7.5, and 10 km, respectively. The results can be seen in Fig. **5**. It is possible to observe that, for the minimum distance considered (which was also the minimum distance observed between two stations already implemented in the RMSP), the optimized stations are concentrated close to the existing stations and in areas with greater population density and concentration of pollutants. This is indicative that the methodology that combines the use of freely available satellite images with global coverage, together with the implementation of a linear optimization method can be efficient and useful to carry out the initial planning of networks of air monitoring stations. As a form of quantitative assessment, Fig. **6** presents a boxplot comparing the distribution of pixel values referring to the product of concentration and population density (Fig. **3**b) where the stations implemented by CETESB and the stations obtained by the optimization methodology are located. It is possible to verify that the values where the stations were allocated through optimization are higher than those where the real stations are, showing that, in this case, the monitoring network would be located in areas that would have greater representation. However, to have a higher level of confidence, it is interesting to consider a longer period for assessing concentrations of pollutants, to detect seasonal variations in different months, for example.

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

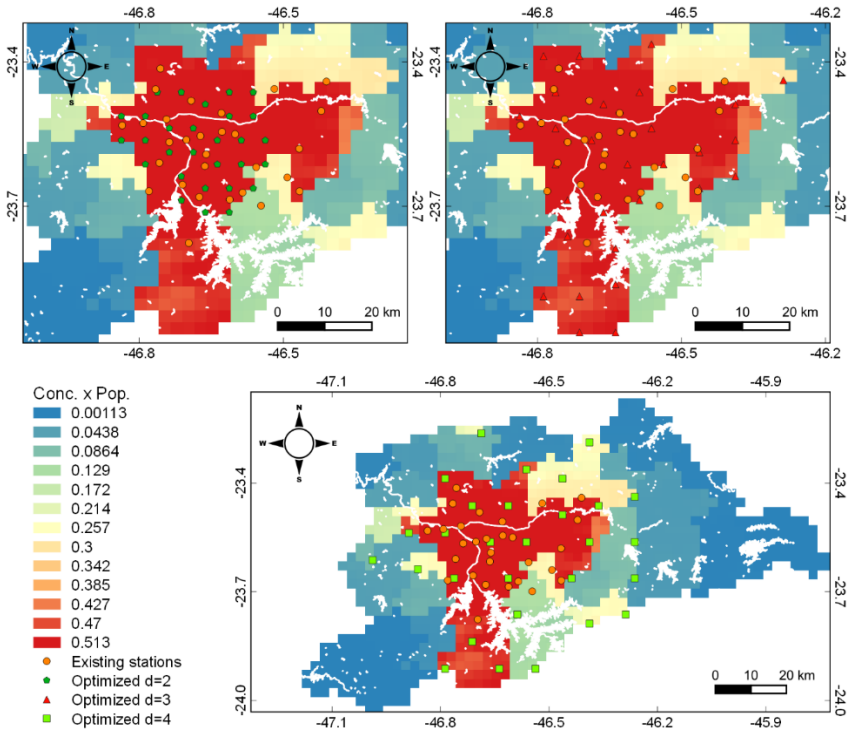


Fig. 5 Location of existing monitoring stations in the São Paulo Metropolitan Region compared to the allocation obtained by the optimization methodology for minimum distances between stations of (a) 2, (b) 3, and (c) 4 pixels. White areas refer to the location of water bodies that were not considered as possible locations for the implementation of a monitoring station

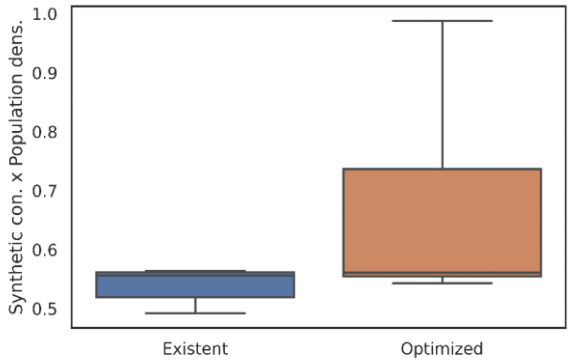


Fig. 6 Boxplot of the distribution of pixel values referring to the product of the synthetic concentration and demographic density in which the CETESB stations are located and the location of stations obtained by the optimization method

Regarding the optimized allocation of stations with minimum distances between them of 3 and 4 pixels, there is a greater spacing, but they are concentrated, for the most part, in the city of São Paulo. Only the scenario with a distance of 4 pixels had a greater number of stations located in the most adjacent areas of the city. The minimum distance between each station must be defined based on criteria established by the project planning agents. The final value of the objective function, as well as the execution time of the algorithm, for each case, is presented in Tabela 1. It can be observed that, the greater the minimum distance defined as a restriction, the greater the execution time of the algorithm until the optimal solution. In the case of the third scenario (d = 4),

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

an optimized solution was obtained, but not optimal, since the execution time reached the defined time limit. However, optimized solutions, even not being optimal, present good answers, being even better than solutions obtained qualitatively.

