Air pollution sources identification precisely based on remotely sensed aerosol and Glowworm Swarm Optimization

Yunping Chen *, Weihong Han, Wenhuan Wang, Yaju Xiong, Ling Tong
School of Automation Engineering
University of Electronic Science and Technology of China,
Chengdu, China
*E-mail address: chenyp@uestc.edu.cn

***Abstract*—In this paper, we developed a novel method to identify air pollution sources based on remotely sensed aerosol data and Glowworm Swarm Optimization (GSO). In practice, it is usually to identify the air pollution sources to certain industries, such as transportation, power plants, biomass burning, and et.al. To our knowledge, the problem of locating and quantifying the pollution to the specified factories is faced for the first time. In this study, the aerosol retrieved from remotely sensed image and GIS were used to locate and quantify the pollution to each enterprise in the study area based on an improved Glowworm Swarm Optimization and meteorological condition. As a result, the gross and intensity of every enterprise in the study area were achieved. Therefore, the polluting contribution of each factory could be listed, and the most polluting factories could be found. Some experiments were carried out to validate the method, and the Key monitoring factories by local authority was ferreted out accurately.**

Keywords—Glowworm Swarm Optimization; pollution source identification; remote sensing; aerosol

I. INTRODUCTION (*HEADING 1*)

Nowadays, the ambient air quality is a serious matter which puzzles many countries, especially developing countries, such as China, India. The status of air quality is described according to biological, chemical and physical properties. Based on the properties, the quality of air can be expressed via a numerical index such as an air pollutant index (API), by combining measurements of selected air quality variables. Poor air quality has both acute and chronic effects, especially to human health[1]. Hence, here is a compelling need for quantification, identification and apportionment of these pollutants in order to facilitate their reduction through proper management plans.

Generally, the source identification methods can be cataloged into two classes, the top-down approach and the bottom-up approach. Emission inventories and chemistry-transport models are the classical method in bottom-up approach, while receptor model, one of top-down approach, is the most widely used model. All of them are the important tools to the evaluation of air pollution. However, these models have some limitations. To Emission inventories, test data from individual sources are not always available and, even then, they may not reflect the general pattern for actual emissions over time[2]. To receptor model, except the expensive sampling cost, some shortages also limit it's wider using, i.e., the source with similar compositions is hard to distinguish.

With the development of satellite technology, remote sensing provides a new way to monitor the air quality of a large area. Compared with the ground-based measurements, remote sensing image, due to its large spatial coverage and reliable repeat overpass, provides an effective way to monitor air quality on large spatial scales periodically. One important and common aerosol parameter retrieved from satellite remote sensing image is AOT (Aerosol Optical Thickness), which is mainly determined by the vertical distribution of aerosols[3]. AOT, as an aerosol optical property, can be retrieved from satellite remote sensing image obtained by Space-borne sensor, such as Moderate Resolution Imaging Spectroradiometer (MODIS)[4, 5], Polarisation and Directionality of Earth's Reflectance (POLDER)[6], Mulit angle Imaging Spectro-Radiometer (MISR)[7], HJ-1[8], Landsat TM [9]. In general, a higher AOT value indicates that the area has higher aerosol concentration and lower visibility. So, some studies had used AOT to represent the aerosol and the relationship with air quality [10].

The objective of this study is to develop a new mothed to identify air pollution source without the emission inventories and in-situ sampling. The retrieved AOT from remote sensing image is assumed as air pollutant, and an improved GSO algorithm was employed to implement the pollution sources identification process.

II. MATERIALS AND METHODS

A. Study Sites

The study area is the area of Chengdu, which is a central city of the west of China, enjoying the reputation of "the heart of the west". However, with the rapidly development of the city, some environmental issues have begun to emerge. With the addition of the significant feature of the climate (cloudy and mist, shortage of sunshine, high air humidity) and the enclosed geographic condition (located in the center of Sichuan basin),

978-1-5090-1893-2/15 $31.00 © 2015 IEEE
DOI 10.1109/SmartCity.2015.56

air pollutants in the city cannot diffuse rapidly, but be easy to accumulate in the city and suburban. According to the resolution of the aerosol retrieved by MODIS data, two areas in Chengdu's Dujiangyan county and Xinjing county were selected as the experiment areas (see Fig. 1).

