Thai Journal of Mathematics
Volume 17 (2019) Number 1 : 125–140
http://thaijmath.in.cmu.ac.th
ISSN 1686-0209

Identification of Atmospheric Pollution Source Based on Particle Swarm Optimization

W. Chaiwino[†] and T. Mouktonglang[‡,§,1]

[†]Master Degree Program in Mathematics, Department of Mathematics
Faculty of Science, Chiang Mai University
Chiang Mai 50200, Thailand
e-mail : wipawineech@cmu.ac.th

[‡]Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd.
Bangkok 10400, Thailand

[§]Department of Mathematics, Faculty of Science, Chiang Mai University
Chiang Mai 50200, Thailand
e-mail : thanasak.m@cmu.ac.th

Abstract : In this paper, we develop a mathematical model for retrieving the positions and the rates of the emission of the pollutant sources. We begin by measuring the pollutant concentrations from pollutant sensors and then we apply particle swarm optimization (PSO) minimizing the differences between the theoretical and the measured concentration. The results of the test cases show that the mathematical model is capable of retrieving the positions and the rates of the emissions of the pollutant sources. We also design the placement of pollutant sensors for the pollutant detecting system to be able to retrieve the positions and the emission rates of the pollutant sources effectively.

Keywords : particle swarm optimization; atmospheric model.

2010 Mathematics Subject Classification : 47H09; 47H10.

[1]Corresponding author.

1 Introduction

An air pollution problem is one of the most significant problems in the high-technology era [1]. A lot of industrial developments could turn factories into the pollutant sources which release chemical substances into the atmosphere. By using results from scientific technique to detect location sources of chemical substances, the local authorities could use these informations to initiate an initial warning to the factories and/or begin further investigation.

Even in the developing country, the burning of rice and cone residues after harvest, in stead of alternatives such as burying the stubble or residues back into the ground, has a number of negative consequences and effects on the environment. The smoke from this particular burning could produce a toxic cloud of particulates diffusing into the city.

The aim of this paper is to propose an optimization technique based on an approximated solution of a diffusion equation to determines the emission rates and source locations accurately. It is a standard optimization application [2–4] which is quite important in a certain practical pollutant situation where the emission locations or rates cannot be measured directly at the sources, but rather only can be measured indirectly at a distance from the sources.

In the next section, the particle swarm optimization (PSO), the probabilistic optimization technique, is briefly discussed. The construction of the approximated solution of the atmospheric model is derived in the third section. The algorithm for detecting the emission rates and source locations are explained in the fourth section. Then, the results of numerical experiments are presented in the fifth section. We also propose where to place air pollution sensors around the city of Chiang Mai, Thailand, to monitor the air quality effectively. Finally, the last section is concluding remark.

2 Particle Swarm Optimization (PSO)

PSO is an evolutionary computation technique using individual improvement together with population competition, which is based on the simulation of simplified social models, such as bird flocking or fish schooling [5–7]. The particle swarm conception is originally motivated from the simulation of social behavior. PSO requires only basic mathematical operators, which are computationally inexpensive in terms of both memory requirements and time. The technique has a characteristic of fast convergence to local and global optimal position for within small number of iterations. A swarm in PSO consists of particles. Each particle represents a feasible or candidate solution to the optimization problem. Each particle moves to a new position according to the new velocity which includes its previous velocity, and the moving vectors according to the past best solution and global best solution. The best solution is then kept; each particle accelerates in the directions of not only the local best solution but also the global best position. If a particle discovers a new probable solution, other particles will move closer

to it in order to explore the region. In general, there are three attributes, the particles current position, current velocity, and past best position, for particles in the search space to present their features. Each particle in the swarm is updated according to the aforementioned attributes. Currently, a commonly used version is the one proposed by Shi [23], in which an adaptive parameter, named inertia weight, is used in order to enhance the performance of the original version of PSO.

