

# Optimal design of air quality monitoring network around an oil refinery plant: a holistic approach

Kh. Zoroufchi Benis · E. Fatehifar

Received: 15 March 2014/Revised: 20 October 2014/Accepted: 17 November 2014/Published online: 6 January 2015  
© Islamic Azad University (IAU) 2015

**Abstract** In this study, a multi-objective method for allocating the number and configuration of an air quality monitoring network based on non-dominated sorting genetic algorithm II has been presented. The multiple cell approach based on the solution of an Eulerian Model built on K-theory was used to predict the dispersion of emitted pollutants ( $\text{SO}_2$ ,  $\text{CO}$ ,  $\text{NO}_x$ ) from different emission sources. The multi-objective optimization method proposed in this study utilized two objectives: (1) maximum coverage area with respect to continuity of covered area and minimum overlap among coverage areas and (2) detection of violations over ambient standards. The concept of sphere of influence was used to determine the spatial area coverage of the monitoring station, and a weighing function was employed to measure the capability of a designed network to detect violations of air quality standards. The results show that three stations are suitable for the study region with coverage efficiency of 80 %. Analyzing the effect of cutoff correlation coefficient  $r_c$  shows that, when the  $r_c$  increases, although the coverage area decreases, the covered region will be well represented and overlap region will decrease.

**Keywords** Air quality · Genetic algorithm · Multi-objective · Multi-pollutant · Optimization

## Introduction

Ambient air quality monitoring is important in air quality management and plays a key role in identifying the air quality problems and formulating policies to control emission sources of air pollutants. It is also important in many other aspects such as to clarify local/regional specific sources and sinks of air pollutants, to study the dynamic behavior of air pollutants, to verify dispersion models, to check compliance of statistical models, and to risk assessment of traffic-related air pollution (Sheng and Tang 2013; Venkanna et al. 2014). Therefore, air quality monitoring network (AQMN) provides a database which enables policy-makers to take informed decisions to protect public health and the environment.

The optimal design of AQMN is highly desirable, since in this case AQMN will be able to meet air quality objectives while reducing the overall number measurement devices which minimizes costs (Nejadkoorki et al. 2011).

Primary studies on monitoring site planning were done based on an ad-hoc fashion by installing pollutant sensors in hot spots but the majority of multi-objective and multi-pollutant approaches which were considered relevant parameters such as population density and pollution variability have been developed in recent researches (Nejadkoorki and Baroutian 2012). Generally, the objectives of AQMN can be summarized in the terms of spatial representativity (i.e., siting criteria, including fixed or mobile sites and numbers of sites), time resolution, and measurement accuracy (Kuhlbusch et al. 2013).

Modak and Lohani (1985a) performed the design of an AQMN, based on the objectives of maximum violation detection and coverage monitoring for single and multiple pollutants in Taipei City, Taiwan, in which the design principles of a Minimum Spanning Tree algorithm were

Kh. Zoroufchi Benis · E. Fatehifar (✉)  
Faculty of Chemical Engineering, Environmental Engineering  
Research Center, Sahand University of Technology, P.O. Box  
51335-1995, Tabriz, Iran  
e-mail: fatehifar@sut.ac.ir



illustrated. They proposed two approaches: the first one was based on the utility function and the other one was based on the principles of sequential interactive compromise. The comparison showed that utility function approach was more effective than the sequential approach in the case of Taipei City. Noll and Mitsutomi (1983) developed an approach that locates monitoring stations using ambient dosage as an index which ranks potential air monitoring sites according to their ability to represent the ambient dosage. The results of application of this method to a power plant located in Northern Illinois denoted that 15 stations are required to monitor the area around the power plant. Liu et al. (1986) applied the concepts of 'sphere of influence' (SOI) and 'figure of merit' (FOM) to determine the number and disposition of ambient air quality stations in a monitoring network. The proposed method also is useful for modification (through addition or relocation) of an existing network. However, in this research, the method had not been applied to any case study. Jain and Sharma (2002) proposed a simple and generalized method for designing an optimum AQMN based on entropy concepts, which are central to the information theory. This methodology was applied to the existing network of nine stations in Delhi. The results suggested different optimal networks for each pollutant. This method addressed the following issues for AQMN design of Delhi: (1) priority locations for sampling and (2) optimal size of network. Chang and Tseng (1999) presented a new approach based on grey compromise programming for siting of new air quality monitoring stations. Results of application of this method to the AQMN for the city of Kaohsiung in South Taiwan indicated that the grey compromise programming model is a useful tool in evaluating the expansion alternatives with respect to different types of decision-making processes. Tseng and Chang (2001) in a companion study of Chang and Tseng (1999) developed a GA-based compromise programming technique to assess relocation strategy of air quality monitoring stations. Experience gained in this study pointed out that as the number of pollutants and objectives increases simultaneously, the higher number of candidate location are selected in the relocation strategy. Mofarrah and Husain (2010) and Mofarrah et al. (2011) presented an objective methodology for determining the optimum number of ambient air quality stations in a monitoring network. They used the fuzzy analytic hierarchy process (FAHP) with triangular fuzzy numbers (TFNs) and concept of sphere of influences (SOIs). The expansion of AQMN of Riyadh city in Saudi Arabia was used as a case study. Based on the results of this study, ten optimal locations for monitoring stations were proposed. McElroy et al. (1986) presented an objective methodology that applies the concepts of SOI and FOM developed by Liu et al. (1986) for determining the

optimum number and disposition of ambient air quality stations in a monitoring network. The application of presented method for determining the optimum number and disposition of monitoring stations for CO<sub>2</sub> in the Las Vegas, Nevada area indicated that networks with fewer stations would be selected if smaller minimum detection capabilities of concentration variations are acceptable, and vice versa. Kao and Hsieh (2006) performed the design of an AQMN in an industrial district (Toufen Industrial District in Miaoli County, Taiwan) in which the objectives of pollution detection, dosage, coverage, and population protection were considered. This study compared the effects of various objectives on the selection of monitoring sites, with the intention to devise a suitable monitoring network and to demonstrate the applicability of the established model. Elkamel et al. (2008) proposed a heuristic optimization approach to determine the optimum number and location of air quality monitoring stations in which multiple cell approach (MCA) was used to simulate monthly distributions for the concentrations of the pollutants emitted from different emission sources. The approach was applied to a network of existing oil refinery stacks, and the results showed that three stations provide a total coverage of more than 70 %. Serón Arbeloa et al. (1993) presented a methodology based on a heuristic approach introduced by Liu et al. (1986) and on the ideas presented by Modak and Lohani (1985a, b) in which the objectives of prediction of the spatial and temporal patterns of the concentration field and detection of violations over legal standards were considered. The implementation of this method for optimization of a network around a hypothetical potash plant indicated that six stations are required for monitoring and control of the pollution. Nejadkoorki et al. (2011) proposed a cost-effective approach for designing a long-term air pollution monitoring network for an urban area. The methodology was applied to the existing monitoring stations of Yazd, Iran in different seasons. The results indicated that the existing monitoring network is not ideal and could benefit from being redesigned and the minimum number of required sites for fall, winter, spring, and summer were 28, 21, 9, and 10, respectively. Zheng et al. (2011) established a methodological framework for site location optimization in designing a regional AQMN, on the basis of analyzing various constraints such as cost and budget, terrain conditions, administrative district, population density, and spatial coverage. The results showed that the presented method can be used as a reference to guide site location optimization of a regional AQMN design in China or other regions of the world. Lozano et al. (2011) developed a method to design or adjust air quality networks for monitoring NO<sub>2</sub> and ozone in compliance with the legislation. The method was applied to the optimization of the air quality assessment network in

Granada. The results indicated that one traffic-orientated and one background control station were necessary for NO<sub>2</sub> assessment, as well as one control station for O<sub>3</sub>. Littidej et al. (2012) applied mathematical model and GIS to determine a proper zone of air quality monitoring stations to monitor CO and NO<sub>x</sub> concentration in a municipality area. Based on the results, it can be concluded that proper location of a monitoring station can be found effectively by using optimization multi-objective decision analysis.