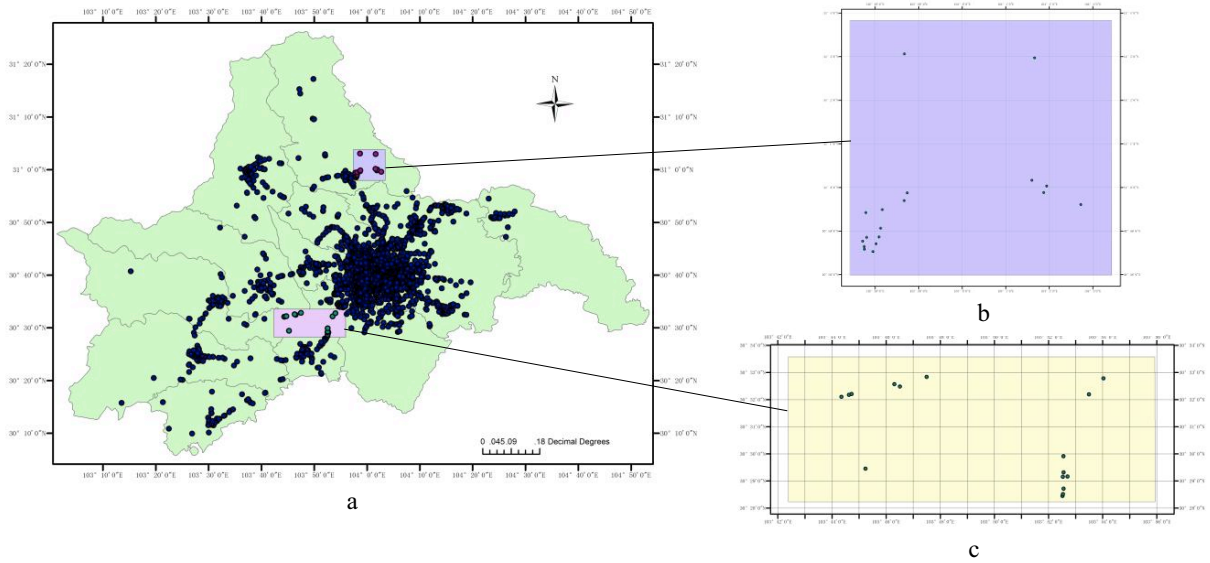

Fig. 1 the study region in Chengdu. a: the two study's location in Chengdu area; b: the second experiment area; c: the first experiment area.

B. Data Collection

According to the update date of GIS data, the collection 6 MODIS L1B product (MOD021KM), h04/v30 on March 17, 2009, was used to study the experiment area. The DDV (Dense Dark Vegetation) algorithm was employed to retrieve aerosol from the MODIS L1B product. The GIS information of study area's enterprises was collected up to 2009.

C. Glowworm Swarm Optimization

Glowworm swarm optimization (GSO) algorithm is the one of the newest nature inspired heuristics for optimization problems developed by Krishnanand and Ghose[11, 12]. It is improved from ACO approach to continuous optimization. GSO was first used for optimizing multimodal functions with equal or unequal objective function values. In GSO, each artificial glowworm carries a luminescence quantity called luciferin along with them and has its own vision, called local-decision range. The luciferin level is associated with the objective function of the agent's positon. The glowworm with a better position is brighter than others, and therefore, has a higher luciferin level value and is very close to one of the optimal solutions. The luciferin level L_j is updated using the following equation:

$$L_j(t) = (1-\rho)L_j(t-1) + \gamma F(p_j(t)) \quad (1)$$

Where, $L_j(t-1)$ is the previous luciferin level for glowworm j; ρ is the luciferin decay constant ($\rho \in (0, 1)$); γ is the luciferin enhancement fraction, and $F(p_j(t))$ represents the objective function value for glowworm j at current glowworm position (p_j); t is the current iteration.

