Stoch Environ Res Risk Assess (2016) 30:779–793
DOI 10.1007/s00477-015-1189-z

ORIGINAL PAPER

Design of a sensitive air quality monitoring network using an integrated optimization approach

Khaled Zoroufchi Benis[1] · Esmaeil Fatehifar[1] · Sirous Shafiei[1] · Fatemeh Keivani Nahr[1] · Yaser Purfarhadi[1]

Published online: 9 December 2015

Abstract A new method is developed to design a multi-objective and multi-pollutant sensitive air quality monitoring network (AQMN) for an industrial district. A dispersion model is employed to estimate the ground level concentration of the air pollutants emitted from different emission sources. The primary objective of AQMN is providing the maximum information about the pollutant with respect to (1) maximum coverage area, (2) maximum detection of violations over ambient air standards and (3) sensitivity of monitoring stations to emission sources. Ant Colony Optimization algorithm (ACO) and Genetic Algorithm (GA) are adopted as the optimization tools to identify the optimal configuration of the monitoring network. The comparison between the results of ACO and GA shows that the performance of both algorithms is acceptable in finding the optimal configuration of AQMN. The application of the method to a network of existing refinery stacks indicates that three stations are suitable to cover the study area. The sensitivity of the three optimal station locations to emission sources is investigated and a database including the sensitivity of stations to each source is created.

Keywords Air quality monitoring network · Ant colony optimization algorithm · Genetic algorithm · Sensitivity of monitoring station

✉ Esmaeil Fatehifar
fatehifar@gmail.com; fatehifar@sut.ac.ir

[1] Faculty of Chemical Engineering, Environmental Engineering Research Center, Sahand University of Technology, P.O. Box: 51335-1995, Tabriz, Iran

1 Introduction

Air quality monitoring plays a crucial role in developing policies and strategies in order to achieve the aims of environmental policies. The ultimate goal of air quality monitoring is to collect data by which scientists, politicians and planners are enabled to make appropriate decisions in managing and boosting global environmental quality (Gurjar et al. 2010; Liu et al. 1986; Mofarrah and Husain 2010). Recent developments in online air quality monitoring and short-term measuring have led to the introduction of advanced warning systems and immediate notifications. These systems can reduce emissions during pollution episodes and help vulnerable inhabitants to cope with these conditions. Such developments, and also successful air pollution control policies, have led to dramatic improvements in air quality, public health, and life quality over the last few decades (Hsu et al. 2013; Kuhlbusch et al. 2014).

The aim of design of air quality monitoring network (AQMN) is to determine the number and locations of stations (configuration) which is highly essential in achieving the air pollution control. Considering the high cost of monitoring stations, i.e. equipment, maintenance and operating personnel, optimization of AQMN is crucial for air pollution control managers (Bayraktar and Turalioglu 2005; Modak and Lohani 1985a; Mofarrah and Husain 2010).

The primary tasks in the optimization of AQMN are based on empirical and quantitative approaches (Kao and Hsieh 2006; Nejadkoorki and Baroutian 2012). However, recent literatures have shown increasing tendencies to apply multi-objective and systematic approaches in the design of AQMN (Gómez-Losada et al. 2014; Liu et al. 1986; Lozano et al. 2009a, b, 2010; Maria Grazia et al.

1999; Mazzeo and Venegas 2008; Nejadkoorki et al. 2011; Pires et al. 2008a; Zarandi et al. 2008). Table 1 summarizes such descriptive studies.

The studies done in the field of AQMN design can be classified based on: (1) target objectives for the network, (2) intended pollutants for monitoring, (3) method to determine the concentration distribution of pollutants over the study area, (4) method of finding the number and locations of stations (design technique), and (5) intended areas for monitoring.

According to Table 1, most of the studies have proposed multi-objective planning methods and cost reduction has been considered as a sub-objective. In other words, the design and/or optimization aim has been to achieve the desired objective(s) with minimum cost, i.e. number of stations. Measuring the maximum concentrations is usually one of the general goals of AQMN (Zhang et al. 2014). The common fault of the methods when this objective is considered as the only criterion in the design procedure is that the proposed stations will not be truly representative of the whole area.

Having in mind the high cost of stations simultaneous measurement of several pollutants at one station is preferred. Therefore, newly proposed network design methods have been established in a way in which multiple pollutants are considered. In addition, if more than one pollutant is measured at a common station, estimating missing values will be possible by using cross-correlations between the other pollutants (Sarigiannis and Saisana 2008). Various techniques have been used to determine the concentration distribution of pollutants over the studied area. In the case of AQMN design, the area is usually divided into grids and then the pollutants concentration is determined for each grid. The centroid of each grid is considered as a potential location for a monitoring station (Nejadkoorki et al. 2011). The pollutant concentration in each grid is estimated by using different methods. Dispersion models with respect to the distribution of pollution sources are used to determine the concentration of pollutants in each grid in the source-oriented siting procedures. Sometimes, available input data makes the application of dispersion models and complex numerical tools impossible; thus, the statistical approaches such as interpolation techniques and Land Use Regression (LUR) models are used to estimate the pollutants concentration in which the available data like existing monitoring stations, environmental measurement data, and satellite turbidity data are considered as input data (Elkamel et al. 2008; Gerharz and Pebesma 2013; Liu et al. 2014b; Sarigiannis and Saisana 2008). Generally, in most methods, a score is given to each grid based on the calculated concentrations. In the case of multi-objective design (e.g., monitoring areas of high population density and detection of violations over ambient standards) another score is given to each grid based on the considered objective. After this stage, each of the presented methods will take different route. In most studies as seen in Table 1, after determining the objective function and its value for each grid, the grid ranking (ranking potential station locations) is used to select the optimal location. In these methods, simultaneous determination of the optimal number and locations of the stations is difficult and usually the number of stations has been specified in advance based on budget constraints. The methods in which the number and locations of the stations are offered simultaneously, i.e. joint solution, would be more practical. Another major difference between the proposed methods is the type of monitoring area, where different types of monitoring areas (such as industrial or urban) will lead to different objectives, pollutants, dispersion models, and design and/or optimization techniques.

Emission of wide variety of pollutants from industrial plants such as oil refineries has made them one of the largest sources of air pollution (Abdul-Wahab et al. 2011). The present study introduces an integrated approach to design a sensitive AQMN for industrial plants (i.e. oil refinery and petrochemical plants). This method optimizes the number and locations of air quality monitoring stations with respect to the sensitivity of monitoring stations to emission sources. The sensitive monitoring station has the capability of quick diagnosis of stochastic accidents in the emission sources which lead to an increase in the pollutants concentration. This character attribute of monitoring stations enables the authorities to specify the defective emission sources and to take prompt measures to control them. The optimization algorithm is developed considering (1) maximum coverage area of the monitoring network, (2) continuity of coverage area, (3) least overlap among coverage areas, (4) maximum detection of violations over ambient air standards, and (5) sensitivity of monitoring stations to emission sources. The developed integrated method leads to a source-specific monitoring network which can be used effectively to control emission sources.

The multiple cell approach (MCA), which was described previously (Fatehifar et al. 2006, 2007, 2008) was expanded and used to simulate air pollution dispersion and determine ground level concentrations. The outputs of the MCA model were used as input data for calculation of the fitness function of optimization algorithms in order to find the optimal configuration of AQMN. The solution with the least value of the fitness function offers the best configuration of monitoring stations. In order to get the best solution to the problem, Ant Colony Optimization Algorithm (ACO) and Genetic Algorithm (GA), integrated with the MCA model, were used separately as the optimization tools. ACO and GA are robust methods capable of locating near global optimal solutions for complex problems. The proposed methodology was implemented as a MATLAB program which is flexible and expandable.