In this research, a GA-based method is presented to represent multi-objective multi-pollutants AQMN design. The optimization algorithm is an extension of methodology presented by Modak and Lohani (1985a, b) which was used and extended by Serón Arbeloa et al. (1993). A mathematical model based on the MCA, delineated in Fatehifar et al. (2006), Kahforoshan et al. (2008) and Fatehifar et al. (2007), was developed to be used for simulating pollutant distributions and determining ground level concentration of the pollutants emitted from different emission sources. The MCA model is based on the solution of the three-dimensional diffusion equation to predict the dispersion of emitted pollutants from refinery stacks. The output results of MCA were used as the input of optimization algorithm. A MATLAB program was written based on the methodology discussed in this paper. The objective of this AQMN design is optimization of station locations with respect to: (1) maximum coverage area with respect to continuity of covered area and minimum overlap among coverage areas and (2) detection of violations over ambient standards. The continuity of coverage area also was taken into account to increase the accuracy of monitoring. The research was done in Tabriz Oil Refining Company, from 2010 to 2012.

**Materials and methods**

The strategy for design and operation of AQMNs are dependent on monitoring objectives. Typical monitoring objectives reported in the WHO (1999) include: (1) Population exposure and health impact assessment; (2) Identifying threats to natural ecosystems; (3) Determining compliance with national or international standards; (4) Informing the public about air quality and establishing alert systems; (5) Providing objective input to air quality management, and also transporting and land-use planning; (6) Identifying and apportioning sources; (7) Developing policies and setting priorities for management actions; (8) Developing and validating management tools such as models and geographical information systems; and (9) Quantifying trends to identify future problems or progress in achieving management or control targets.

A monitoring station is generally located in a place where it can measure the pollution distributions in the best

way. Figure 1 shows the pollutant concentration profile on ground level for a single point source. The potential zone has been defined as the zone with pollutant concentration of more than a definite threshold value. It has been suggested that a monitoring station should be located in a potential zone (Kao and Hsieh 2006).

In this paper, firstly, the dispersion model and meteorological uncertainty will be explained. Then, the description of the applied optimization model will be followed, which includes objective formulations and non-dominated sorting genetic algorithm II (NSGA-II) method.

**Dispersion model**

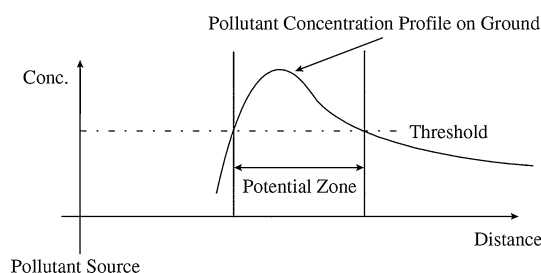
The MCA model, which has been developed and introduced in former papers (Fatehifar et al. 2006, 2007, 2008; Kahforoshan et al. 2008), is improved in a way that considers wind direction in the simulation of pollutants dispersion in order to make it more comfortable for the AQMN optimization. A brief description of main considerations and assumptions of the model is included below.

The MCA model is based on the solution of an Eulerian Model built on K-theory to predict the dispersion of emitted pollutants from point sources. The model uses graphical user interface (GUI-MATLAB program based), and it is applicable for network of refinery stacks, petrochemical complexes, and urban and industrial stacks. The basis of the model is mass conservation equation in which the transfer and diffusion of pollutants from point source (at Cartesian coordinates, constant wind velocity, and turbulent diffusivities) are described by advective–diffusive equation (Eq. 1) (Fatehifar et al. 2008):

$$\begin{aligned} \frac{\partial C^S}{\partial t} = & - \frac{\partial(U_x C^S)}{\partial x} - \frac{\partial(U_y C^S)}{\partial y} - \frac{\partial(U_z C^S)}{\partial z} \\ & + \frac{\partial}{\partial x} \left( K_x \frac{\partial C^S}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_y \frac{\partial C^S}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_z \frac{\partial C^S}{\partial z} \right) \\ & + E^S - (k_1^S + k_2^S) C^S + Q(C^S) \end{aligned}$$

$S = 1, 2, \dots, q$

(1)



**Fig. 1** Pollutant concentration profile on ground level (Kao and Hsieh 2006)

where  $C^S$  is the concentration of the chemical species involved in the model (CO, NO<sub>x</sub> and SO<sub>2</sub>),  $U$  is wind velocity,  $K_x$ ,  $K_y$ , and  $K_z$  are diffusion coefficients,  $E^S$  is the emission sources,  $k_1^S$  and  $k_2^S$  are deposition coefficients (for the dry deposition and the wet deposition, respectively) and  $Q(C^S)$  represents chemical reactions.

Some assumptions and approximations should be taken into account in order to solve the problem formulated by the Eulerian approach (Seinfeld and Pandis 2006). The following items are employed:

1. The initial conditions are arbitrarily set to zero. The initial conditions have been found to be important only for the initial period of modeling.
2. Transport by bulk in wind direction exceeds diffusion in that direction i.e., molecular diffusion is negligible compared with turbulent diffusion ( $K_x = 0$ ).
3. The wind velocity is constant, a function of  $z$  and only in  $x$  direction ( $U_y = U_z = 0$ ). The pollutants are advected downwind and are diffused vertically and laterally by the turbulent eddies in the atmosphere. The wind and diffusivity profiles vary with height above the ground and are dependent on the net heat flux to the air and the local roughness of the surface (Ragland and Dennis 1975).
4. There is no chemical reaction in the system ( $Q = 0$ ). This assumption is valid, given that only primary pollutants were considered in this study (Olcese and Toselli 2005).
5. There is no deposition in the system ( $k_1^S = k_2^S = 0$ ). This assumption is valid considering that study area is relatively flat, dry, and smooth with negligible grass and vegetation. Also, the modeling is not done on rainy days.

Dispersion parameters such as meteorological conditions and emission rate are approximately constant while the plume is in route from the source to the receptor, which leads to considering the steady state condition ( $\frac{\partial C^S}{\partial t} = 0$ ). Applying following boundary conditions leads to the finite difference representation as given in Eq. 2: At  $x = 0$ ,  $C(0, j, k) = 0$  (considering the specific location of refinery and the prevailing wind directions, the background concentration is zero with a good approximation).