After that, GSO algorithm gets into movement phase. Each glowworm decides, using a probabilistic mechanism, to move toward a neighbor that has a luciferin value higher than its own. That is, glowworms are attracted to neighbors that glow brighter. For each glowworm i, the probability of moving toward a neighbor j is given by:

$$p_{ij}(t) = \frac{L_j(t) - L_i(t)}{\sum_{k \in N_j(t)} L_k(t) - L_i(t)} \quad (2)$$

Where, $j \in N_j(t)$, $N_i(t) = \{j : d_{ij}(t) < r_d^j(t); L_i(t) < L_j(t)\}$ is the set of neighbors of glowworm i at time t, $d_{ij}(t)$ represents the Euclidean distance between glowworms i and j at time t, and $r_d^i(t)$ represents the variable neighborhood range associated with glowworm i at time t. Based on equation 2, glowworm i select a glowworm $j \in N_i(t)$ with $p_{ij}(t)$, and the discrete-time model of the glowworm movement can be stated as:

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\left\|x_j(t) - x_i(t)\right\|}\right) \quad (3)$$

Where, $x_i(t) \in \mathbb{R}^m$ is the location of glowworm i, at time t, in the m-dimensional real space $\mathbb{R}^m$, $\|\cdot\|$ represents the Euclidean norm operator, and s (>0) is the step size.

In GSO, the neighborhood range is not constant, but update in every iteration. A substantial enhancement in performance is noticed by using the rule given below:

$$r_d^i(t+1) = \min\{rs, \max\{0, r_d^i(t) + \beta(nt - |Ni(t)|)\}\} \quad (4)$$

Where, β is constant parameter and n_t is a parameter used to control the number of neighbors.

D. Improvement of GSO

In this section, we discuss two new factors which influence the movement phase of GSO, 1) the wind, and 2) the attribute of 'glowworm'. In the original GSO, the movement of a glowworm depends on the luciferin, which is decided by the objective function value and glowworm position. The attribute of each glowworm is not to consider. However, in our case, the attribute of glowworm, i.e. the aerosol value of the point, is an important factor influenced movement, since the points with similar aerosol level are more likely to come from the same pollution source.

For each glowworm *i*, considered the attribute of the glowworm, an additional force except the luciferin of moving toward a neighbor *j* is given by:

$$\Delta_{attribute} = s_a \cdot \left(\frac{x_j(t) - x_i(t)}{\left\| x_j(t) - x_i(t) \right\|}\right) \quad (5)$$

$$x_j = \min\left(\left| x_{att\_k} - x_{att\_i} \right|\right) \quad (6)$$

Where $j \in N_j(t)$, $N_i(t) = \{j : d_{ij}(t) < r_d^i(t); L_i(t) < L_j(t)\}$ is the set of neighbors of glowworm *i* at time *t*, just same as equation 3. $\Delta_{attribute}$ represents the movement forced by the similarity of attribute. s_a is the step size. In our research, it is same as s in

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\left\| x_j(t) - x_i(t) \right\|}\right)$$

equation (3. x_{att_i} is the attribute value of glowworm *i*, and x_{att_k} represents each attribute value of the glowworms in the neighborhood range of glowworm *i*.

Furthermore, the influence of wind was considered in this research, and an additional force caused by wind is given by:

$$\Delta_{wind} = s_w \cdot \overrightarrow{speed}_{wind} \quad (7)$$

Where Δ_{wind} represents the movement caused by wind. s_w is the scale factor, depended on the ratio of the wind speed and the resolution back ground map. In this study, the unit of map is kilometer, while the unit of wind speed is meter. Therefore, s_w=0.001. $\overrightarrow{speed}_{wind}$ is the wind vector, which depended on the wind speed and direction.

Considered wind and the similarity of attribute, as a result, the equation is modified as the following:

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\left\| x_j(t) - x_i(t) \right\|}\right) + \Delta_{attibute} + \Delta_{wind} \quad (8)$$

III. RESULTS AND DISCUSSION

Two areas were selected as the experiment areas. The reason why the two areas were selected is, in these two areas, there is one key monitoring enterprise respectively bulletined by local authority, ChengdDu Environment Protection Bureau (CDEPB). In these experiments, we tried to list the polluting contribution of each enterprise, and ferret out the key monitoring enterprise among the enterprises in this area.