Table 1 Systematic approaches in AQMN designing

No.	AQMN objectives	Considered pollutants	Pollutants concentration/ dispersion model	Design technique	Area/place	Ref.
1	Detection of violations over ambient standards	SO_2	Diffusion model	Ranking potential air monitoring sites	Urban area/Piacenza, Italy	Finzi et al. (1991)
2	Prediction of the spatial and temporal patterns of the concentration field Detection of violations over ambient standards	Single pollutant	ISCST	Spatial correlation analysis technique/concept of potential of violation	Around a hypothetical potash plant	Arbeloa et al. (1993)
3	Monitoring areas of high population density Measuring the maximum concentrations Highest detection of the violations	Overall	ISC	Genetic algorithm	Urban area/Kaohsiung, Taiwan	(Chang and Tseng (1999a), Chang and Tseng (1999b))
4	Measuring the maximum concentrations	Overall	Measuring the concentration of air pollutant with a mobile analytical laboratory/ISC3	Ranking potential air monitoring sites	Urban area/Salerno, Italy	Corti and Senatore (2000)
5	Assessments of adverse human health impacts from exposures to airborne contaminants	Airborne contaminants	ISC3	Ranking potential air monitoring sites/Risk assessment techniques	Urban area	Baldauf et al. (2001, 2002)
6	Providing adequate information for administrative purposes	CO, SO_2, PM_{10}, O_3	Available monitoring data	Shannon information index	Urban area/Santiago, Chile	Silva and Quiroz (2003)
7	Detection of violations over ambient standards Monitoring areas of high population density	NO_2	Available monitoring data/linear regression model	L-A procedure/Using the ESRI's ARC/INFO software	Urban area/Toronto, Canada	Kanaroglou et al. (2005) and Harris et al. (2014) for water quality network
8	Environmental objectives: monitoring highest pollutant (and average) concentration monitoring Largest violations and violations over ambient standards monitoring Largest emission sources (total, industrial, and mobile sources) Social objectives: monitoring Largest population, sensitive receptors, traffic volume, and air pollution petitions	PM_{10}, SO_X, NO_X, NMHC	ISC3	The modified bounded implicit enumeration algorithm/the constraint arrangement method	Urban area/Taoyuan, Taiwan	Chen et al. (2006)

Table 1 continued

No.	AQMN objectives	Considered pollutants	Pollutants concentration/dispersion model	Design technique	Area/place	Ref.
9	Detection pollutant concentration Measuring the maximum concentrations Maximum coverage area Maximum population protection	Overall	ISCST3	Trade-off relationship between objectives/using concept SOI	Industrial district/Toufen Industrial District in Taiwan	Kao and Hsieh (2006)
10	Providing maximum information about multi-pollutants Detection of violations over ambient standards	Overall	A mathematical model based on the multiple cell approach	Spatial correlation analysis and the minimum spanning tree-Sequential interactive compromise	Around an oil refinery/Toufen Industrial District in Taiwan	Elkamel et al. (2008)
11	Detection of violations over reference level	NO_X	DAUMOD, ISCST3	Ranking potential air monitoring sites	Urban area/Buenos Aires, Argentine	Mazzeo and Venegas (2008)
12	Monitoring areas with similar air pollution behavior	SO_2, PM_{10}, CO, NO_2, O_3	Available monitoring data	Principal component analysis (PCA) and cluster analysis (CA)	Urban area/Oporto, Portugal	Pires et al. (2008a, b)
13	Detection of violations over reference level Maximum coverage area Maximum gain of information, in terms of exposure of population, land-use and vulnerable cultural	Overall	Atmospheric turbidity as surrogate for air pollution loading	Linear regression analysis/ranking potential air monitoring sites	Urban area/Brescia, Italy	Sarigiannis and Saisana (2008)
14	Measuring the maximum concentrations	Benzene, Toluene, n-Hexane	Sampling data/interpolation methods	Ranking potential air monitoring sites	Urban area/Murcia, Spain	Ferradás et al. (2010)
15	Measuring the maximum concentrations	NO_2, O_3	Sampling data/interpolation methods	Ranking potential air monitoring sites	Urban area/Jaen, Spain	Lozano et al. (2010)
16	Monitoring areas of high population density	SO_2, CO, NO_X, PM	ISC3	Triangular fuzzy numbers, ranking potential air monitoring sites/using concept SOI	Urban area/Riyadh, Saudi Arabia	Mofarrah and Husain (2010)
17	Capture the maximum variance in measured pollutants	TSP	Available monitoring data/interpolation methods	Ranking potential air monitoring sites	Urban area/Yazd, Iran	Nejadkoorki et al. (2011)
18	Quick detection of areas with highest emission variability	Heavy metals, PAHs, PM_{10}, $PM_{2.5}$, Mercury	Geographic information system/national air quality model MINNI	Neighbourhood statistic function	Italian national monitoring stations/Northern and central Italy	Righini et al. (2014)

2 Materials and methods

2.1 Air pollutant dispersion model

In this study, the multiple cell approach (MCA) based on the solution of the three-dimensional diffusion equation is used to predict ground level pollutant concentrations. The model based on the previous papers (Fatehifar et al. 2006, 2007, 2008) is developed in a way that can consider wind direction in the prediction of ground level pollutant concentrations in order to make it more adapted for AQMN optimization. However, a brief description of its main considerations and assumptions is included below.

The program of the model uses MATLAB-based graphical user interface (GUI) and it is applicable for network of refinery stacks, petrochemical complexes, urban and industrial stacks. The model is based on Eulerian Model and k-theory in which the advection and diffusion of pollutants from point sources in Cartesian Coordinates (considering constant wind velocity and turbulent diffusivities) are expressed by a advective–diffusive equation (Fatehifar et al. 2008). The simplifying assumptions of the model are:

1. Steady state conditions.
2. The initial conditions are arbitrarily set at zero.
3. Transport by bulk in wind direction exceeds diffusion in that direction, i.e., the x-axis eddy diffusion coefficients of the model were neglected.
4. The wind velocity is constant, a function of z and only in x direction.
5. There is no chemical reaction and no deposition in the system.

Figure 1 shows the schematic view of selected domains for modeling of pollutants dispersion. As can be seen, two domains for modeling procedure are introduced: main and modelling domains.

The main domain is defined in a way which includes all the area that pollution gets dispersed over, in different wind directions. The modeling domain is a moving domain which can rotate based on the wind direction, around the central point of the main domain. The network of the point sources is located in the center of the main domain.

The empirical equations (which are dependent on the stability classes of atmosphere and functions of surface roughness length and friction velocity) are used to calculate eddy diffusivities in the lateral and vertical directions and wind velocity at different heights. The modified Holland's equation is used for plume rise calculation. Application of this model to industrial stacks has shown results that agree with observational data within a reasonable accuracy (Fatehifar et al. 2006, 2007, 2008).

2.2 Objectives of the air quality monitoring network

The main goal of designing an AQMN is to collect the most possible information in the least expensive way. So, the monitoring stations must be located in the points where they properly represent pollutant concentration of the relevant surrounding area and provide the maximum coverage area with the minimum number of stations. Therefore, the minimum overlap among covered areas of stations (which leads to the maximum coverage area) can be considered as an objective. Furthermore, the stations must be located in areas which have the maximum air pollutant concentration and also, high fluctuations of pollutant concentration which leads to the estimation of sensitivity of stations to pollutants sources. Maximum population coverage can also be considered when resident safety is a concern, but considering the studied area, which is located far from urban and

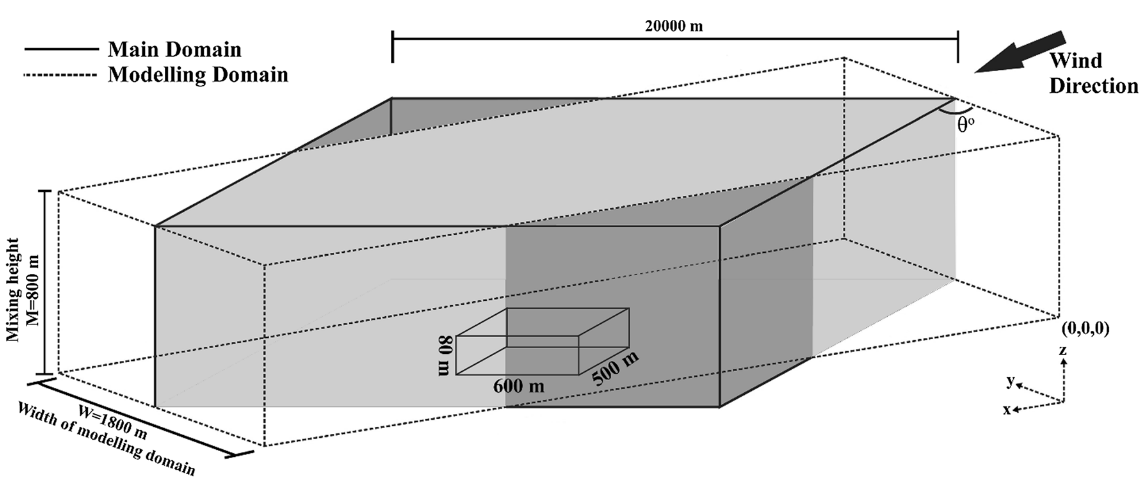


Fig. 1 Schematic of selected domains for modeling

high density areas, population protection is not an objective in this study.