At  $y = 0$  and  $y = W$ ,  $\frac{\partial C}{\partial y} = 0$  (Given that, the emission sources were located at the center of domain and also the large enough width of domain was considered in the modeling; this boundary condition is acceptable).

At  $z = 0$ ,  $\frac{\partial C}{\partial z} = 0$  (there is no mass transfer between the ground surface and atmosphere).

At  $z =$  mixing height,  $\frac{\partial C}{\partial z} = 0$  (the pollutant plume can progress upwards from the release height until the mixing height (Beychok 1994). The mixing height changes on a daily and seasonal basis (Ghermandi et al. 2014) which were considered in the model.  $W$  and mixing height are illustrated in Fig. 2.

$$\begin{aligned} & U_{xk} C_{i+1,j,k}^S \Delta y \Delta z + (K_y)_k \Delta x \Delta z (C_{i+1,j,k}^S - C_{i+1,j+1,k}^S) / \Delta y \\ & + (K_y)_k \Delta x \Delta z (C_{i+1,j,k}^S - C_{i+1,j-1,k}^S) / \Delta y \\ & + (K_z)_{k+1/2} \Delta x \Delta y (C_{i+1,j,k}^S - C_{i+1,j,k+1}^S) / \Delta z \\ & + (K_z)_{k-1/2} \Delta x \Delta y (C_{i+1,j,k}^S - C_{i+1,j,k-1}^S) / \Delta z \\ & = U_{xk} C_{i,j,k}^S \Delta y \Delta z + E_{i,j,k}^S \Delta y \Delta z \end{aligned} \quad (2)$$

where the empirical equations, which are dependent on the stability classes of atmosphere and functions of surface roughness length and friction velocity, were used to calculate eddy diffusivities in the lateral ( $K_y$ ) and vertical ( $K_z$ ) directions and get wind velocity as a function of height above ground ( $U_{xk}$ ). The modified Holland's equation was used for plume rise calculation (Fatehifar et al. 2006, 2007, 2008; Kahforoshan et al. 2008).

Figure 2 presents the schematic view of selected domains for modeling of pollutants dispersion. As can be seen, two domains for modeling procedure were introduced: main and modeling domains. The main domain is defined in a way which includes all the areas that pollution gets dispersed over it 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. Application of the MCA model to industrial stacks has been verified in former papers (Fatehifar et al. 2006; Kahforoshan et al. 2008). The flowchart of the proposed model is shown in Fig. 5.

#### Meteorological uncertainty

Wind velocity and direction as well as the air temperature in the study area are the meteorological conditions that impose significant effect on the distribution of pollutants. Figure 3a shows the wind rose diagram for the study area based on monthly average from 2002 to 2012 and highlights that the prevailing wind directions were from the east and northeast. The percentage of east, northeast, and west wind directions occurrence were 49.25, 42.42, and 8.33, respectively. Thus, only east and northeast wind directions were considered in this study. Figure 3b illustrates the evolution of temperature throughout the year 2012. As



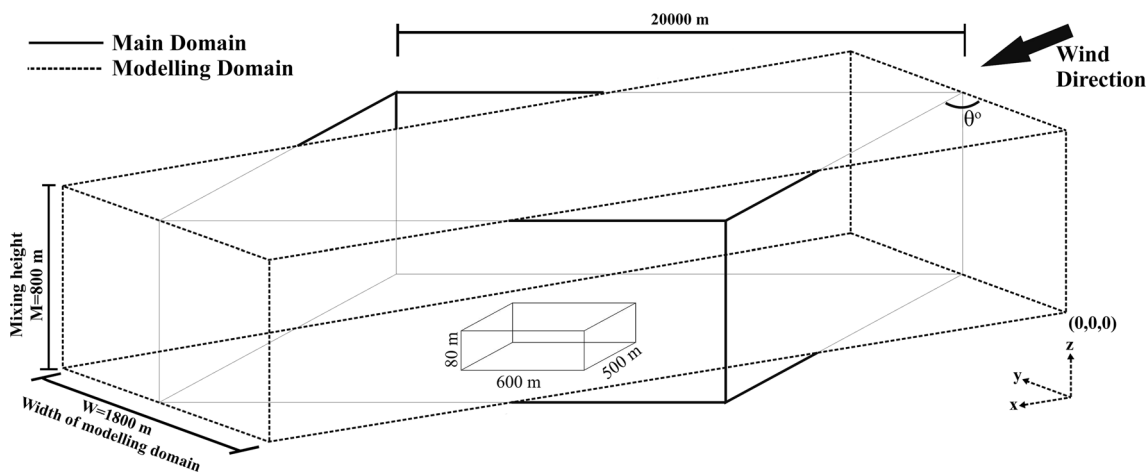


Fig. 2 Schematic of selected domains for simulation

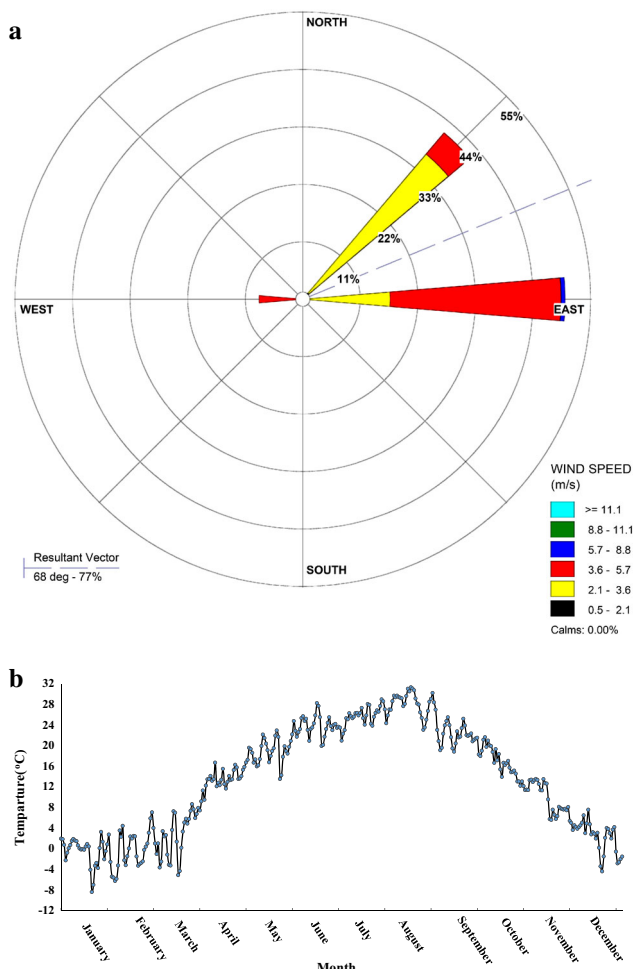


Fig. 3 Meteorological uncertainty of study area: **a** Wind rose for the study area based on monthly average from 2002 to 2012; **b** temperature variation versus time (year 2012) (Fatehifar 2012)

shown in this figure, the study area has relatively cold winters and hot summers.