In the PSO iteration, each feasible solution is presented as a particle with a vector x, and a moving velocity represented as v. As for an n dimensional optimization, the feasible solution and velocity of the i th particle can be represented as $x_i = (x_{i,1}, x_{i,2} \cdots, x_{i,n})$ and $v_i = (v_{i,1}, v_{i,2} \cdots, v_{i,n})$, respectively. After the first iteration as for $t > 0$, each particle has its best position with respect to objective value obtained so far at time t. We refer this particle x^p as pbest. Similarly, the global best (gbest) particle is denoted by x^g, which represents the best particle so far up to time t in the entire swarm. The new velocity of the i particle is updated by

$$v_i(t+1) = w_i v_i(t) + r_1 c_1 (x_i^p(t) - x_i(t)) + r_2 c_2 (x_i^g(t) - x_i(t)), \qquad (2.1)$$

where w_i is called an inertia weight, $v_i(t)$ is the old velocity of the particle i at time t. The acceleration constants c_1 and c_2 in the above equation are adjustable. They represent the amount of tension in PSO system. Usually low values of acceleration constants allow particles to travel from target regions to other target regions. On the other hand, high values result in a sudden movement toward, or past, target regions. The acceleration constants are therefore referred as the cognitive and social rates. They represent the weighting of the acceleration terms that pull the individual particle toward the personal best and global best positions.

3 Derivation of the Gaussian Plume

As for the use of air quality models, it is well-known that the gaussian plume model is one of the state of-the-art mathematical models and it is recommended for inert pollutants. In this paper, since the deficiency of data, we assume that the concentration of the pollutants can be estimated be the Gaussian plume.
In order to derive and understand the Gaussian Plume, there are a few assumptions be made for mathematically and physically clear.

The equation is brought to calculate the steady state concentration of a pollutant at a point (x, y, z) downwind of a point source at (x', y') with height H and mass emission rate Q, and under the situation of a constant wind of speed u and standard deviations in concentrations of $\sigma_y(x' - x)$ and $\sigma_z(x' - x)$ in the crosswind and vertical directions.

The Gaussian Plume is derived from a diffusion equation under appropriate limiting, initial and boundary conditions [8–10]. The limitations on the diffusion equation and its particular solution for a typical urban situation, and the techniques for reducing the solution to the gaussian plume equation are not obvious. By applying some reasonable physical assumptions and mathematical techniques,

one can derive other gaussian plume equations. For more details, see [11]. Consider the diffusion of a non-reacting pollutant in the atmosphere. It is described by

$$\frac{\partial C}{\partial t} + \nabla \cdot (Cu) = \nabla \cdot (K \nabla C) + S, \tag{3.1}$$

where C is the concentration, K is the diffusivity tensor, u is the average wind velocity, and S is a source function. This partial differential equation (3.1) is derived from the conservation of mass in a bounded volume. The diffusion by turbulent eddies is not strictly Fickian. However, the equation (3.1) can be applicable with empirically variable coefficients.

We need to make a few simplifying assumptions that allow us to derive a closed-form of analytic solution:

- The source is emitted at a constant rate Q [kg/s] from a single point source $X = (0, 0, H)$ located at H above the ground surface. For a steady state point source with emission rate Q,

  $$S(x, y, z) = Q\delta(x)\delta(y)\delta(z - H),$$

  where, δ is the Dirac delta function.
- Constant wind in one direction (u = constant, $v = w = 0$)
- Crosswind and vertical diffusion vary with downwind distance only ($K_x(x) = K_y(x) = K_z(x) = K_z(x) = K(x)$,
- The wind velocity is sufficiently large that diffusion in the x-direction is much smaller than advection; then the term $K_x\partial_x^2 C$ can be neglected

With these assumptions, the equation (3.1) reduces to

$$u\frac{\partial^2 C}{\partial x^2} = K\frac{\partial^2 C}{\partial y^2} + K\frac{\partial^2 C}{\partial z^2} + Q\delta(x)\delta(y)\delta(z - H). \tag{3.2}$$

Here, the considered domain is $\{(x, y, z) \in [0, \infty) \times (-\infty, +\infty) \times [0, \infty)\}$. Finally, the boundary conditions of this PDE are:

$$C(0, y, z) = 0, \lim_{x\to+\infty} C(x, y, z) = 0, \lim_{y\to\pm\infty} C(x, y, z) = 0, \lim_{z\to+\infty} C(x, y, z) = 0,$$

and

$$K\frac{\partial C}{\partial x}(x, y, 0) = 0.$$

This PDE (3.2) together with the above boundary conditions represents a well-posed problem for a steady-state contaminate concentration, see [11] for details.