2.2.1 The coverage area

To determine the coverage area of the stations, the concept of sphere of influence (SOI) along with the continuity of coverage is used. The continuity of coverage area leads to an increment of accuracy of monitoring. The concept of SOI has been widely used in the literature (Baldauf et al. 2001, 2002; Chang and Tseng 1999a, b; Elkamel et al. 2008; Kao and Hsieh 2006; Liu et al. 1986; Mofarrah and Husain 2010; Mofarrah et al. 2011; Arbeloa et al. 1993; Tseng and Chang 2001) without considering the continuity of coverage area.

The contiguous SOI of a station is defined as the surrounding area of the station where the air quality data can be considered as representative of the whole area. It is based on the similarity between the information contained in a given station compared with the rest of area. To calculate the contiguous SOI, a spatial correlation coefficient (r) based on the extent of similarity in ground level pollutants concentration is used. The ground level pollutants concentration are estimated by the MCA model. Presuming that $C_1 = (C_{11}, C_{12}, C_{13}, \ldots, C_{1n})$ and $C_2 = (C_{21}, C_{22}, C_{23}, \ldots, C_{2n})$ are the indicators of pollutant concentration in two various network points calculated by the MCA model at the same time, the spatial correlation coefficient for a sample size n can be expressed as:

$$r = \frac{\sum_{i=1}^{n} (C_{1i} - \bar{C}_1)(C_{2i} - \bar{C}_2)}{\sqrt{\sum_{i=1}^{n} (C_{1i} - \bar{C}_1)^2 \sum_{i=1}^{n} (C_{2i} - \bar{C}_2)^2}} \quad (1)$$

where, $\bar{C}_1$ and $\bar{C}_2$ are the average concentrations at locations 1 and 2, respectively.

$$\bar{C}_1 = \frac{1}{n}\sum_{i=1}^{n} C_{1i} \quad and \quad \bar{C}_2 = \frac{1}{n}\sum_{i=1}^{n} C_{2i} \quad (2)$$

To calculate contiguous SOI in each point, the r value obtained for each point is compared to a predefined value called cut-off value (r_c). If the spatial correlation coefficient (r) is higher than r_c, the corresponding points, which are interconnected to each other (Fig. 2), would be considered as correlated. The coverage area of ith candidate location is defined as the number of potential locations placed inside the contiguous SOI of candidate location i can be quantified in terms of pattern scores. Pattern score for an ith candidate location is denoted by N_p^i.

Therefore, determining the optimum number and locations of monitoring stations, where the overall pattern score is maximized, is the first interest in AQMN design. High

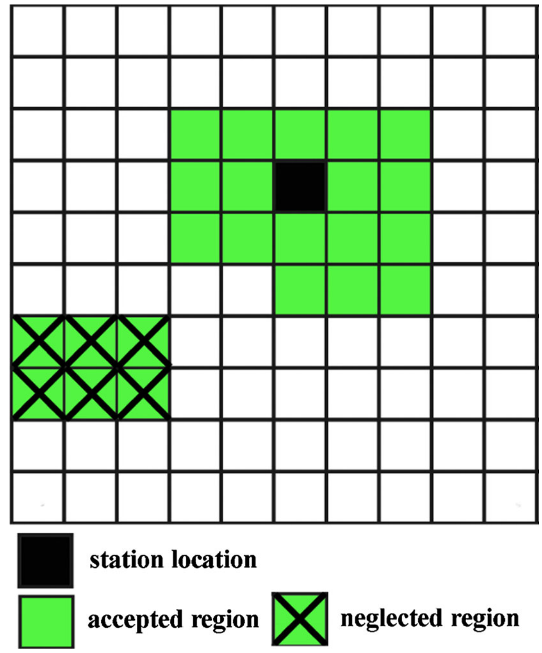


Fig. 2 Continuity of SOI

overall pattern score leads to the optimal coverage area. This objective of the design can be expressed as:

Maximize $\sum_{i=1}^{m} N_p^i$.where, m is the number of stations.

2.2.2 Violation over ambient standards

The concept of violation score is used to determine the extent of violation over ambient standards for each point: the point with higher violation score is assumed to have a high potential of detecting the violations and a high sensitivity to emission sources. A weighted scoring is used to determine the violation scores based on the concentration of pollutants which depends on:

1. The threshold level of pollutant concentration;
2. Weighing factor between each threshold range and weighing function.

A decision on threshold levels and weighing factors is indeed pollutant-specific and dependent upon the considered average time. Several weighing functions such as linear, segmented linear, non-linear, and segmented non-linear can be employed (Elkamel et al. 2008). In this study, a segmented non-linear function is used. The violation score for each candidate location is calculated using the equation:

$$N_v^i = \sum_{i=1}^{T} \sum_{k=1}^{N_t} \frac{(w_{k+1} - w_k)(x_i - x_k)X}{(x_{k+1} - x_k)} \quad (3)$$

where, N_v^i is the violation score for the ith candidate location, w_k is the weighing factor corresponding to threshold x_k, x_k is the kth threshold, $X = 0$ if $(\mathrm{x_i} - \mathrm{x_k}) \leq 0$, otherwise $X = 1$, N_t is the total number of thresholds and T is the total number of simulated observations. The threshold values are chosen based on the United States Environmental Protection Agency (USEPA) limit values of common pollutants (Elkamel et al. 2008; Sarigiannis and Saisana 2008).

The second interest of optimization could therefore be deliberated as the identification of the optimum number (m) and locations of monitoring sites so that the overall violation score is maximized. High overall violation score leads to more sensitivity of stations to pollutants sources and achieving a maximum detection of violations over ambient air standards. This objective of the designing can be expressed as:

Maximize $\sum_{i=1}^{m} N_V^i$

where m is the number of stations.

2.2.3 *Sensitivity of candidate locations to emission sources*

The sensitivity of candidate locations to emission sources is calculated using the following procedure for every meteorological scenario:

Initially, the MCA model is used to calculate the ground level concentration of pollutant at each point for the eth meteorological scenario using the average concentration of pollutant emitted from emission sources. In the next step, the concentration of pollutant is raised 25 % in the first source (concentrations of other sources are constant) and the ground level concentration of pollutant is obtained (using MCA model) at each point again. This procedure is done accordingly for all of the sources at 25, 50, 75 and 100 %. Finally, the score of sensitivity of candidate locations to emission sources are calculated using the following equation:

$$N_s^i = \sum_{e=1}^{E} \sum_{s=1}^{S} Sl_{es} \tag{4}$$

where, N_s^i is the sensitivity score for the ith candidate location, Sl_{es} is the slope of concentration variation at ith candidate location versus concentration changes in pollution source s and meteorological scenario e, S is the number of emission sources, and E is the number of meteorological scenarios.

This objective of the designing can be expressed as:

Maximize $\sum_{i=1}^{m} N_s^i$

where m is the number of stations.

2.2.4 *Fitness function*

Given that having a network with high overall pattern score, overall violation score, and overall sensitivity score is desirable, a combination of functions is used to achieve the simultaneous maximum pattern score, violation score, and sensitivity score. Several forms of combinations (e.g., additive or multiplicative form) could be used for combining these independent variables, out of which the multiplicative form is used:

$$F = (N_p)^{b1} \times (N_v)^{b2} \times (N_s)^{b3} \quad (where\, b_1 + b_2 + b_3 = 1) \tag{5}$$

The parameters b_1, b_2, and b_3 are used to weigh the relative importance given to the three objectives and their value depends on the purpose of the network to be developed. The multiplicative form implies an intensity function, and for positive high correlations, the interests for a location are quite well represented by the product ($N_p \times N_v \times N_S$), rather than the addition (Elkamel et al. 2008; Modak and Lohani 1985b). Also this structure of the objective function ensures that if any of the three sub-objectives become zero, the overall objective function (F) will also has a zero value. For example, if a potential station location has a high violation score, but zero effective pattern score, then it will be discarded in the optimal design.