Tabela 1. Objective function value and execution time for each evaluated scenario

Distance (px)	**2**	**3**	**4**
Objective function	19.23	18.11	18.50
Runtime (s)	140	15580	> 16000

4. Discussion

According to the results presented, it was possible to verify the potential of the methodology to be applied for the prior study of allocations of air monitoring networks, especially in regions that lack air quality information before such studies (they do not have nearby stations or who need to apply models of dispersion/advection phenomena even with little information on pollutant concentrations). Thus, the method proposed by this work seeks to facilitate the performance of such studies, as it uses information available for all regions of the world for free, as well as applying linear optimization methods through open-source language, without the need for hiring specific software, for example. However, it is recommended that more tests be carried out in regions that also have a well-established monitoring network, such as the RMSP, for methodological validations; as well as further validation studies of the data obtained by Sentinel-5P with measured concentrations in soil, as it is still a relatively new product. A later stage with expert evaluation can still be included to improve and verify the validity of the results, as proposed by Zheng et al. [2011]. Changes to the objective function can also be made, to include additional information that is considered relevant by the planning agents. Furthermore, there is the possibility of adding weights of importance for each of the parameters considered by the function.

Some methodological limitations can also be listed, including the need for greater computational power when large matrices are used (large regions and high image resolution). This results in a large number of variables that must be optimized, which can lead to memory problems in the workstation, as well as a long time to obtain an optimized solution. An alternative can be to reduce the image resolution or reduce the study area, if possible. However, other limitations come into being with the adoption of these possibilities. Another restriction is associated with the time of data made available by the Sentinel-5P satellite, as well as the restricted number of mapped pollutants. Because their information is relatively recent, studies that wish to use a longer period to assess the distribution of pollutants are limited to the time available.

As for the optimization method used, the dual simplex proved to be adequate for the problem that had an expressive number of variables and constraints. However, the proposed methodology does not restrict the optimization method to be applied. Other methods can be tested and compared to identify their strengths and weaknesses.

5. Final remarks

This study proposed a methodology for optimizing the allocation of air quality monitoring stations using the information on pollutant concentrations obtained via Sentinel-5P satellite combined with a linear programming method. From the grouping of concentrations observed in a given period and population density information, the method sought to maximize the objective function that contains such parameters, in addition to using as restrictions the number of stations, minimum distance between stations, and terrain conditions (presence of water resources). The application of the methodology in the Metropolitan Region of São Paulo showed satisfactory results, in which the allocation of stations given by the optimization was similar to the location of the existing monitoring network. The scope and availability of the data, combined with a relatively simple method of optimization, allow for wide methodological applicability.

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

References

Apparicio, P., Gelb, J., Jarry, V., e Lesage-Mann, É. (2021). Cycling in one of the most polluted cities in the world: Exposure to noise and air pollution and potential adverse health impacts in Delhi. *International Journal of Health Geographics*, 20(1):1–16. https://doi.org/10.1186/s12942-021-00272-2

Carvalho e Freitas, P. C. de. (2006). Map locator of the Metropolitan Region of São Paulo, SP.

CETESB. (2013). Redes de Monitoramento. Qualidade do Ar. https://cetesb.sp.gov.br/ar/redes-de-monitoramento/. Accessed 6 June 2021

CETESB. (2020). Qualidade do ar no Estado de São Paulo.

Elkamel, A., Fatehifar, E., Taheri, M., Al-Rashidi, M. S., e Lohi, A. (2008). A heuristic optimization approach for Air Quality Monitoring Network design with the simultaneous consideration of multiple pollutants. *Journal of Environmental Management*, 88(3):507–516. https://doi.org/10.1016/j.jenvman.2007.03.029

ESA. THE EUROPEAN SPACE AGENCY. (2017). Sentinel-5P. Sentinel. https://sentinel.esa.int/web/sentinel/missions/sentinel-5p

Estado de São Paulo. (2010). Cobertura da Terra. Data Geo. Sistema Ambiental Paulista. https://datageo.ambiente.sp.gov.br/app/#. Accessed 4 June 2021