Description of optimization objectives

Spatial coverage representation

The concept of SOI suggested by Liu et al. (1986) was employed to determine the special area coverage of the monitoring station. SOI is defined as the zone over which the air quality data for a given monitoring location can be considered representative (Elkamel et al. 2008). The spatial correlation coefficient ( $r$ ) was used to calculate the SOI. This coefficient is based on the similarity between values of pollutant concentration at a given potential location and the corresponding values at its neighboring points. For instance, if  $C_1 = (C_{11}, C_{12}, \dots, C_{1n})$  and  $C_2 = (C_{21}, C_{22}, \dots, C_{2n})$  denote the pollutant concentrations at two potential locations calculated by the MCA model at the same time, then the spatial correlation coefficient for a sample size  $n$  can be expressed as:

$$r = \frac{\sum_{i=1}^n (C_{1i} - \bar{C}_1) (C_{2i} - \bar{C}_2)}{\sqrt{\sum_{i=1}^n (C_{1i} - \bar{C}_1)^2 \sum_{i=1}^n (C_{2i} - \bar{C}_2)^2}} \tag{3}$$

and

$$\bar{C}_1 = \frac{1}{n} \sum_{i=1}^n C_{1i} \quad \text{and} \quad \bar{C}_2 = \frac{1}{n} \sum_{i=1}^n C_{2i} \tag{4}$$

$\bar{C}_1$  or  $\bar{C}_2$  is the average pollutant concentration calculated at point 1 or 2, respectively.

The spatial correlation coefficient ( $r$ ) for concentration fluctuations decreases by increasing the distance from the monitoring station. Therefore, a cutoff distance could be established in a way that the correlation coefficient in that distance is expected to be less than a certain cutoff value  $r_c$ . In order to determine SOI of each potential location, the calculated spatial correlation ( $r$ ) coefficient for each point was compared with a pre-defined cut off value ( $r_c$ ). If spatial correlation coefficient ( $r$ ) was higher than  $r_c$ , the corresponding points would be considered as correlated. The coverage area of a potential location is defined as the number of potential locations placed inside the SOI that can be quantified in terms of pattern scores. The pattern scores for an  $i$ th candidate location are denoted by  $N_p^i$ . This approach was also applied by Liu et al. (1986), Serón Arbeloa et al. (1993), Kao and Hsieh (2006) and Mo-farrah and Husain (2010) to determine coverage areas.

In this study, as shown in Fig. 4, continuity of SOI was considered to increase the accuracy of monitoring. Thus, the pattern score for an  $i$ th candidate location was redefined as the number of potential locations placed inside the SOI of candidate location  $i$  and is contiguous to each other.

Therefore, the first interest of AQMN designing is to determine the optimum number and location of monitoring stations, where the overall pattern score of the network is maximized.

#### Violations over ambient standards

The terms of violation scores were considered to measure the capability of a designed network to detect violations of

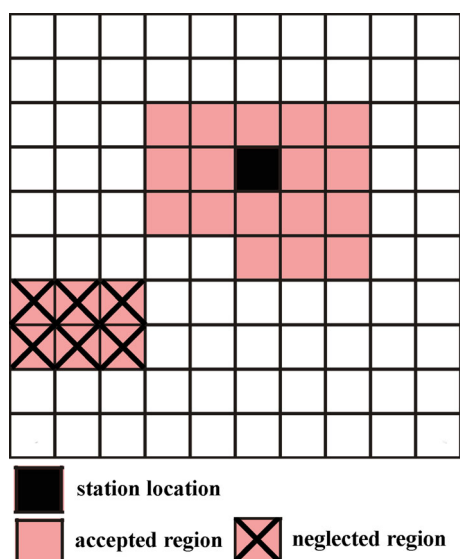


Fig. 4 Continuity of SOI

air quality standards which are denoted by  $N_v$ . A weighted scoring was used to determine the violation scores based on the concentration of  $\text{SO}_2$ ,  $\text{NO}_x$ , and  $\text{CO}$ . In this approach, all of the violations do not have the same severity. Therefore, these scores are dependent on the threshold levels and the weighing factors between each threshold range and the weighing function.

In this work, among the various weighing functions to calculate the violation score of a site, a segmented non-linear weighing function (Modak and Lohani 1985a, b) was chosen. The violation score for each candidate location is calculated by the equation below:

$$N_v^i = \sum_{i=1}^T \sum_{k=1}^{N_i} \frac{(w_{k+1} - w_k)(x_i - x_k)X}{(x_{k+1} - x_k)} \quad (5)$$

where  $N_v^i$  is the violation score for the  $i$ th candidate location,  $w_k$  is the weighing factor corresponding to threshold  $x_k$ ,  $x_k$  is the  $k$ th threshold,  $X = 0$  if  $(x_i - x_k) \leq 0$ ,  $X = 1$  otherwise,  $N_i$  is the total number of thresholds, and  $T$  is the total number of simulated observations. The threshold values based on the US-EPA limit values of common pollutants are shown in Table 1.

The second interest of optimization problem 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.

#### Fitness function

In order to meet the two objectives described above, a composite objective function is formulated as

$$F = N_p \times (N_v)^b \quad (\text{with } b \geq 0) \quad (6)$$

The value of parameter  $b$  depends on the purpose of the network to be developed and is used to weigh the relative importance given to each objective. This structure of the objective function ensures that if any of the two sub-objectives becomes zero, the overall objective function ( $F$ ) will also have a zero value.

A weighted sum objective function was used to extend the function to multiple pollutants:

**Table 1** Air pollution index assigned to weighing factors values for the violation (Elkamel et al. 2008)

$\text{SO}_2$ ( $\mu\text{g}/\text{m}^3$ )	$\text{NO}_x$ ( $\mu\text{g}/\text{m}^3$ )	$\text{CO}$ ( $\mu\text{g}/\text{m}^3$ )	Weighing factors
80	30	4,000	0.5
120	80	6,000	1
140	100	8,000	2.5
160	130	13,000	3
190	160	20,000	5



$$UF_i(u) = \sum_{j=1}^p w_j F_j^i \tag{7}$$

and

$$\sum_{j=1}^p w_j = 1 \tag{8}$$

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  $i$ .

Since the interest of the optimization problem is to achieve maximum coverage effectiveness and overall violation score at a minimum overlap, two dimensionless parameters were defined as follows:

$$UF_n = \frac{\max(UF) \times ns}{\sum_{i=1}^{ns} UF_i(u)} \tag{9}$$

$$OL_n = \frac{\text{number of overlapped grids}}{\text{maximum number of possible overlapped grids}} \tag{10}$$

where  $UF_n$  is the utility function number of the network,  $OL_n$  is overlap number,  $ns$  is the number of stations, and  $\max(UF)$  is the maximum value of  $UF(u)$  among all  $UF(u)$ s of the potential locations.

Therefore, the optimization problem is transformed into the identification of the location of stations in which  $UF_n$  and  $OL_n$  are in minimum amounts. The NSGA-II was used to find the optimal solution for optimization problem.

#### Non-dominated sorting genetic algorithm (NSGA-II)

Non-dominated sorting genetic algorithm II which is proposed by Deb et al. (2002) is the most famous multi-objective optimization algorithm, which is widely used for generating the Pareto frontier (that is a set of solutions which would represent the best trade-off among the objectives) and satisfying both goals of Pareto multi-objective optimization (Tao et al. 2014; Erfani et al. 2013). Fast non-dominated sorting procedure, fast crowded distance estimation procedure, and simple crowded comparison operator are the special characteristics of NSGA-II (Deb et al. 2002).