It should be noted that the effect of weighting factors on F, N_p, N_v, and N_s values of each pollutant were investigated to give appropriate weight value for each component.

Since the objective of this research is to design an AQMN for multiple pollutants, an aggregating approach is used:

$$UF_i(u) = \sum_{j=1}^{p} w_j F_j^i \tag{6}$$

and

$$\sum_{j=1}^{p} w_j = 1 \tag{7}$$

where, w_j is the importance associated with pollutant j, F_j^i is the objective function for pollutant j at location i and $UF_i(u)$ is the additive/overall objective for P pollutants at location i. Considering that each pollutant had the same importance (in this case study), the equal weight values were used to give the same tendency for each pollutant.

Now, given the defined equations, the objective of the network design is to achieve the maximum sum of the UF_i of stations. In order to reach this maximum value, the minimum number of overlap points among the coverage areas of stations (which is inconsistent with this maximum value) is used. For the sake of design a fitness function of ACO and GA the two dimensionless parameters are defined as:

$$UF_n = \frac{UF_{\max} \times m}{\sum_{i=1}^{m} UF_i(u)} \tag{8}$$

$$OL_n = \frac{number\ of\ overlaped\ grids}{maximum\ number\ of\ possible\ overlaped\ grids} \quad (9)$$

where, UF_n is Overall Objective number of network, OL_n is Overlap Number, m is the number of stations and UF_{max} is the maximum value of $UF(u)$ among all $UF(u)$s of the potential station locations.

Finally, since obtaining least UF_n and OL_n simultaneously is desirable in network design (leading to more coverage area), an aggregate function, is used as follow:

$$G = g_1 \times OL_n + g_2 \times UF_n \quad (g_1 + g_2 = 1) \quad (10)$$

where, g_1 and g_2 are the weight coefficients defined according to the relative importance of each parameter. The variation of the coverage efficiency versus g_1 ($g_2 = 1 - g_1$) was investigated in order to find the best weight values.

In order to find the least value of G where the sum of the UF_i of stations is maximum and the overlap among the coverage areas of stations is minimum, GA and ACO are used separately and then the results are compared in order to confirm the obtained results.

ACO is a metaheuristic optimization algorithm based on the behavior of ants seeking a path between their colony and a source of food. This collective trail-laying and trail-following behavior in which an ant is influenced by the pheromone trail left by other ants is the source of inspiration for ACO. This algorithm benefits from artificial ants which deposit pheromone based on the fitness and goodness of the identified trial solutions (Emami Skardi et al. 2015; Fayaed et al. 2013). The GA is a stochastic search algorithm, which mimics evolution in nature and it is an efficient tool in searching for the global optimum. This algorithm is widely used to optimize functions in various geocomputational applications (Liu et al. 2014a; Shad et al. 2009).

Overall, the AQMN designing procedure with sensitivity of monitoring stations to emission sources can be summarized by a flowchart as shown in Fig. 3.

3 Illustrative case

The algorithm defined above was used for optimizing AQMN around Tabriz Oil Refining Company. The company is located in the southwestern part of Tabriz, East Azerbaijan province, Iran. The refinery has 14 refining units and 10 utility units which include 26 furnaces. Given that the exhausts of some of furnaces are joined to each other, 20 stacks disperse the exhaust pollutants over the surrounding area. The MCA model was used to create spatial distributions for the concentrations of the pollutants (SO_2, NO_x, CO) using 1 year's collected data. Normally 1 year is considered to be the minimum period which can reflect the meteorological fluctuations (Elkamel et al. 2008; Arbeloa et al. 1993). A flue gas analyzer (MRU Varioplus, Germany) was used to measure the temperature, velocity and concentration of the flue gas species. The measurement was done periodically every week during the year 2012. Totally, 48 data sets were collected. The other information of emission source that needs to be input into the model are the physical stack dimensions such as height, location and internal diameter. It should be noted that the effect of background concentration is negligible with regard to the refinery location and prevailing wind directions. Table 2 summarizes the emission sources information.

In addition, the meteorological data such as wind speed, wind direction, ambient temperature, height of the mixing layer, and stability class, registered at a closest weather station of Tabriz Meteorological Organization, were used as the input data. Figure 3a shows the location of Tabriz Oil Refining Company and the positions of air pollution sources. Location 1 includes 18 sources and each of the locations 2 and 3 include only one source. Investigation of the wind regimes of the study area between the years 2005 and 2012 indicated that the prevailing wind directions throughout the study area were eastern (90°) and north eastern (45°) for more than 80 % of the time. Thus, only these two directions were used in the modelling and optimization procedures. Tabriz has a semi-arid climate with dry and semi-hot summer and snowy cold winter. The temperature varied from −12 and 39 °C during the year 2012. It is worth noting that Tabriz Oil Refining Company is located at a smooth plate.

4 Results and discussion

At the beginning of AQMN designing, the study area was divided into 100 × 100 m continuous grids in which each grid represented a potential site for monitoring stations. According to the prevailing wind directions throughout the study area, the area was divided into 3245 grids (59 × 55) and they were numbered sequentially according to Fig. 4a. The MCA model was used to model the distribution of pollutants concentrations in the area based on the 48 collected data. A matrix of 48 × 3245 for SO_2, NO_x, and CO concentrations that has been generated at the specified 3245 candidate locations was used as an input to the optimization algorithm. Figure 4b schematically shows the modeling grids, location of sources and potential zones for locating monitoring stations according to the prevailing wind directions.

As an example, Fig. 5 shows the ground level concentration of pollutant for a selected domain at different wind directions and meteorological conditions.