It is common practice to replace the independent variable x with the new independent variable

$$r = \frac{1}{u}\int_0^x K(\xi)d\xi,$$

which has a unit of m^2. When the reflecting boundary condition is imposed and K is a constant, then we have

$$C(r,y,z) = \frac{Q}{4\pi ur}\exp(\frac{y^2}{4r})\exp(-\frac{(z-H)^2}{4r}) - \exp(-\frac{(z-H)^2}{4r}).$$

Here, by direct calculation, it is easy the see that

$$r = Kx/u.$$

4 Inverse Model

The recognition of the locations and the emission rate of unknown sources, starting from the detecting of the concentration of the pollutants, is known as inverse model [3, 4]. This mathematical technique could be used to identify the most relevant pollution sources which release the illegal chemical substance in the atmosphere.

In this inverse model, we apply a least square formulation technique minimizing the square of the difference between concentration measurements from the sensors and the theoretical concentrations which in this case we use the gaussian plume model. By starting from a finite set of concentration measurements, $i = 1, 2, \cdots n$ where n is the number of sensors, the least square method used to estimate parameters of sources (locations and emission rate) from n observations of concentration measurements of μ_i and from m unknowns of sources is based on the minimization of the sum of square of residuals represented by the function J.

Let $X_i = (x_i, y_i, z_i),\ i = 1, 2, \cdots, m$ be vectors of location coordinates of pollutant sources and $Q = (q_1, \cdots, q_m)$ be a vector of emission rates of each source. X_i and Q are decision variables.

$$J(X_1, \cdots X_m, Q) = \frac{1}{2}\left[\sum_{i=1}^{n}\mu_i - C_T(\bar{X}_i, X_1, \cdots X_m, Q)\right]^2, \tag{4.1}$$

where $\bar{X}_i = (\bar{x}_i, \bar{y}_i, \bar{z}_i),\ i = 1, 2, \cdots, n$ are vectors of location coordinates of pollutant sensors. These vectors are given data and

$$C_T(\bar{X}_i, X_1, \cdots X_m, Q) = \sum_{j=1}^{m}\bar{C}(\bar{X}_i, X_j, q_i),$$

with

$$\bar{C}(\bar{X}_i, X_j, q_i) = \frac{q_i}{4\pi K X_j}\exp\left[\frac{-u(\bar{y}-y_j)^2}{4x_j}\right]\left[\exp\left[\frac{-u(\bar{z}-z_j)^2}{4z_j}\right] + \exp\left[\frac{-u(\bar{z}+z_j)^2}{4z_j}\right]\right].$$

5 Numerical Experiment: Identification of the Pollution Sources

In this section, the algorithm is tested on a number of examples, ranging from 10 kilometers to 50 kilometers in north of Chiang Mai area. The artificial concentration data is employed from the Gaussian Plume. On each measurement by sensors, we utilize the reading from the Gaussian Plume. Therefore, theoretically the optimal value could be as low as zero. At first,we put 5×5 and 5×10 imaginary grids on the domain that we consider for possible locations $p_{ij}, i, j = 1, 2, 3, 4$ of pollutant sensors as shown in figures 1 and 2. A rectangular coordinating system is also imposed to locate the exact locations of studied objects. A unit is in meter.

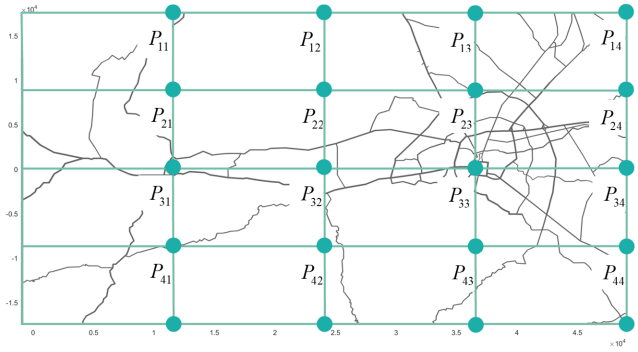

Figure 1: 5 by 5 grids and possible locations of sensors

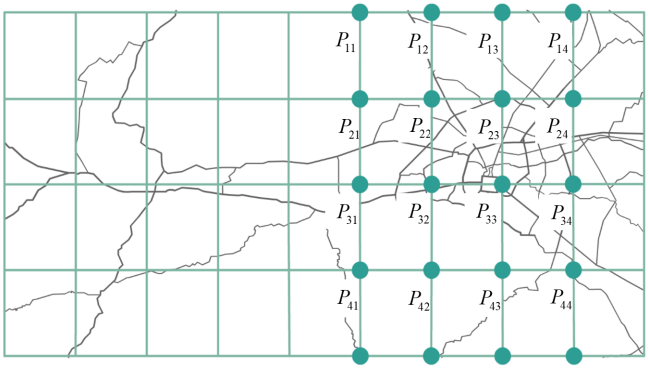

Figure 2: 5 by 10 grids and possible locations of sensors

One of our goals is to find an optimal location to place 4 air pollution sensors around the city of Chiang Mai, Thailand, to monitor the air quality effectively.