As mentioned above, the MODIS L1B product, on March 17, 2009, was used to retrieve aerosol. The conventional aerosol retrieval algorithm, DDV, was employed to derive AOT from MODIS image. The resolution of retrieved aerosol is 1 km. After that, the grid AOT product was converted to points for the use of the improvement GSO (see Fig. 2.and Fig. 3 Left). In Fig. 2 and Fig. 3, each dot represents an AOT value. The enterprise positions were converted from the GIS vector points to raster point with the tif format. In Fig. 2 and Fig. 3, each little rectangle represents an enterprise, while some enterprises are merged to one because of the relative coarse AOT resolution. The AOT product map and enterprise map were registered strictly.

An improved GSO algorithm was implemented using Matlab to carry out pollution source identification process.

The algorithm parameters values have been determined based on extensive numerical experiments. Luciferin initial value L_0=2, the sensor range rs=2, adaptive decision-range initial value r_0=3, neighborhood threshold n_t=5, update rate of dynamic decision-range β=0.08, step-size s=0.03, dissipation rate of luciferin value ρ=0.2, update rate of luciferin value γ=0.6. The iteration time is 200.

A. Evaluation Measures

In order to characterize the pollution situation of enterprises, three indexes were calculated named as the pollution gross, the pollution intensity, and the area normalization pollution. The pollution gross, abbreviated as PG, referred to the ATO sum of the pollution points surrounded on a pollution source. The calculation equation is as follow:

$$PG_i = \sum_k^n a_k \qquad k \in A \quad (9)$$

Where a_k represents the AOT value of a point, A presents the set of the points which clustered in the specified range around the pollution source *i*.

The pollution intensity (PI) was defined as:

$$PI_i = \frac{\sum_k^n a_k}{n} = \frac{PG_i}{n} \qquad k \in A \quad (10)$$

Obviously, PI_i is the average value of the points which clustered in the specified range around the pollution source *i*.

Considered that the distribution of enterprises is not uniform, some enterprise is isolated, while some are clustered. Sometimes, the isolated enterprise is more likely to attract the pollution points. In order to avoid this effect, the area normalization pollution was proposed, and defined as:

$$ANP_i = \frac{\sum_k^n a_k}{area_i} = \frac{PG_i}{area_i} \qquad k \in A \quad (11)$$

Where $area_i$ represent the exclusive maximum area by the pollution source *i*, that is, in this area, there are just one

pollution source. The radius of the area is the distance from the pollution source *i* to the nearest pollution source.