Estado de São Paulo. (2019). Densidade demográfica. Data Geo. Sistema Ambiental Paulista. https://datageo.ambiente.sp.gov.br/app/#. Accessed 4 June 2021

Gillies, S., et al. (2014). Rasterio: geospatial raster I/O for Python programmers. https://github.com/mapbox/rasterio

Gupta, S., Pebesma, E., Mateu, J., e Degbelo, A. (2018). Air quality monitoring network design optimisation for robust land use regression models. *Sustainability (Switzerland)*, 10(5):1–27. https://doi.org/10.3390/su10051442

HARP Developers. (2015). HARP manual. HARP Python interface. https://stcorp.github.io/harp/doc/html/index.html. Accessed 4 June 2021

Huang, J., Pan, X., Guo, X., e Li, G. (2018). Impacts of air pollution wave on years of life lost: A crucial way to communicate the health risks of air pollution to the public. *Environment International*, 113:42–49. https://doi.org/10.1016/j.envint.2018.01.022

IBGE. Insituto Brasileiro de Geografia e Estatística. (2018). Estimativas da população residente nos municípios brasileiros com data referência em 1o de julho de 2018. ftp://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2018/estimativa_dou_2018.pdf. Accessed 6 June 2021

Jeong, U., e Hong, H. (2021). Assessment of Tropospheric Concentrations of NO2 from the TROPOMI/Sentinel-5 Precursor for the Estimation of Long-Term Exposure to Surface NO2 over South Korea. *Remote Sensing*, 13(10):1877.

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873

Kostina, E. (2002). The long step rule in the bounded-variable dual simplex method: numerical experiments. *Mathematical Methods of Operations Research*, 55(3):413–429.

Liu, J. (2021). Mapping high resolution national daily NO2 exposure across mainland China using an ensemble algorithm. *Environmental Pollution*, 279:116932.

Liu, X., Zhu, H., Hu, Y., Feng, S., Chu, Y., Wu, Y., et al. (2016). Public's health risk awareness on urban air pollution in Chinese megacities: The cases of Shanghai, Wuhan and Nanchang. *International Journal of Environmental Research and Public Health*, 13(9). https://doi.org/10.3390/ijerph13090845

Mehmood, K., Bao, Y., Petropoulos, G. P., Abbas, R., Abrar, M. M., Mustafa, A., et al. (2021). Investigating connections between COVID-19 pandemic, air pollution and community interventions for Pakistan employing geoinformation technologies. *Chemosphere*, 272, 129809.

Mitchell, S., Consulting, S. M., & Dunning, I. (2011). PuLP: A Linear Programming Toolkit for Python.

Modak, P. M., & Lohani, B. N. (1985). Optimization of ambient air quality monitoring networks. *Environmental monitoring and assessment*, 5(1):1–19.

Noll, K. E., Miller, T. L., Norco, J. E., & Raufer, R. K. (1977). An objective air monitoring site selection methodology for large point sources. Atmospheric Environment (1967), 11(11):1051–1059. https://doi.org/10.1016/0004-6981(77)90234-7

Park, H., Jeong, S., Park, H., Labzovskii, L. D., e Bowman, K. W. (2021). An assessment of emission characteristics of Northern Hemisphere cities using spaceborne observations of CO2, CO, and NO2. *Remote Sensing of Environment*, 254, 112246.

Sarigiannis, D. A., & Saisana, M. (2008). Multi-objective optimization of air quality monitoring. *Environmental Monitoring and Assessment*, 136(1–3):87–99. https://doi.org/10.1007/s10661-007-9725-z

Verhoelst, T., Compernolle, S., Pinardi, G., Lambert, J. C., Eskes, H. J., Eichmann, K. U., et al. (2021). Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. *Atmospheric Measurement Techniques*, 14(1):481–510. https://doi.org/10.5194/amt-14-481-2021

Wang, K., Ding, Y., Zhao, H., Hou, L., e Sun, F. (2013). Optimization of air pollutant monitoring stations based on genetic algorithm. Proceedings - 4th International Conference on Emerging Intelligent Data and Web Technologies, EIDWT 2013, 680–684. https://doi.org/10.1109/EIDWT.2013.124

Zheng, J., Feng, X., Liu, P., Zhong, L., e Lai, S. (2011). Site location optimization of regional air quality monitoring network in china: Methodology and case study. *Journal of Environmental Monitoring*, 13(11):3185–3195. https://doi.org/10.1039/c1em10560d

DOI: 10.59254/sbpo-2023-174873
http://live-proceedings.pantheonsite.io/p/174873