First, we place the exact location of the contamination source at the coordinate $(288, 77)$. We then place 4 sensors on a 5×5 grids, and measure the performance

of our algorithm. We assume that the contamination coincides with the Gaussian plume. There are 625 different ways to place 4 sensors on possible location p_{ij}. On each setting, we execute the algorithm 5 times, measure the errors and then take the average of the errors. With respect to average errors, the top 10 setting are listed in table 1 along with the predicted location as in table 2.

No.	TYPE	1st	2nd	3rd	4th	5th	AVG
1	189	0.96	1.10	0.47	2.05	0.04	0.93
2	558	0.48	1.38	3.92	0.25	0.14	1.23
3	303	0.69	2.87	1.56	0.47	1.70	1.46
4	363	0.89	0.29	8.66	2.45	3.10	3.08
5	283	2.18	17.43	2.36	0.34	4.27	5.32
6	62	148.78	0.25	2.05	0.15	0.28	30.30
7	436	0.59	148.10	1.26	0.55	0.50	30.38
8	187	1.29	0.44	0.40	1.54	148.49	30.43
9	338	0.64	1.53	0.56	1.39	149.10	30.82
10	564	0.49	1.08	3.67	149.52	0.22	30.10

Table 1: Top 10 average distances error between exact and predicted locations in meters with 4 sensors on 5 by 5 grid

No.	Location	1st	2nd	3rd	4th	5th
1	x	287.70	288.51	287.72	286.91	288.04
	y	76.09	77.98	76.62	75.26	77.01
2	x	288.29	287.35	290.27	287.89	288.09
	y	77.38	75.785	80.19	76.78	77.11
3	x	288.52	290.12	286.97	288.34	286.84
	y	77.45	78.94	75.83	77.32	75.76
4	x	288.66	287.78	294.35	289.87	290.34
	y	77.60	76.80	82.90	78.59	79.04
5	x	289.58	301.03	286.36	287.80	284.90
	y	78.50	88.58	75.30	76.71	74.07
6	x	284.87	288.15	286.83	287.90	288.00
	y	-71.75	77.20	75.32	76.89	76.72
7	x	288.35	284.94	287.37	288.44	287.75
	y	77.47	-71.97	75.94	77.33	76.56
8	x	287.49	288.36	287.84	287.03	284.78
	y	75.81	77.26	76.64	75.81	-71.46
9	x	288.41	287.00	288.40	287.01	283.97
	y	77.49	75.85	77.39	76.03	-72.94
10	x	288.16	288.61	286.07	285.28	287.89
	y	77.47	77.89	73.88	-72.49	76.81

Table 2: Top 10 results of the coordinate of the predicted location of the source using 4 sensors on 5 by 5 grid.

By considering only the top 10 candidate settings, we wish to determine which of the top 10 settings can predict the location of the source more accurately if the source get moved. Therefore, we consider the pollutant sources at coordinate $(288, 8750)$ and $(288, -8750)$ and repeat the experiments. Table 3 presents average errors for each setting on two locations of the sources.

Rank.	TYPE	AVG Error (288,8750)	TYPE	AVG Error (288,-8750)
1	436	741.74	283	2.51
2	564	1463.70	558	7.11
3	189	3494.40	187	13.53
4	283	3711.88	436	405.67
5	338	4143.90	189	1612.70
6	62	4499.88	62	1820.24
7	558	6524.12	564	3234.91
8	187	6994.36	363	17190.58
9	363	10701.26	338	17664.22
10	303	11132.55	303	24645.90

Table 3: The table shows average distance errors for different types of locations setting of the sensors for both locations of the contamination sources.

Table 3 displays the average distance errors for different types of location setting of the sensors for both (288,8750), and (288, -8750) coordinates of the contamination sources. From all three locations we found that the best sensor location has the coordinates listed in table 4.

x	11500	24000	36500	49000
y	8750	0	0	-17500

Table 4: The optimal coordinates for placing a set of 4 sensors for detecting a single pollutant source on 5 by 5 grid

Figure 3 displays an aerial photo of a the northern part of Chiang Mai indicating with the exact location of one source, approximated location of the sources and the optimal 4 sensor locations. The size of the depicted area is roughly 30 x 50 km. Figure 4 shows the closer look of how accurate the method is.