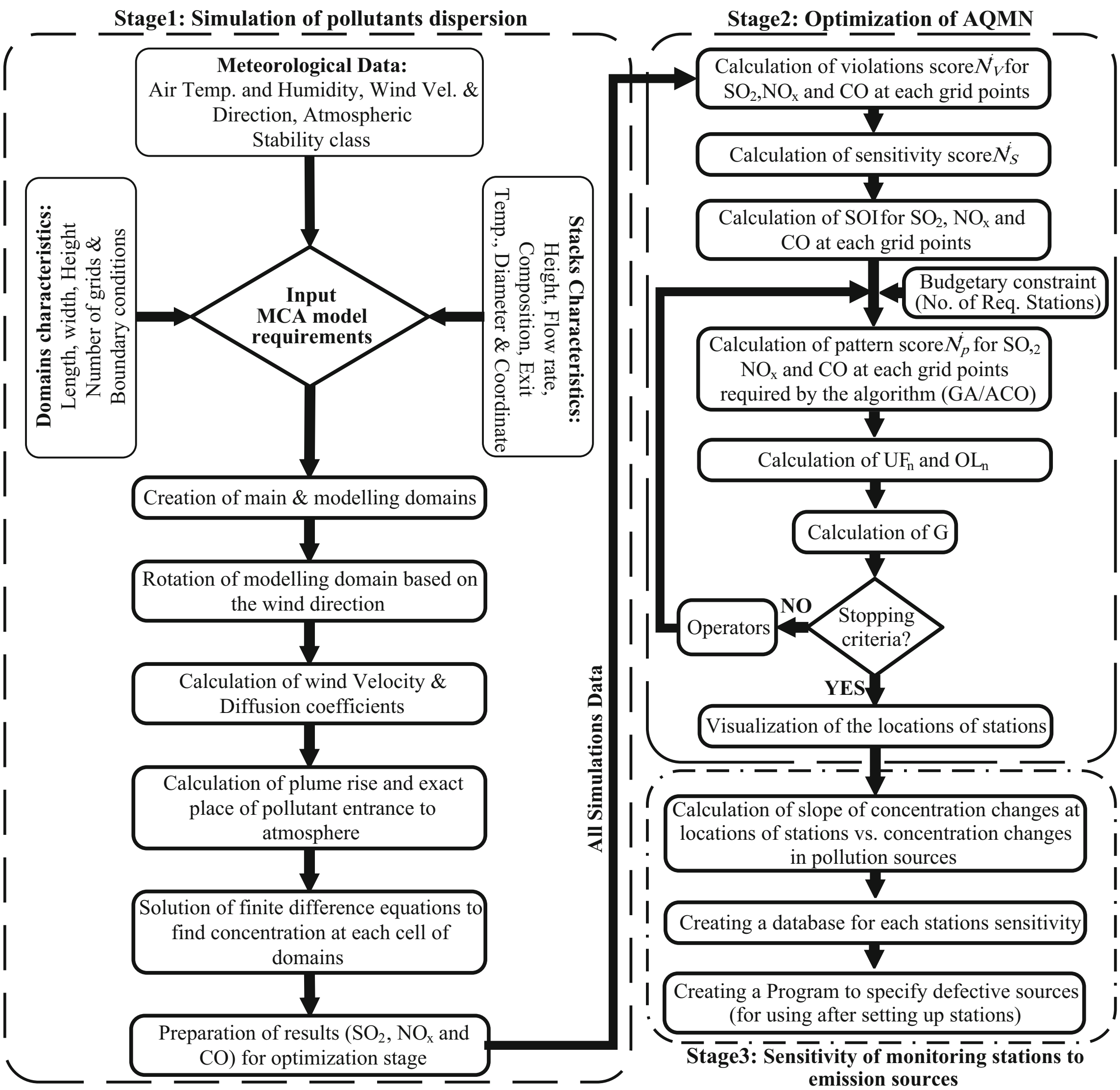


Fig. 3 Flowchart of proposed method for optimal allocation of monitoring stations with sensitivity to emission sources

Given the air pollution dispersion, a “potential zone” is recommended for locating the stations. A potential zone is an area in which the pollutant concentrations are higher than a predefined threshold value (Kao and Hsieh 2006). According to the Fig. 5, a potential zone can be diagnosed from the pollutant concentration profile at ground level of simulation examples.

Considering the intended objectives for the AQMN, the candidate locations with high value of N_p, N_v and N_S will occur in the potential zone. So, the output results of the MCA model determine the location of the potential zone in each modelling scenario and subsequently the behavior of integrated optimization algorithms (ACO and GA). The length and the distance of the potential zone from the emission sources are dependent on the pattern of pollutant concentration which is a function of meteorological conditions and emission rates. Accordingly, the effects of important meteorological parameters such as wind velocity, atmospheric stability, air temperature and dispersion coefficient and also effective parameters of emission sources on the dispersion of pollution and the location of the potential zone were investigated. The results showed that when the wind velocity increases, the pollution dispersion decreases and pollutants go far from the stacks. The distribution of pollutants increases with increasing atmospheric instability, so the pollutants do not go far from the emission sources. The dispersion of pollutants increases with increasing air temperature and pollutants come down near the emission sources. Increasing the stack exit velocity and temperature leads to a decrease of the ground level concentration. Therefore, it can be concluded that the distance of potential zone from the emission sources

Table 2 Information of Emission sources

Source	Coordinate** x	Coordinate** y	Stack height (m)	Stack diameter (m)	Stack temperature (°C)*	Stack velocity (m/s)*	SO_2 CON (ppm)*	CO CON (ppm)*	NO_x CON (ppm)*
1	0	655	73	3.5	179.4	6.8	148.3	17.3	101.0
2	24	655	73	3.5	146.4	9.4	14.4	14.2	51.3
3	46	655	73	3.5	153.6	6.1	50.6	3.3	19.6
4	125	674	36.6	1.9	327.9	5.4	26.5	11.2	56.2
5	128	854	43	3.57	308.2	6.9	30.5	25.6	91.7
6	139	854	36.6	2.18	337.8	10.1	27.5	32.5	70.8
7	162	677	36.6	0.92	407.6	6.2	17.3	91.4	66.5
8	162	685	46	1.81	290.6	4.2	3.8	7.8	62.0
9	162	694	46	2.18	242.4	6.0	3.3	22.2	71.2
10	220	667	36.6	2.2	565.9	10.6	56.6	194.6	62.3
11	246	667	36.6	4.35	488.3	8.6	39.4	56.3	60.4
12	272	667	36.6	2.35	310.6	2.6	13.7	20.7	41.5
13	365	676	52	2.52	442.1	4.1	3.8	252.4	44.9
14	406	672	73.2	3.58	167.3	5.3	78.1	75.6	55.3
15	396	672	36.6	1.5	549.7	6.6	144.5	29.3	54.7
16	429	677	52	2.38	345.3	6.9	224.5	13.2	55.6
17	435	677	53	1.5	374.9	5.4	148.0	6.7	49.6
18	436	817	36.6	1.58	365.9	7.2	52.3	23.9	54.7
19	351	667	36.6	3	586.6	5.4	72.2	35.6	47.4
20	200	515	60.8	2.35	467.3	6.8	2966.1	1474.4	99.1

* Average of collected values

** Coordinate system transformed to the center of modeling domain

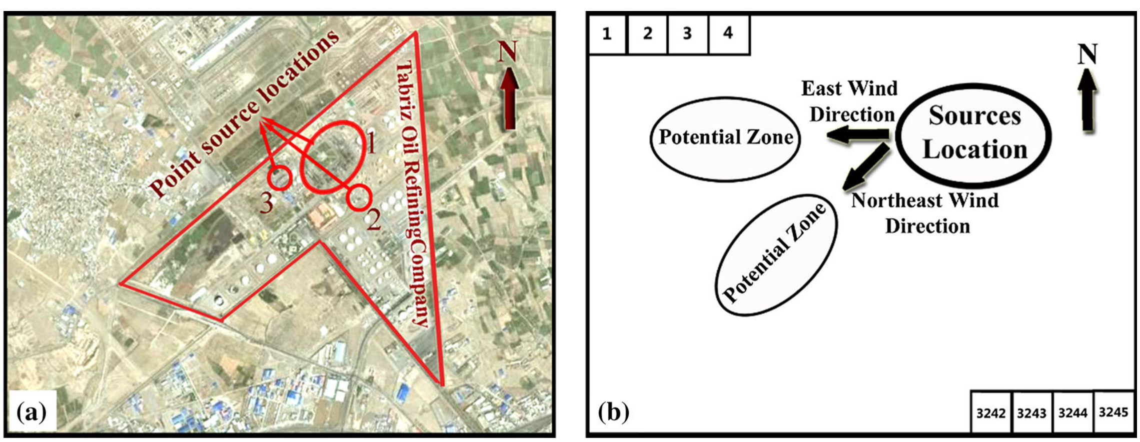


Fig. 4 **a** Location of industrial plant and emission sources, **b** Modeling grids, location of emission sources, potential zones and wind directions

increases by increasing the wind velocity and atmospheric stability, and decreasing the stack exit velocity and temperature and the air temperature.

In order to verify the predictions of the model, a comparison of model output with the measured ambient concentration data was done. The measured data was collected around the refinery simultaneously with stack sampling. The measurement was done twice at downwind distances of 400, 800, and 1200 m from the emission sources. Note that the measurements at sampling points were not done simultaneously.

Figure 6 illustrates a comparison between measured data and model results. As can be seen, there was good agreement between the measured data and modeling