In sum, the algorithm can be outlined as follows:

Initial generation:

1. Generating random parent population  $P_0$  (Each chromosome of population includes a set of stations)

2. Finding cost of each chromosome (Cost is calculated based on the  $UF_n$  and  $OL_n$ )
3. Non-dominated sorting of initial population and calculate crowding distance
4. Creating an offspring population  $Q_0$  of size  $N$  by using the usual binary tournament selection, recombination and mutation operators

The  $i$ th generation:

5. Combining parent and offspring ( $Q_t$ ) to form  $R_t$  ( $R_t = P_t \cup Q_t$ )
6. Sorting the population  $R_t$  according to non-domination (each solution is assigned a fitness (or rank) equal to its non-domination level:  $F_1$  is the best level,  $F_2$  is the next best level, and so on)
7. Forming the population  $P_{t+1}$  (If the size of  $F_1$  is smaller than  $N$ , all the members of  $F_1$  should be placed in  $P_{t+1}$  and the remain members of  $P_{t+1}$  from subsequent non-dominated fronts are placed in the order of their ranking until the parent population is filled)
8. Using the population of  $P_{t+1}$  to create an offspring population  $Q_{t+1}$  by using selection, crossover and mutation
9. Checking the stopping criteria (If the end condition is satisfied, stop and show the current population as the best solution, otherwise go to step 5) (Deb et al. 2002). Crowding distance uses in selection operator to keeps the population diverse and helps the algorithm to explore the fitness landscape by making sure each member stays a crowding distance apart (Carlos A. Coello Coello et al. 2007). Details of the method can be found in Deb et al. (2002).

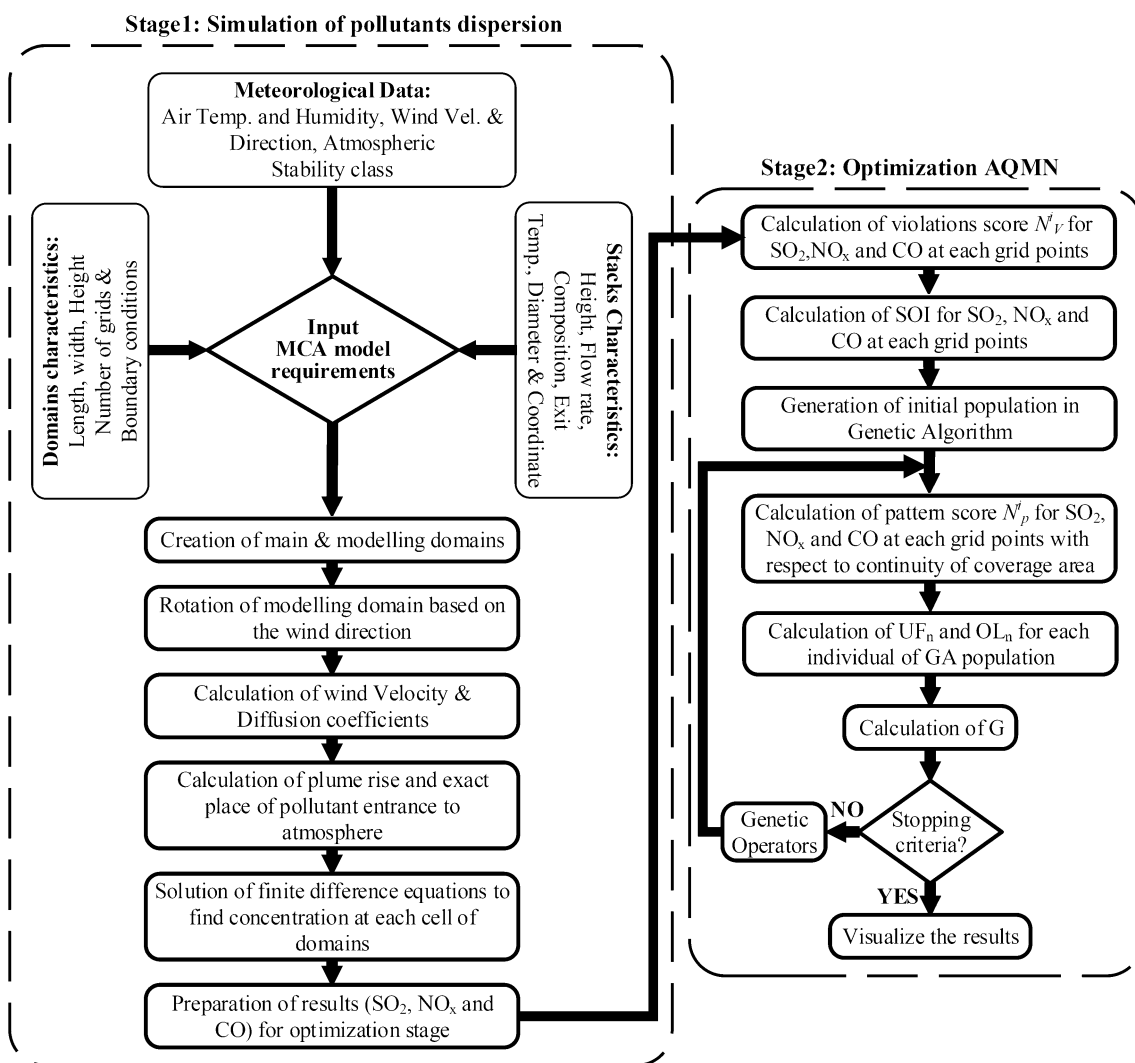
Overall, the AQMN designing procedure can be summarized by a flowchart as shown in Fig. 5. This algorithm has been implemented in MATLAB.

In the following section, the multi-objective model with varied weight sets is applied in a case study for planning an AQMN for a network around the oil refinery plant.

## Results and discussion

### Illustrative case

The multi-objective and multi-pollutant optimization method outlined above was applied around Tabriz Oil Refining Company. The refinery is situated on a plot area of 1.5 square kilometers located at 15 km of Tabriz-Azarshahr road, East Azarbayjan province, Iran. The refinery was originally built to process 80,000 BPD of



**Fig. 5** Flowchart of proposed method for optimal allocation of monitoring stations

Ahvaz-Asmari's crude oil to meet the region requirements. However, the capacity was enhanced to 110,000 BPD in recent years. The refinery complex, as a whole, contains refining units, utility services, waste water treatment plant, and storage tanks. In sum, CO, NO<sub>x</sub>, and SO<sub>2</sub> release from 20 stacks and one flare of refinery. In order to predict pollutant concentration at potential locations, 96 sets of collected data of pollutants concentrations which have been measured in the stacks and also wind velocities and temperatures of vicinity of the refinery stacks for 8 years (from 2005 to 2012) were used in MCA model. A matrix of 96 × 3,245 for SO<sub>2</sub>, NO<sub>x</sub>, and CO concentrations that has been generated at the specified 3,245 candidate locations was used as an input to the optimization algorithm. Figure 6a shows Tabriz Oil Refining Company and the locations of air pollution sources. Location 1, 2, and 3 include 19 stacks, one stack, and one flare, respectively. It is noted

that Tabriz Oil Refining Company is located at a smooth plate. Figure 6b–d shows the concentration of pollutant at ground level for selected main domain and different wind blowing angels in different conditions.