We try to put 4 sensors on a finner grids, the 5×10 grids. We repeated the similar experiments to obtain the optimal coordinates for placing a set of 4 sensors locating closer city on a 5×10 grid. Table 5 gives the optimal coordinate for placing a set of 4 sensors locating closer to city.

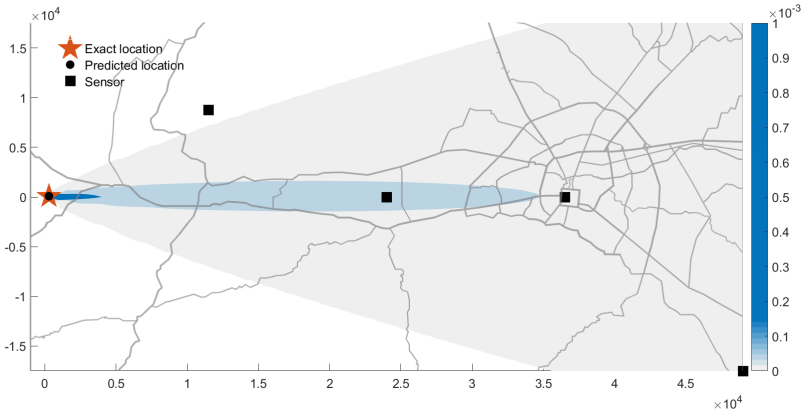

Figure 3: Optimal location for sensors on 5 by 5 grids detecting one pollutant source

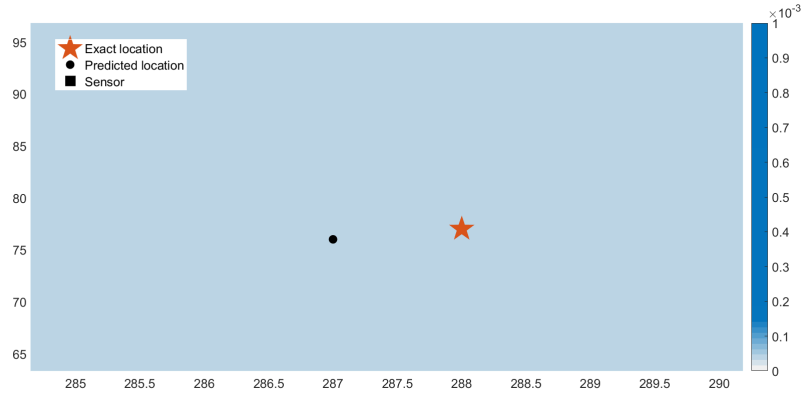

Figure 4: Enlarged image of exact location and predicted location

x	26777.78	32333.33	37888.89	43444.44
y	0	8750	0	0

Table 5: The optimal coordinate for placing a set of 4 sensors locating closer to city.

Figure 5 displays an aerial photo of a the exact location of one source, approximated locations of the sources and the optimal locations for sensors closer to the city. Figure 6 shows the closer look of how accurate the method is.

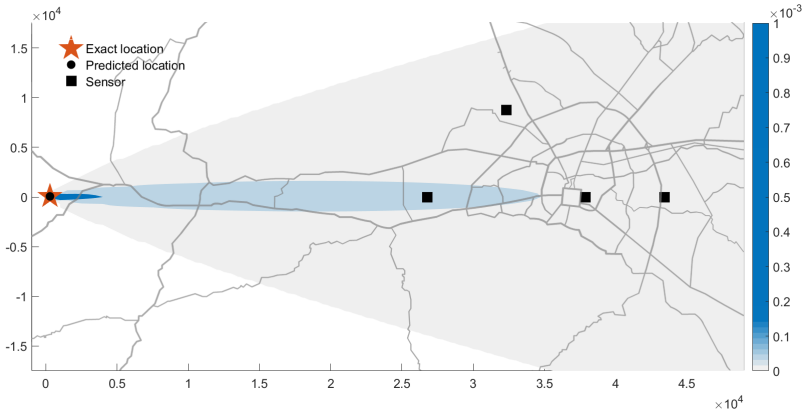

Figure 5: Optimal location for sensors on 5 by 10 grids detecting one pollutant source