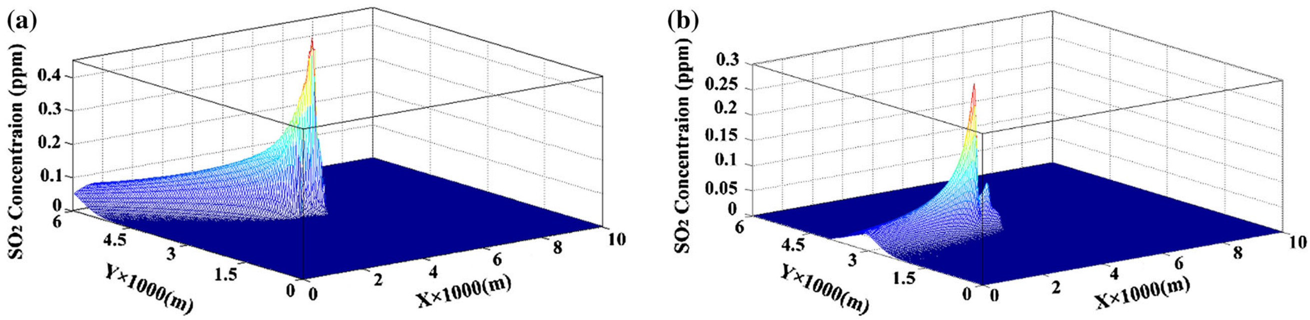
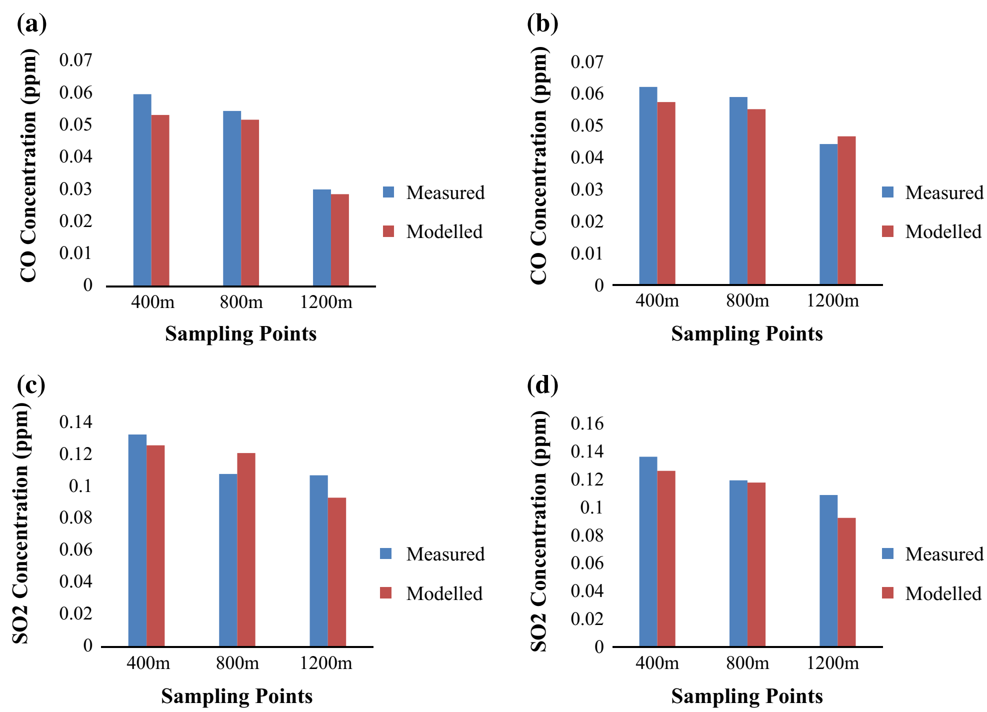


Fig. 5 Concentration of pollutants at ground level with wind blowing direction: **a** 45, **b** 90

Fig. 6 Comparison of measured and modelled concentration **a**, **c** first data sets, **b**, **d** second data sets



results. The mean error between measured and modelled values was less than 10 %.

In order to give appropriate weight value for each component in Eqs. (5) and (10), their sensitivity was investigated. In the case of Eq. (5), the results showed that by increasing the b_1 (decreasing b_2 and b_3) the overall N_v and N_s values of the network decrease and the F value and the coverage efficiency of the network increase. Increasing the b_2 and b_3 (decreasing b_1) showed almost the same trend (decreasing the coverage efficiency and overall N_p of the network). For b_1 values lower than 0.8, the coverage efficiency decreases significantly. Having in mind that the low coverage efficiency is not appropriate for AQMN and it leads to a network with more monitoring stations which is not economically feasible, the 0.8 was considered as the minimum value of b_1 which causes a network with acceptable coverage efficiency. For the b_1 values higher than 0.9 the overlap among coverage areas of stations increases markedly and also the overall N_v and N_s values of the network decrease. Finally in order to have a sensitive monitoring station with high capability of detection of violations over ambient air standards, the equal weight values were used for N_v and N_s ($b_1 = b_2 = 0.1$ and $b_2 = 0.8$). It should be noted that the selected weight values are not applicable to any case study. Because the N_p, N_v and N_s values are dependent to the air pollutants dispersion and the characteristics of emission sources.

The results of sensitivity analysis of Eq. (10) showed that for the g_2 values lower than 0.92 the coverage efficiency (and UF_n) decreased markedly due to the increasing overlap among covered areas of stations. On the other hand for the g_1 values of higher than 0.8 the OL_n and also the coverage efficiency had acceptable values. So, the $g_1 = 0.8$ and $g_2 = 0.92$ were selected as the weight values to reach

Table 3 Optimal locations for air quality monitoring stations, obtained using ACO and GA ($r_c = 0.75$ and $r_c = 0.95$)

Station location		1	2	3	4	5	6
ACO	$r_c = 0.75$	1036					
		720	1036				
		1036	954	1153			
		716	1094	594	1550		
		540	793	859	1559	1440	
		1129	1272	2539	110	1791	654
	$r_c = 0.95$	1152					
		1267	962				
		1383	3071	719			
		1497	906	2081	848		
		1786	1150	1023	840	2146	
		3015	971	2254	2599	781	1387
GA	$r_c = 0.75$	1210					
		853	1210				
		853	1095	1210			
		1210	892	1324	853		
		1210	841	663	2978	1095	
		1210	1393	1154	667	1209	1124
	$r_c = 0.95$	1324					
		1324	853				
		853	1324	1210			
		968	1324	853	1617		
		1324	1502	2494	2075	853	
		1439	841	955	1082	1210	2024

(The numbers indicate grid points according to the Fig. 4b.)

the best configuration of AQMN, in which the optimization results (e.g., Figure 9) confirm the selected weight values.

4.1 Comparison of ACO and GA results

The optimization algorithm was implemented for different cut-off values r_c (0.75–0.95) and different station numbers (1–6). Table 3 shows the locations of stations which were obtained by using ACO and GA for cut-off values 0.75 and 0.95. Due to the different structure of the algorithms, the obtained station locations were different.

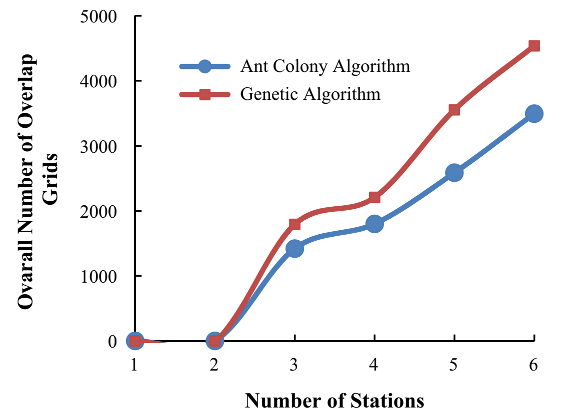


Fig. 8 Overall number of overlapped grids versus number of stations ($r_c = 0.85$)

Figure 7 shows the overall UF_i of Network for different configurations denoted in Table 3. As shown in the figure, by decreasing the r_c, the overall UF_i of network increases, which causes an enlargement of the network coverage area. For a stipulated budget, an air quality monitoring organization can maintain either a high or a low r_c value based network. A high r_c based network may not necessarily cover the entire region, but the covered region could be well represented. A low r_c based network, on the other hand, would offer more coverage of the area, but the covered area may not be satisfactorily represented (Elkamel et al. 2008).

Considering Figs. 7 and 8 it becomes obvious that the overall UF_i and overlap among coverage areas of designed network by GA were higher than that of ACO. However, concurrent comparison of overall UF_i of the network, the overall number of overlapped grids and also the number of covered grids of network showed that both algorithms up to three stations had the same trend. Figure 7 indicates that by increasing the number of stations (from 4 to 6), increasing rate of overall UF_i of network decreased. In contrast Fig. 8 shows a sharp increase of overall number of the overlapped

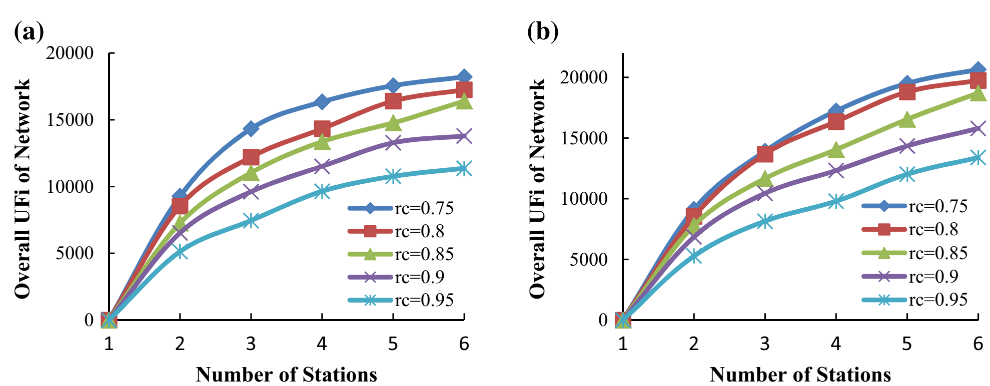


Fig. 7 Overall UF_i of network versus number of stations as a function of r_c **a** ACO **b** GA