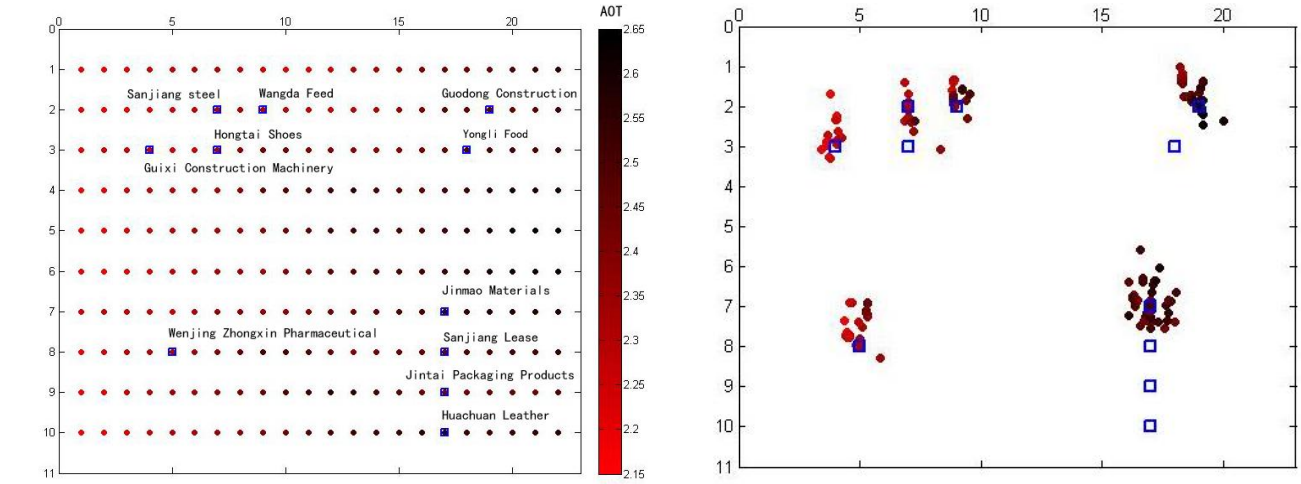

Fig. 2. The pollution source position and the pollutant distribution in first area. The pollutant distribution before running algorithm (left). The pollutant distribution after running algorithm (right). The colorbar represents AOT (aerosol optical thickness).

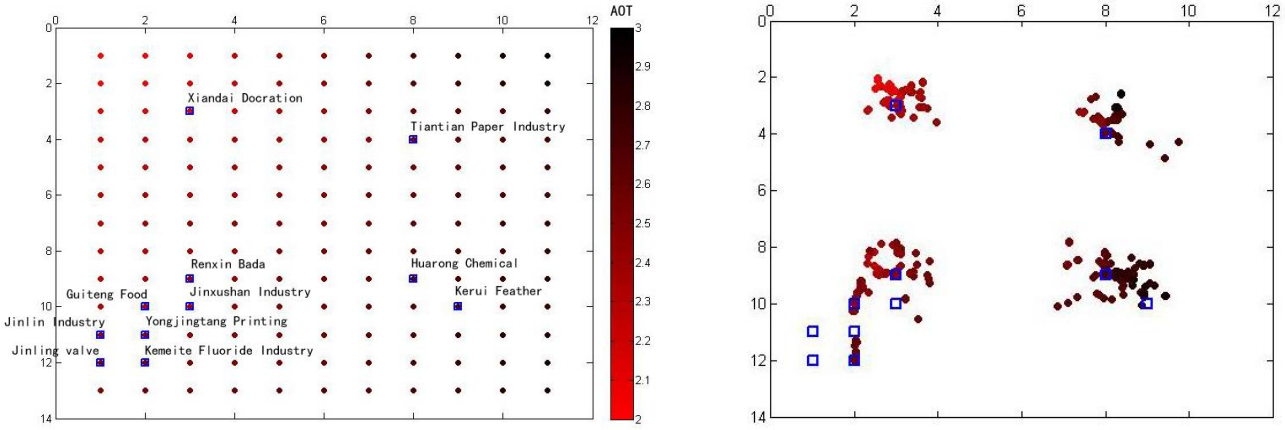

Fig. 3 The pollution source position and the pollutant distribution in second area. The pollutant distribution before running algorithm (left). The pollutant distribution after running algorithm (right). The colorbar represents AOT.

B. Result

In Table 1 and Table 2, the pollution indexes of 11 enterprises were listed in the two experiment area respectively. Guodong Construction is the key monitoring enterprise bulletined by CDEPB in the first area. From table 1, we could find that Guodong construction has the highest pollution gross, highest pollution intensity, and highest area normalization pollution in that area. And some enterprise, such as Jinmao Materials, Wangda Feed, Sanjiang Steel, are also the key enterprises contributed to pollution.

Table 1 the pollution statistics for each enterprise in the first area

Enterprises	Pollution Gross	Pollution intensity	Area normalization pollution
Guixi Construction Machinery	36.31	2.24	1.28
Wenjin Zhongxin Pharmaceutical	38.83	2.30	0.49
Sanjiang Steel	42.12	2.32	13.41
Hongtai Shoe Industry	9.39	2.34	2.99
Wangda Feed	42.71	2.39	3.40
Jinmao Materials	62.74	2.50	19.97
Sanjiang lease	9.48	2.23	3.02
Jintai packaging products	0	0	0
Huachuan leather	0	0	0
Yongli Food	0	0	0
Guodong Construction	65.78	2.51	20.94

In Table 2, Huarong chemical is the key monitoring enterprise bulletined by CDEPB. From table 2, we could also

find that Huarong chemical has the highest pollution gross, highest pollution intensity, and highest area normalization pollution in second experiment area.