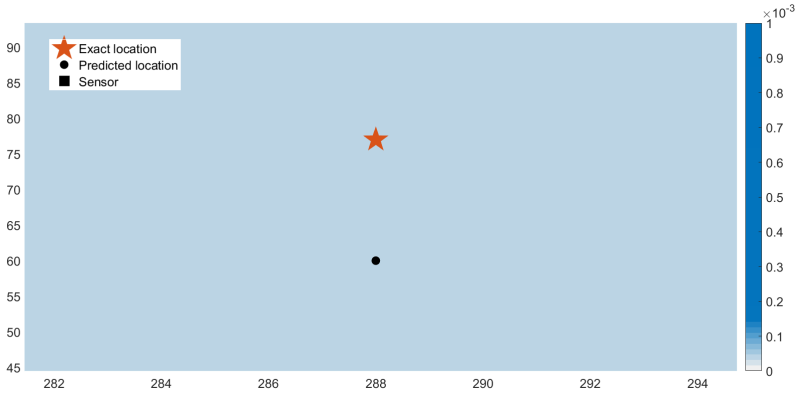

Figure 6: Enlarged image of exact location and predicted location

We also try to retrieve information on both locations and emission rates of the contamination sources. Since predicting the emission rate requires an additional

unknown, the problem of predicting both locations and emission rate of the sources becomes more complex than the problem of predicting only for the locations. Therefore, to obtain an information on both rates of emission and locations of each source, the accuracy is most likely dropped. We run our algorithm for solving such a problem, where the feasible solution is the location of the source (x, y) and the emission rate Q. For each setting, we execute our algorithm and the find its average.

To retrieve both the exact location and emission rate, we need to determine the suitable parameters $w : c$ in the equation 2.1 for this problem. To achieve that, we run all possible combinations of $w : c$ where $w, c \in \{0, 0.1, 0.2, \cdots, 1\}$ on a number of test problems. According to our experiments, the best ratio between w and c is 0.8:0.4. Thus we use this ratio to find a source location and emission rate.

We then preform experiments of finding a location for a set of 4 sensors to detect the source locations and emission rate most accurately. From the experiments, we discover that the best location for 4 sensors has coordinate location listed in table 6.

x	11500	24000	36500	49000
y	-8750	0	0	8750

Table 6: The optimal coordinate for placing a set of four sensors to retrieve the location and emission rate of the source

No.	TYPE	AVG location Error (meters)	AVG emission rate Error (g/min)
1	439	15.37	0.01
2	308	19.51	0.02
3	314	27.86	0.02
4	303	49.18	0.01
5	313	59.70	0.05
6	315	69.39	0.02
7	189	73.32	0.02
8	322	74.10	0.06
9	311	78.30	0.08
10	339	78.52	0.07

Table 7: Average distance and emission rate errors for 5 $\times$ 5 grids

Table 7 confirms that our algorithm can retrieve information on both the location of the source and the emission rate as well. Figure 7 displays an aerial photo of the exact location of one source, approximated location of the sources and the optimal sensor locations. The size of the depicted area is roughly 30 x 50 km. Figure 8 shows the closer look of how accurate the method is.

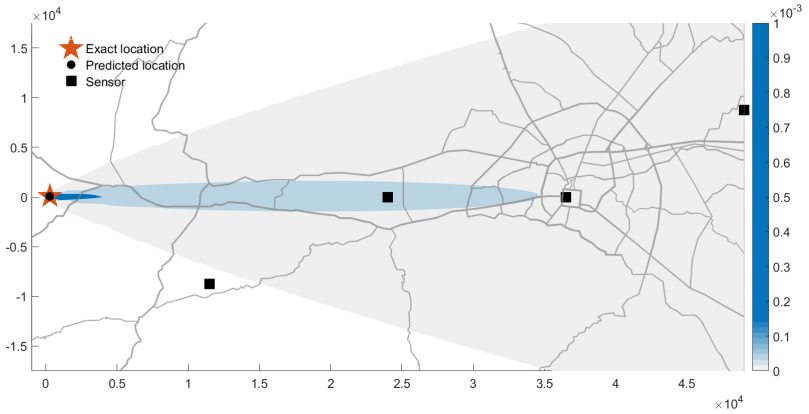

Figure 7: Exact and predicted location with rate prediction by the best 4 sensor position located