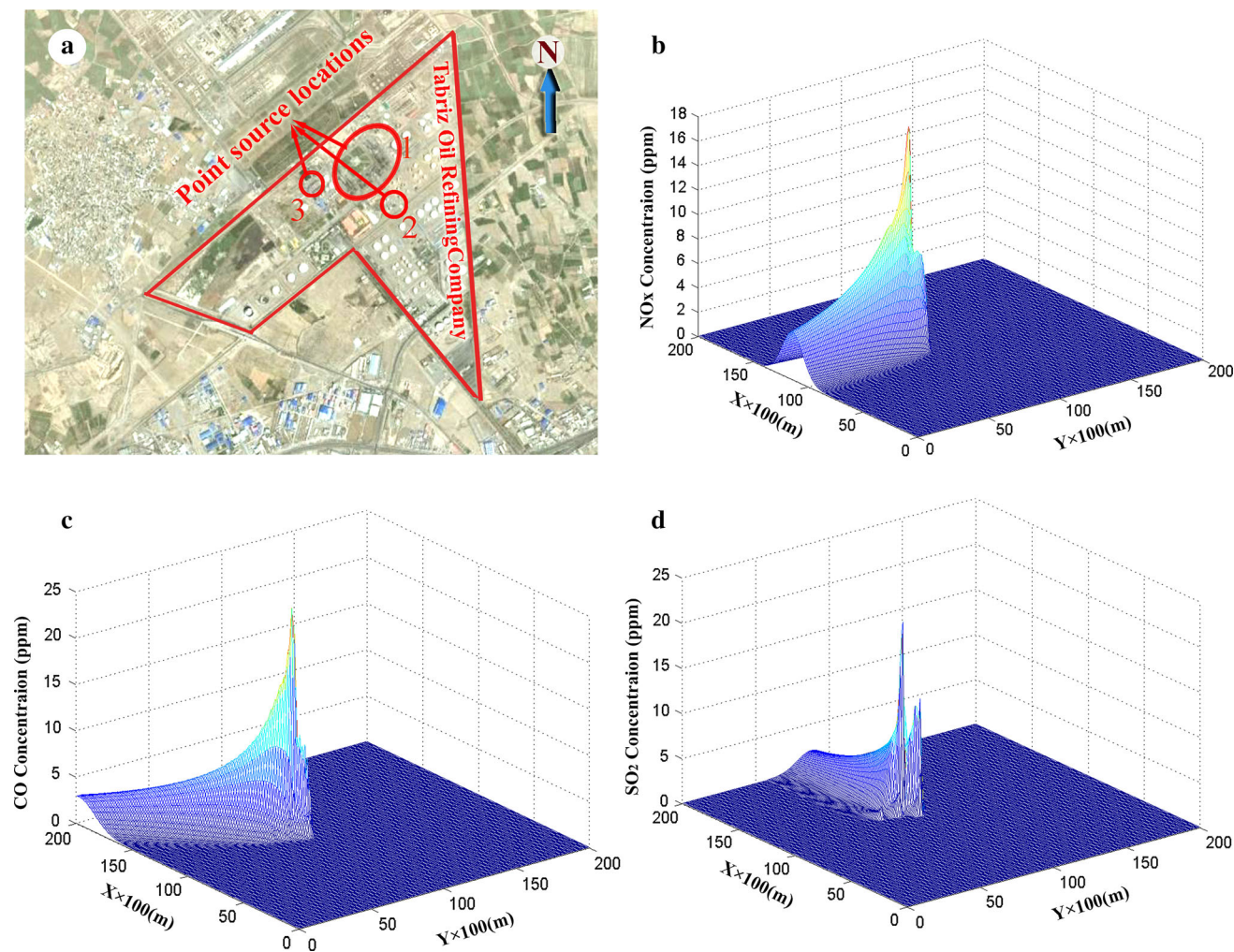
According to the Fig. 6, potential zone can be diagnosed from pollutant concentration profile on ground level of simulation examples.

#### Optimization results

The heuristic optimization algorithm was implemented for different cutoff  $r_c$  values in the correlation coefficient matrix. The value of  $r_c$  varied from 0.75 to 0.95 in order to study the effect of  $r_c$  on the coverage effectiveness of the monitoring networks. The outputs of optimization program are Pareto set of solution. The solution with maximum number of covered grids was selected as the best solution.







**Fig. 6** Location of industrial district and air pollution sources (a) and concentration of pollutants at ground level with wind blowing angles: b 10, c 45, d 90

Table 2 shows the optimal locations of the monitors for a maximum number of six stations.

Figure 7 illustrates the coverage efficiency of the AQMN for different  $r_c$  values. As shown in the figure, three stations for AQMN are suitable because by increasing the number of stations from 3 to 6, there is no significant increase in coverage efficiency. A concurrent comparison of  $UF_i$  of network and number of overlapped grids indicates that by increasing the number of stations (from 4 to 6), the increasing rate of overall  $UF_i$  of network decreases and in contrast the number of overlapped grids increases sharply. In general, it can be concluded that increasing of overall  $UF_i$  of network is due to covering 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. However, the final decision in such a case is of course left to the air quality monitoring organization.

For illustration purposes, three stations locations were selected and, Fig. 8 shows the location and coverage area of stations and overlap region for pollutant  $NO_x$ . As shown in the figures, when the cutoff value increases, the coverage region decreases, but the covered region will be well represented and overlap region will decrease.

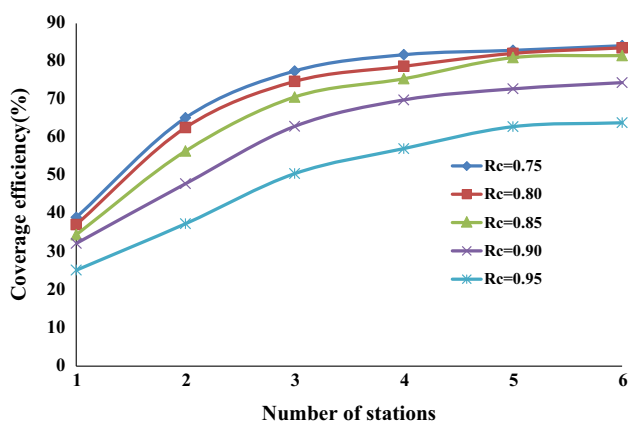
For a stipulated budget, the air quality monitoring organization (i.e., Iran EPA, US-EPA) could maintain either a high or a low value  $r_c$  based network. A high  $r_c$  based network may not necessarily cover the entire region, but the covered region will be well represented. A low  $r_c$  based network on the other hand would offer more coverage of the region, but the covered region may not be satisfactorily represented (Elkamel et al. 2008).