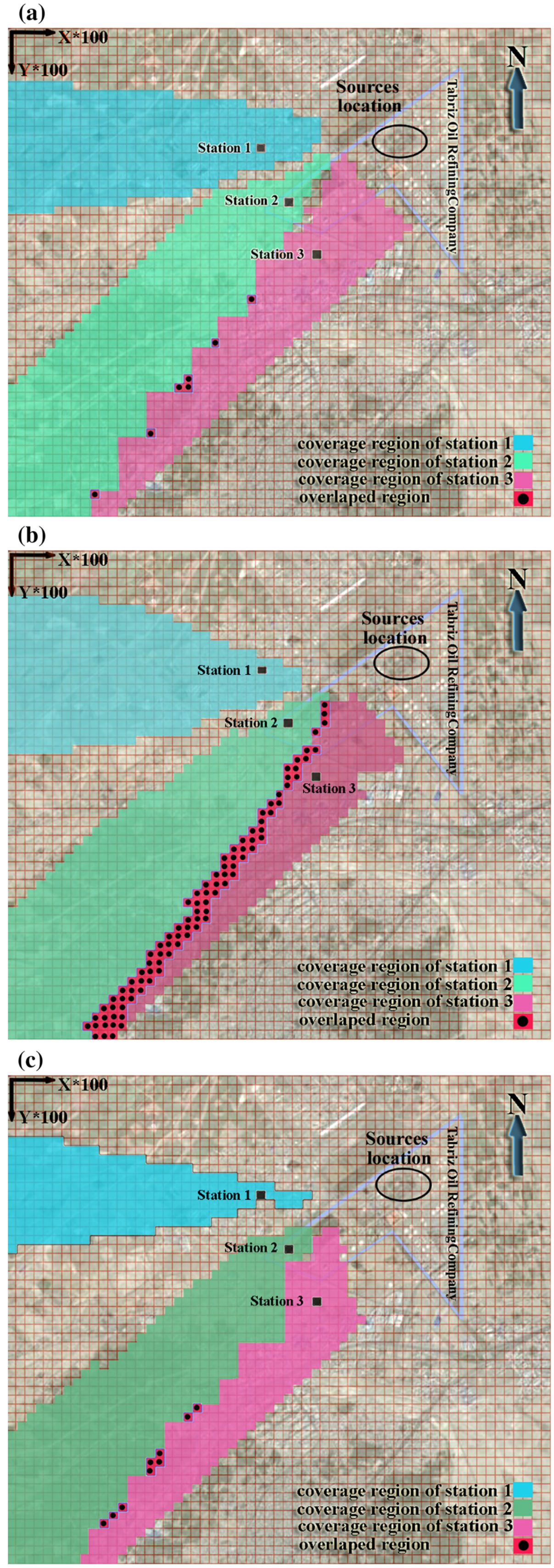


Fig. 9 Coverage area of AQMN for selected location of stations: ▶ **a** Pollutant NO_x, **b** Pollutant SO_2, and **c** pollutant CO

grids. In general, it can be concluded that increase of overall UF_i of the network was due to the coverage of identical points by several stations, or in the other words, increasing the number of overlapped grids. In conclusion, establishing a network with more than three stations would not be economically justified for the study area.

A comparison between the results of the current study and a previous study (Elkamel et al. 2008) was made in order to evaluate the performance of the proposed method. The results of the comparison indicated that the new method performed better than the previously reported method. The new method led to a sensitive network with more coverage area and high accuracy of monitoring while the number of proposed stations was the same in both studies. Higher coverage area with the same number of monitoring devices leads to more information and decreasing costs of equipment, maintenance and operating personnel. The reduction in costs is due to the fact that even by increasing (almost doubling) the coverage area, AQMN was designed in a way that requires the same number of stations. However, considering the doubling of the coverage area, if the previously reported method was used at least six stations would be required.

4.2 Sensitivity of monitoring stations to emission sources

Using values obtained at sensitive monitoring stations and having the knowledge of the sensitivity of the monitoring stations to emission sources will allow to quickly identify the defective sources, and then take the necessary measures to control them. Defective sources are the emission sources in which, due to the unexpected or stochastic changes, the amount of released pollutants has been higher than normal value.

For this purpose, a database of the sensitivity of the location of monitoring stations to emission sources was created and a program capable of ranking the defective sources by using the created sensitivity database and changes in measured concentration at monitoring station was designed to be used after setting up the stations. Three candidate locations (795, 1152 and 1505) with $r_c = 0.85$ were selected for the placement of the monitoring stations. Development of the database was done using the data which are used to calculate the sensitivity scores of selected optimal locations.

Figure 9 shows the coverage area of AQMN for selected location of stations for $r_c = 0.85$ and pollutants NO_x, SO_2 and CO.

5 Conclusions

Designing air quality monitoring networks is an important task for environmental protection authorities. It is necessary for AQMN to be planned effectively and systematically. An improved methodology is presented to optimize the number and location of air quality monitoring stations with respect to multi-pollutants and multi-objective model.

In this approach, the developed MCA model was employed to model the distribution of pollutants concentrations, and the results of dispersion model were used as inputs for the optimization algorithm. The generated monitoring network provides the optimal number and location of monitoring stations with respect to maximum coverage area with the minimum overlap among coverage areas as well as the maximum detection of violations over ambient standards and the maximum sensitivity of monitoring stations to emission sources for primary gaseous pollutants such as SO_2, NO_x and CO. Ant colony optimization algorithm and Genetic algorithm were used in the optimization procedure. This approach was applied to the case of Tabriz Oil refining Company.

The results of the optimization procedure were investigated for different values of the correlation coefficient (r_c) which showed that as the cut-off correlation coefficient was increased, the coverage area of stations and the overlap region were decreased.

Comparison of the performance of the ACO and GA indicated that ACO has better ability for reduction of overlap among coverage areas of monitoring station than GA. However both algorithms have good ability in finding the best configuration of AQMN and maximizing the coverage area of monitoring stations.

A comparison between the results of the current study and a previous study (Elkamel et al. 2008) indicated that the new method provides more coverage area. This increase in coverage efficiency will be remarkable when the increase of network coverage area (due to considering the prevailing wind directions) and accuracy of monitoring (due to considering the continuity of coverage area) be considered. Finally, three sensitive stations were suggested for AQMN of study area which enables the authorities to specify defective sources and then take prompt measures to control them.

Acknowledgments This project was funded by the Tabriz Oil Refining Company. The authors would like to thank the company Research and Development Department and HSE Department for providing facilities and successful cooperation.

References

Abdul-Wahab S, Ali S, Sardar S, Irfan N, Al-Damkhi A (2011) Evaluating the performance of an integrated CALPUFF-MM5 modeling system for predicting SO_2 emission from a refinery. Clean Technol Environ Policy 13:841–854. doi:10.1007/s10098-011-0360-6

Arbeloa FJS, Caseiras CP, Andrés PML (1993) Air quality monitoring: optimization of a network around a hypothetical potash plant in open countryside. Atmos Environ Part A Gen Top 27:729–738

Baldauf RW, Lane DD, Marote GA (2001) Ambient air quality monitoring network design for assessing human health impacts from exposures to airborne contaminants. Environ Monit Assess 66:63–76

Baldauf R, Lane D, Marotz G, Barkman HW, Pierce T (2002) Application of a risk assessment based approach to designing ambient air quality monitoring networks for evaluating non-cancer health impacts. Environ Monit Assess 78:213–227. doi:10.1023/A:1019920302108

Bayraktar H, Turalioglu FS (2005) A Kriging-based approach for locating a sampling site—in the assessment of air quality. Stoch Environ Res Risk Assess 19:301–305. doi:10.1007/s00477-005-0234-8

Chang N-B, Tseng CC (1999a) Optimal design of a multi-pollutant air quality monitoring network in a metropolitan region using kaohsiung, Taiwan as an example. Environ Monit Assess 57:121–148. doi:10.1023/A:1005992712569