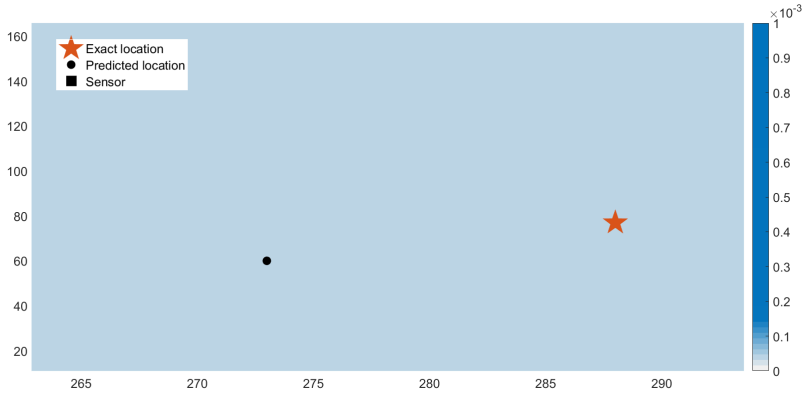

Figure 8: Enlarged image of finding location with rate prediction by the best sensor position

The error of the prediction of the emission rate is about $\pm 10^{-2}$ g/min. Table 8 is a optimal locations for placing a set of 4 sensors.

x	26777.78	32333.33	37888.89	43444.44
y	-8750	8750	8750	-8750

Table 8: The optimal coordinate for placing a set of 4 sensor to detect location and emission rate of the source for the inner city of Chiang Mai.

Figure 9 presents optimal location of the sensors, and Figure 10 displays how close of the prediction to the exact location.

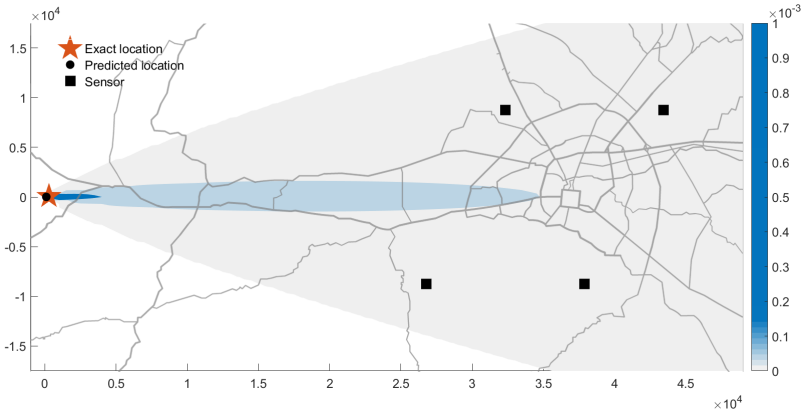

Figure 9: Location of the 4 sensors within the city detecting the location and the emission rate

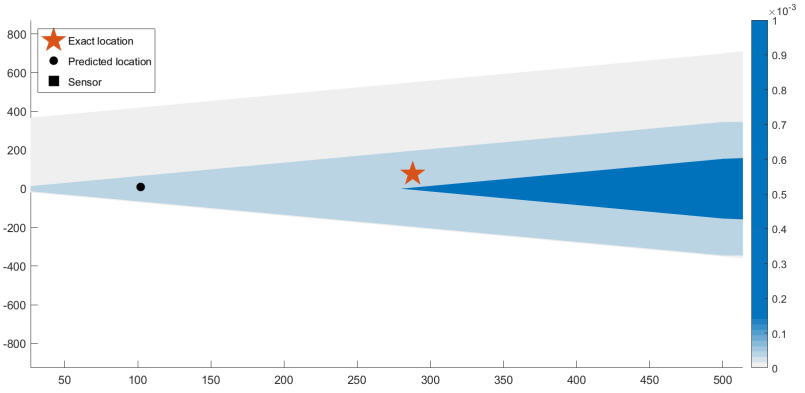

Figure 10: Enlarged image of the exact and the predicted location of the source

The algorithm can also predict the locations of the multiple contamination sources accurately. Table 10, figure 11 and 12 confirm the effectiveness of our algorithm.

x	26777.78	32333.33	37888.89	43444.44
y	-8750	8750	8750	-8750

Table 9: The optimal coordinate for placing a set of four sensors to detect location of the two source for the inner city of Chiang mai.

	source 1st	source 2nd
Error of distance (m)	32.21	216.17
x (m)	332.16	3784.24
y (m)	-5001.74	4986.78