Table 2 the pollution statistics for each enterprise in the second area

Enterprises	Pollution Gross	Pollution intensity	Area normalization pollution
Jinlin Industry	0	0	0
Jinling valve	0	0	0
Guiteng Food	10.63	0.72	3.38
Yongjingtang Printing	6.27	1.01	2.00
Kemeite Fluoride Industry	13.63	2.02	4.34
Xiandai Decoration	40.80	2.24	0.52
Renxin Bada	41.40	2.41	13.18
Jinxushan Industry	4.07	2.04	1.30
Tiantian Paper Industry	40.86	2.52	0.52
Huarong Chemical	57.21	2.66	18.21
Kerui Feather	9.17	2.51	2.92

IV. CONCLUSION

In this paper, a novel pollution source identification method was proposed. Based on this method, only remote sensing data and enterprise's GIS information were needed to relatively quanlify the pollution of each factory without extremely expensive sampling or the emission inventory. As a result, the key pollution enterprise would be find out. It is very useful to the local government to control the pollution. However, as a primary research, the pollution sources are just factories, which could be regarded as points. Other important pollutions, such as city road, were not considered. There should be considered in further research.

ACKNOWLEDGMENT

This research was funded by the Prior Research Program of the 12th Five-year Civil Aerospace Plan (D040201-04) and the State Grid Corporation of China Research Funds (521997140007).

REFERENCES

[1] Moustris, K.P., I.C. Ziomas, and A.G. Paliatsos, 3-Day-ahead forecasting of regional pollution index for the pollutants NO2, CO, SO2, and O3 using artificial neural networks in Athens, Greece. Water, Air, & Soil Pollution, 2010. 209(1-4): 29-43.

[2] Elbir, T., A. Müezzinoğlu, and A. Bayram, Evaluation of some air pollution indicators in Turkey. Environment International, 2000. 26(1): 5-10.

[3] Kaufman, Y.J. and R.S. Fraser, Light extinction by aerosols during summer air pollution. Journal of climate and applied meteorology, 1983. 22(10): 1694-1706.

[4] Chu, D.A., Y. Kaufman, G. Zibordi, et al., Global monitoring of air pollution over land from the Earth Observing System‐Terra Moderate Resolution Imaging Spectroradiometer (MODIS). Journal of Geophysical Research: Atmospheres (1984–2012), 2003. 108(D21).

[5] Engel-Cox, J.A., C.H. Holloman, B.W. Coutant, et al., Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. Atmospheric Environment, 2004. 38(16): 2495-2509.

[6] Kacenelenbogen, M., J.-F. Léon, I. Chiapello, et al., Characterization of aerosol pollution events in France using ground-based and POLDER-2 satellite data. Atmospheric Chemistry and Physics, 2006. 6(12): 4843-4849.

[7] Liu, Y., J.A. Sarnat, V. Kilaru, et al., Estimating ground-level PM2. 5 in the eastern United States using satellite remote sensing. Environmental science & technology, 2005. 39(9): 3269-3278.

[8] Zhongting, W., L. Qing, T. Jinhua, et al., Monitoring of aerosol optical depth over land surface using CCD camera on HJ-1 satellite. China Environmental Science, 2009. 29(9): 902-907.

[9] Retalis, A., Assessment of the distribution of aerosols in the area of Athens with the use of Landsat Thematic Mapper data. International Journal of Remote Sensing, 1999. 20(5): 939-945.

[10] Fraser, R.S., Y.J. Kaufman, and R. Mahoney, Satellite measurements of aerosol mass and transport. Atmospheric Environment (1967), 1984. 18(12): 2577-2584.

[11] Krishnanand, K. and D. Ghose, Glowworm swarm optimisation: a new method for optimising multi-modal functions. International Journal of Computational Intelligence Studies, 2009. 1(1): 93-119.

[12] Krishnanand, K. and D. Ghose. Detection of multiple source locations using a glowworm metaphor with applications to collective robotics. in Swarm Intelligence Symposium, 2005. SIS 2005. Proceedings 2005 IEEE. 2005. IEEE.