Chang NB, Tseng CC (1999b) Optimal evaluation of expansion alternatives for existing air quality monitoring network by grey compromise programing. J Environ Manag 56:61–77

Chen C-H, Liu W-L, Chen C-H (2006) Development of a multiple objective planning theory and system for sustainable air quality monitoring networks. Sci Total Environ 354:1–19

Corti A, Senatore A (2000) Project of an air quality monitoring network for industrial site in Italy. Environ Monit Assess 65:109–117

Elkamel A, Fatehifar E, Taheri M, Al-Rashidi MS, Lohi A (2008) A heuristic optimization approach for air quality monitoring network design with the simultaneous consideration of multiple pollutants. J Environ Manag 88:507–516

Fatehifar E, Elkamel A, Taheri M (2006) A MATLAB-based modeling and simulation program for dispersion of multipollutants from an industrial stack for educational use in a course on air pollution control. Comput Appl Eng Educ 14:300–312. doi:10.1002/cae.20089

Fatehifar E, Elkamel A, Taheri M, Anderson W, Abdul-Wahab S (2007) Modeling and simulation of multipollutant dispersion from a network of refinery stacks using a multiple cell approach. Environ Eng Sci 24:795–811

Fatehifar E, Elkamel A, Osalu AA, Charchi A (2008) Developing a new model for simulation of pollution dispersion from a network of stacks. Appl Math Comput 206:662–668

Fayaed S, El-Shafie A, Jaafar O (2013) Reservoir-system simulation and optimization techniques. Stoch Environ Res Risk Assess 27:1751–1772. doi:10.1007/s00477-013-0711-4

Ferradás EG, Miñarro MD, Terrés IMM, Martínez FJM (2010) An approach for determining air pollution monitoring sites. Atmos Environ 44:2640–2645. doi:10.1016/j.atmosenv.2010.03.044

Finzi G, Zani L, Zecchini M (1991) A decision support system for the design of air quality monitoring networks. Environ Softw 6:55–60

Gerharz L, Pebesma E (2013) Using geostatistical simulation to disaggregate air quality model results for individual exposure estimation on GPS tracks. Stoch Environ Res Risk Assess 27:223–234. doi:10.1007/s00477-012-0578-9

Gómez-Losada Á, Lozano-García A, Pino-Mejías R, Contreras-González J (2014) Finite mixture models to characterize and refine air quality monitoring networks. Sci Total Environ 485–486:292–299. doi:10.1016/j.scitotenv.2014.03.091

Gurjar BR, Molina LT, Ojha CSP (2010) Air pollution: health and environmental impacts. Taylor & Francis, New York

Harris P, Clarke A, Juggins S, Brunsdon C, Charlton M (2014) Geographically weighted methods and their use in network re-designs for environmental monitoring. Stoch Environ Res Risk Assess 28:1869–1887. doi:10.1007/s00477-014-0851-1
Hsu A, Reuben A, Shindell D, de Sherbinin A, Levy M (2013) Toward the next generation of air quality monitoring indicators. Atmos Environ 80:561–570. doi:10.1016/j.atmosenv.2013.07.036
Kanaroglou PS, Jerrett M, Morrison J, Beckerman B, Arain MA, Gilbert NL, Brook JR (2005) Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. Atmos Environ 39:2399–2409. doi:10.1016/j.atmosenv.2004.06.049
Kao J-J, Hsieh M-R (2006) Utilizing multiobjective analysis to determine an air quality monitoring network in an industrial district. Atmos Environ 40:1092–1103
Kuhlbusch TAJ et al (2014) New directions: the future of European urban air quality monitoring. Atmos Environ 87:258–260. doi:10.1016/j.atmosenv.2014.01.012
Liu MK, Avrin J, Pollack RI, Behar JV, McElroy JL (1986) Methodology for designing air quality monitoring networks: I. Theoretical aspects. Environ Monit Assess 6:1–11
Liu D, Guo S, Shao Q, Jiang Y, Chen X (2014a) Optimal allocation of water quantity and waste load in the Northwest Pearl River Delta, China. Stoch Environ Res Risk Assess 28:1525–1542. doi:10.1007/s00477-013-0829-4
Liu F, Wen Z, Zheng L (2014b) Mathematical modeling for air quality management under stochastic and interval uncertainties. Stoch Environ Res Risk Assess, pp 1–14. doi:10.1007/s00477-014-0934-z
Lozano A, Usero J, Vanderlinden E, Raez J, Contreras J, Navarrete B (2009a) Air quality monitoring network design to control nitrogen dioxide and ozone, applied in Malaga, Spain. Microchem J 93:164–172
Lozano A, Usero J, Vanderlinden E, Raez J, Contreras J, Navarrete B, El Bakouri H (2009b) Design of air quality monitoring networks and its application to NO_2 and O_3 in Cordova, Spain. Microchem J 93:211–219
Lozano A, Usero J, Vanderlinden E, Raez J, Contreras J, Navarrete B, El Bakouri 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
Maria Grazia P, Daniela R, Mario CC, Renato C (1999) An optimal design for air pollution monitoring network. Environmetrics 10:351–360
Mazzeo N, Venegas L (2008) Design of an air-quality surveillance system for buenos aires city integrated by a NO x monitoring network and atmospheric dispersion models. Environ Model Assess 13:349–356. doi:10.1007/s10666-007-9101-y
Modak P, Lohani BN (1985a) Optimization of ambient air quality monitoring networks (part I). Environ Monit Assess 5:1–19. doi:10.1007/bf00396391
Modak P, Lohani BN (1985b) Optimization of ambient air quality monitoring networks (part II). Environ Monit Assess 5:39–53. doi:10.1007/bf00396393
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
Mofarrah A, Husain T, Alharbi BH (2011) Design of urban air quality monitoring network: fuzzy based multi-criteria decision making approach. Air Qual Monit Assess Manag. doi:10.5772/16716
Nejadkoorki F, Baroutian S (2012) Forecasting extreme PM 10 concentrations using artificial neural networks. Int J Environ Res 6:277–284
Nejadkoorki F, Nicholson K, Hadad K (2011) The design of long-term air quality monitoring networks in urban areas using a spatiotemporal approach. Environ Monit Assess 172:215–223. doi:10.1007/s10661-010-1328-4
Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM, Martins FG (2008a) Management of air quality monitoring using principal component and cluster analysis–Part I: SO_2 and PM10. Atmos Environ 42:1249–1260
Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM, Martins FG (2008b) Management of air quality monitoring using principal component and cluster analysis–Part II: CO, NO_2 and O_3. Atmos Environ 42:1261–1274
Righini G, Cappelletti A, Ciucci A, Cremona G, Piersanti A, Vitali L, Ciancarella L (2014) GIS based assessment of the spatial representativeness of air quality monitoring stations using pollutant emissions data. Atmos Environ 97:121–129. doi:10.1016/j.atmosenv.2014.08.015
Sarigiannis D, Saisana M (2008) Multi-objective optimization of air quality monitoring. Environ Monit Assess 136:87–99. doi:10.1007/s10661-007-9725-z
Shad R, Mesgari MS, Abkar A, Shad A (2009) Predicting air pollution using fuzzy genetic linear membership kriging in GIS Computers. Environ Urban Syst 33:472–481. doi:10.1016/j.compenvurbsys.2009.10.004
Silva C, Quiroz A (2003) Optimization of the atmospheric pollution monitoring network at Santiago de Chile. Atmos Environ 37:2337–2345
Skardi ME, Afshar A, Saadatpour M, Solis SS (2015) Hybrid ACO–ANN-based multi-objective simulation–optimization model for pollutant load control at basin scale. Environ Model Assess 20:29–39. doi:10.1007/s10666-014-9413-7
Tseng CC, Chang N-B (2001) Assessing relocation strategies of urban air quality monitoring stations by GA-based compromise programming. Environ Int 26:523–541
Zarandi SM, Khajevandi M, Damez-Fontaine D, Ardestani M (2008) Determination of air pollution monitoring stations. Int J Environ Res 2:313–318
Zhang Y et al (2014) Air quality in Lanzhou, a major industrial City in China: characteristics of air pollution and review of existing evidence from air pollution and health studies. Water Air Soil Pollut 225:1–17. doi:10.1007/s11270-014-2187-3