A comparison between the results of current study and previous paper (Elkamel et al. 2008) was done to evaluate the performance of proposed method. In comparison with previous study, an evolutionary algorithm (NSGA-II) was

**Table 2** Optimal locations for AQMN for multi-pollutant ( $x$  and  $y$  are in m)

Station number	$r_c = 0.75$		$r_c = 0.8$		$r_c = 0.85$		$r_c = 0.90$		$r_c = 0.95$	
	$x$	$y$	$x$	$x$	$x$	$y$	$y$	$y$	$x$	$y$
<b>1 Station</b>										
1	2,100	3,000	2,100	3,000	2,100	3,000	2,100	3,000	2,300	2,600
<b>2 Station</b>										
1	1,800	3,300	2,600	2,700	2,300	2,900	2,100	3,000	2,300	2,600
2	1,500	3,000	1,300	1,600	2,000	700	1,500	2,700	1,500	2,700
<b>3 Station</b>										
1	1,900	3,200	1,400	2,800	1,800	3,200	3,900	2,600	<b>2,600</b>	<b>2,300</b>
2	1,400	3,100	2,600	3,400	1,400	2,900	1,500	2,700	<b>1,900</b>	<b>1,600</b>
3	2,200	3,400	2,000	3,100	3,800	2,400	2,000	3,100	<b>2,600</b>	<b>3,400</b>
<b>4 Station</b>										
1	1,900	3,200	2,600	3,400	3,400	2,600	3,400	2,600	2,600	3,400
2	700	500	2,600	2,300	2,300	2,600	2,300	2,600	2,600	2,300
3	2,700	3,800	1,000	1,700	1,000	1,700	1,000	1,700	1,000	1,700
4	1,500	3,200	1,900	1,600	1,900	1,600	1,900	1,600	1,900	1,600
<b>5 Station</b>										
1	1,900	1,000	4,500	3,500	2,000	3,100	3,800	2,500	1,900	3,000
2	4,400	2,000	2,000	3,100	900	1,500	2,300	3,200	1,500	2,400
3	1,500	2,700	1,400	2,900	2,100	900	1,500	2,600	600	3,600
4	4,400	300	700	400	1,600	3,000	2,300	2,700	3,300	2,500
5	1,800	3,400	3,600	2,500	4,100	2,500	2,200	3,800	3,100	3,400
<b>6 Station</b>										
1	1,400	2,500	4,700	4,900	2,000	3,100	3,300	2,300	1,400	2,500
2	2,400	2,900	2,900	3,100	900	1,500	3,100	2,900	2,400	2,900
3	2,600	3,200	4,700	600	2,100	900	4,700	300	2,600	3,200
4	800	1,300	2,000	1,500	1,600	3,000	1,400	2,900	800	1,300
5	2,500	2,400	1,800	3,400	4,100	2,500	900	2,200	2,500	2,400
6	1,700	2,700	1,500	3,100	2,900	3,100	1,900	3,100	1,700	2,700

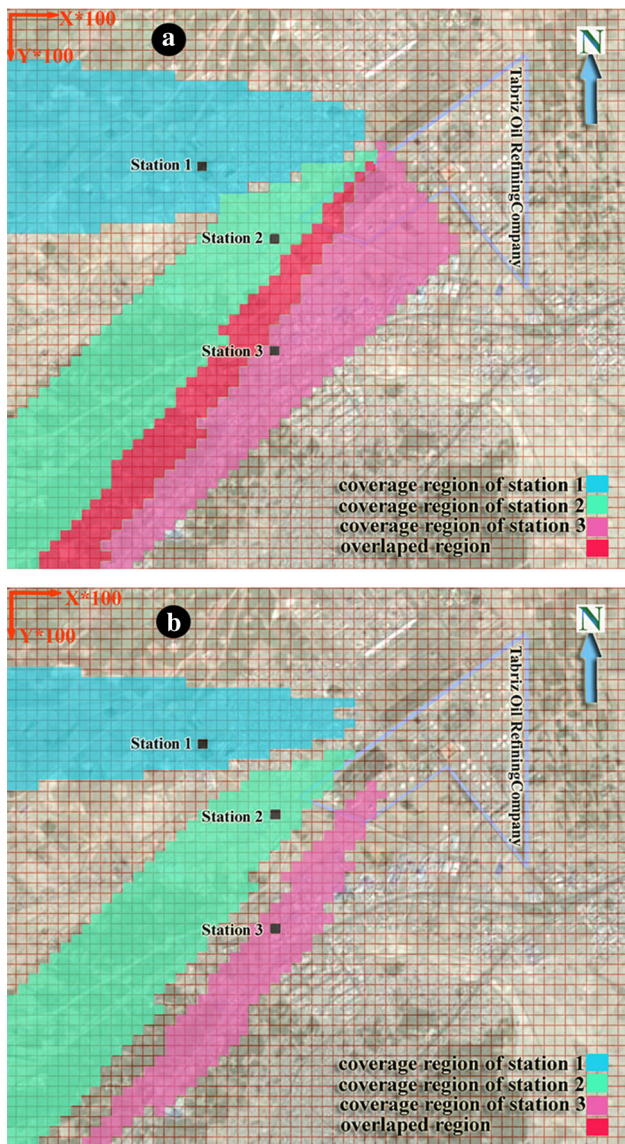
Bold values are the selected locations for monitoring stations



**Fig. 7** Coverage efficiency versus number of stations for multi-pollutants as a function of  $r_c$

used instead of sequential minimum Spanning Tree algorithm. Using NSGA-II allowed consideration of two objectives simultaneously: (1) maximization of objective function (2) minimization of overlap among coverage areas, which causes to maximize network performance and reduces the number of stations. Also, in the current study, the continuity of coverage area of stations was considered which leads to enhance the quality of monitoring. The MCA model which was used in the current study was developed in a way which can consider wind blowing angle and easily be used at optimization procedure. Table 3 shows some general characteristics of both studies.

As shown in the table, new method provides more coverage efficiency in comparison with the previous study. This increase in coverage efficiency will be remarkable



**Fig. 8** Location and coverage region for three selected stations: **a**  $r_c = 0.75$ , **b**  $r_c = 0.95$

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) being considered. In the other words, the coverage area of designed network in current study is almost double the coverage area of designed network of former paper while both networks have three stations which leads to low cost and more information.

**Conclusion**

Air quality monitoring network is an essential tool for air quality management and control sources of air pollutants. This study presented an approach for determining the

**Table 3** Some general characteristics and results of both studies

	Current study	Our previous study (Elkamel et al. 2008)
Model of air pollution dispersion	MCA	MCA
Wind direction	Considered	Not considered
Wind velocity	Considered	Considered
Air temperature	Considered	Considered
Used data set	96 set	72 set
Optimization method	NSGA-II	Minimum Spanning Tree algorithm
Continuity of coverage area	Considered	Not considered
Overlap among coverage areas	Minimized	Not considered systemically
Number of proposed stations	3 stations	3 stations
Coverage efficiency of proposed network ( $r_c = 0.75$ )	80 %	75 %

optimal configuration of AQMN. This approach optimizes the location and number of monitoring stations with respect to maximum coverage area with minimum overlap and detection of violation over ambient standards for primary gaseous pollutants such as  $SO_2$ ,  $CO$ , and  $NO_x$ . In order to increase the accuracy of monitoring, the continuity of coverage area of monitoring stations was considered in the optimization procedure.