Table 10: Predicted locations of two sources and errors

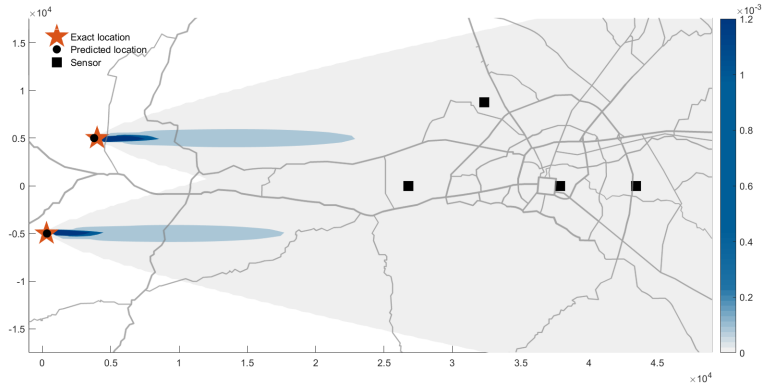

Figure 11: Locations of by the sensors detecting 2 sources

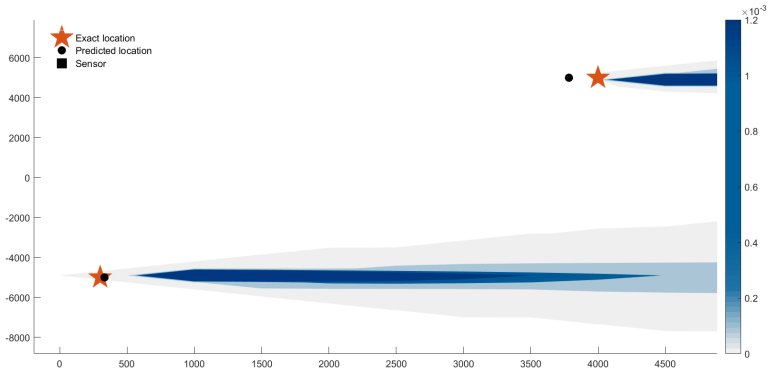

Figure 12: Enlarged image of exact and retrieved source locations

6 Concluding Remarks

In this research, we are able to use optimization technique to find a potential solution to an environmental problems. The numerical experiments show that the method has a good performance and computationally inexpensive for retrieving both source locations and rate of emission. The applications of this technique can be extended greatly. Instead of using Gaussian Plume estimation, other numerical method can be brought to improve the problem more realistic. All boundary conditions, wind directions can be varied input. It could also be implemented to give real time results.

Acknowledgement : This research is supported by the Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok, Thailand, 10400.

References

[1] A. Tiwary, J. Colls, Air Pollution, Routledge, London 2010.

[2] S.P. Arya, Air Pollution Meteorology and Dispersion, Oxford University Press, New York, 1999.

[3] A. Cantelli, F. D'Orta, A. Cattini, F. Sebastianelli, L. Cedola, Application of genetic algorithm for the simultaneous identification of atmospheric pollution sources, Atmos. Environ. 115 (2015) 36-46.

[4] A. Cantelli, G. Leuzzi, P. Monti, P. Viotti, An inverse modelling approach for estimating vehicular emissions in Urban Coastal areas of the Messina strait, Int.J. Environ. Pollut. 50 (1) (2012) 274-282.

[5] J. Kennedy, R. Eberhart, Particle swarm optimization, Proceedings of IEEE International Conference on Neural Networks IV (1995) 1942-1948.

[6] X.S. Yang, S. Deb, S. Fong, Accelerated particle swarm optimization and support vector machine for business optimization and applications, NDT 2011, Springer CCIS 136 (2011) 53-66.

[7] Y. Zhang, A comprehensive survey on particle swarm optimization algorithm and its applications, Mathematical Problems in Engineering (2015) http://dx.doi.org/10.1155/2015/931256.

[8] C.V. Chrysikopoulos, L.M. Hildemann, P.V. Roberts, A three-dimensional steadystate atmospheric dispersion-deposition model for emissions from a ground-level area source, Atmos. Environ. 26A (1992) 747-757.

[9] J.H. Seinfeld, S.N. Pandis, Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, John Wiley & Sons, New York, 1998.

[10] F.B. Smith, The problem of deposition in atmospheric diffusion of particulate matter, J. Atmospheric Sci. 19 (1962) 429-435.

[11] J.M. Stockie, The Mathematics of atmospheric dispersion modeling, SIAM REVIEW 53 (2) (2011) 349-372.

(Received 14 September 2018)
(Accepted 16 October 2018)