In this study, a mathematical model based on MCA was developed in a way that considers wind direction in the prediction of ground level pollutant concentrations in order to make it more comfortable for the AQMN optimization. The model was used to determine ground level concentration of pollutant in the study area (Tabriz Oil Refining Company) for 96 different scenarios. The output results of MCA were employed to determine the pattern score and violation score of the potential locations. An algorithm was formulated to find optimal number and location of monitoring stations using NSGA-II.

The effect of the correlation coefficient on total coverage and effectiveness of the network was also studied. The results showed that for the area under study, three stations are suitable and for  $r_c = 0.75$  can give a coverage efficiency of about 80 %. Evaluation of performance of proposed method indicated that the method has good efficiency in designing of AQMN.

The proposed design for the network around the oil refinery plant provides a cost-effective solution to environmental monitoring. The presented method is a suitable and effective method of designing a proper AQMN around an oil refinery which can be used for other industrial process plants such as petrochemical complexes and power plants.

**Acknowledgments** This project was funded by the Tabriz Oil Refining Company with the Grant Number of financial support of 88-21/Date: 10/13/2009. The authors would like to thank the company Research and Development Department for their kindly cooperation and for providing the case study.

## References

- Beychok MR (1994) Fundamentals of stack gas dispersion. Milton R. Beychok Irvine, Irvine
- Chang NB, Tseng CC (1999) Optimal evaluation of expansion alternatives for existing air quality monitoring network by grey compromise programming. *J Environ Manag* 56(1):61–77
- Coello CAC, Lamont GB, Veldhuizen DAV (2007) Evolutionary algorithms for solving multi-objective problems. Genetic and evolutionary computation series, 2nd edn. Springer. doi:10.1007/978-0-387-36797-2
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evol Comput IEEE Trans* 6(2):182–197. doi:10.1109/4235.996017
- 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(3):507–516
- Erfani T, Utyuzhnikov S, Kolo B (2013) A modified directed search domain algorithm for multiobjective engineering and design optimization. *Struct Multidiscip Optim* 48(6):1129–1141. doi:10.1007/s00158-013-0946-1
- Fatehifar E (2012) Design of optimal number and location of air quality monitoring networks (sensitive to emission sources) in Tabriz oil refinery. Technical report. Tabriz
- 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(4):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(6):795–811
- Fatehifar E, Elkamel A, Alizadeh Osalu A, Charchi A (2008) Developing a new model for simulation of pollution dispersion from a network of stacks. *Appl Math Comput* 206(2):662–668
- Ghermandi G, Teggi S, Fabbi S, Bigi A, Zaccanti MM (2014) Tri-generation power plant and conventional boilers: pollutant flow rate and atmospheric impact of stack emissions. *Int J Environ Sci Technol* 1–12. doi:10.1007/s13762-013-0463-1
- Jain VK, Sharma M (2002) Air quality monitoring network design using information theory. *J Environ Syst* 29(3):245–267
- Kahforoshan D, Fatehifar E, Babalou AA, Ebrahimi AR, Elkamel A, Soltanmohammadzadeh JS (2008) Modeling and evaluation of air pollution from a gaseous flare in an oil and gas processing area. Paper presented at the WSEAS conferences, Cantabria, Spain
- Kao J-J, Hsieh M-R (2006) Utilizing multiobjective analysis to determine an air quality monitoring network in an industrial district. *Atmos Environ* 40(6):1092–1103
- Kuhlbusch TJ, Quass U, Fuller G, Viana M, Querol X, Katsouyanni K, Quincey P (2013) Air pollution monitoring strategies and technologies for urban areas. In: Viana M (ed) *Urban air quality in Europe*, vol 26. The handbook of environmental chemistry. Springer, Berlin, pp 277–296. doi:10.1007/698\_2012\_213
- Littidej P, Sarapirome S, Aunphoklang W (2012) Air pollution concentration approach to potential area selection of the air quality monitoring station in Nakhon Ratchasima Municipality, Thailand. *J Environ Sci Eng A* 1(4):484–494
- 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):1–11
- Lozano A, Usero J, Vanderlinden E, Raez J, Contreras J, Navarrete B, El Bakouri H (2011) Air quality monitoring network design to control nitrogen dioxide and ozone, applied in Granada, Spain. *Ozone Sci Eng* 33(1):80–89. doi:10.1080/01919512.2011.536741
- McElroy JL, Behar JV, Meyers TC, Liu MK (1986) Methodology for designing air quality monitoring networks: II. Application to Las Vegas, Nevada, for carbon monoxide. *Environ Monit Assess* 6(1):13–34
- Modak P, Lohani BN (1985a) Optimization of ambient air quality monitoring networks (part I). *Environ Monit Assess* 5(1):1–19. doi:10.1007/bf00396391
- Modak P, Lohani BN (1985b) Optimization of ambient air quality monitoring networks (part II). *Environ Monit Assess* 5(1):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(3):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(1):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(1):215–223. doi:10.1007/s10661-010-1328-4
- Noll KE, Mitsutomi S (1983) Design methodology for optimum dosage air monitoring site selection. *Atmos Environ* (1967) 17(12):2583–2590
- Olcese LE, Toselli BM (2005) Development of a model for reactive emissions from industrial stacks. *Environ Model Softw* 20(10):1239–1250
- Ragland KW, Dennis RL (1975) Point source atmospheric diffusion model with variable wind and diffusivity profiles. *Atmos Environ* (1967) 9(2):175–189. doi:10.1016/0004-6981(75)90066-9
- Seinfeld JH, Pandis SN (2006) *Atmospheric chemistry and physics: from air pollution to climate change*. Wiley, Hoboken
- Serón Arbeloa FJ, Pérez Caseiras C, Latorre Andrés PM (1993) Air quality monitoring: optimization of a network around a hypothetical potash plant in open countryside. *Atmos Environ Part A Gen Top* 27(5):729–738
- Sheng N, Tang UW (2013) Risk assessment of traffic-related air pollution in a world heritage city. *Int J Environ Sci Technol* 10(1):11–18. doi:10.1007/s13762-012-0030-1
- Tao T, Wang J, Xin K, Li S (2014) Multi-objective optimal layout of distributed storm-water detention. *Int J Environ Sci Technol* 11(5):1473–1480. doi:10.1007/s13762-013-0330-0
- Tseng CC, Chang N-B (2001) Assessing relocation strategies of urban air quality monitoring stations by GA-based compromise programming. *Environ Int* 26(7–8):523–541
- Venkanna R, Nikhil GN, Siva Rao T, Sinha PR, Swamy YV (2014) Environmental monitoring of surface ozone and other trace gases over different time scales: chemistry, transport and modeling. *Int J Environ Sci Technol* 1–10. doi:10.1007/s13762-014-0537-8
- WHO (1999) *Monitoring ambient air quality for health impact assessment*, vol 85. World Health Organization regional publications, European series
- Zheng J, Feng X, Liu P, Zhong L, Lai S (2011) Site location optimization of regional air quality monitoring network in China: methodology and case study. *J Environ Monit* 13(11):3185–3195. doi:10.1039/C1EM10560D