Atmospheric Environment 81 (2013) 188–198

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Comparison and improvements of optimization methods for gas emission source identification

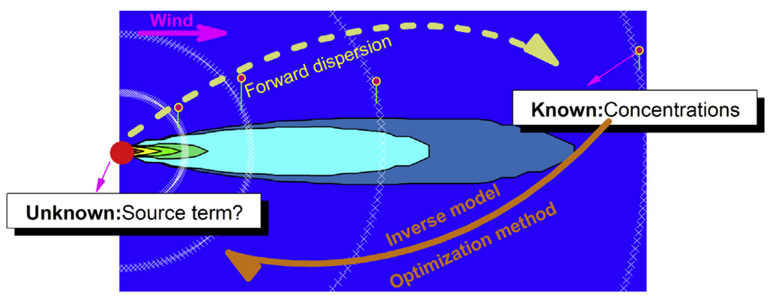
Denglong Ma [a, b], Jianqiang Deng [b], Zaoxiao Zhang [a, b, *]

[a] State Key Laboratory of Multiphase Flow in Power Engineering, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an 710049, PR China
[b] School of Chemical Engineering and Technology, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an 710049, PR China

HIGHLIGHTS

- Different optimization methods were used to identify source term.
- The effect of sensor number, distribution forms and other factors were discussed.
- Modification of cost function increases the estimation accuracy.
- New intelligent forward model improves the location estimation.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 9 May 2013
Received in revised form
3 September 2013
Accepted 5 September 2013

Keywords:
Gas leakage
Source identification
Optimization methods
Neural network
Atmosphere dispersion

ABSTRACT

Identification of gas leakage source term is important for atmosphere safety. Optimization is one useful method to determine leakage source parameters. The performances of different optimization methods, including genetic algorithm (GA), simulated annealing (SA), pattern search (PS) method, Nelder–Mead simplex method (N–M simplex) and their hybrid optimization methods, were discussed. It was seen that GA–PS hybrid optimization has the best performance for location and source strength estimation while the hybrid methods with N–M simplex is the best one when time cost and robustness are added into consideration. Moreover, the performances of these optimization methods with different initial values, signal noise ratios (SNR), sensor numbers and sensor distribution forms were discussed. Further, experiment data test showed that the less deviation of forward simulation model from the real condition, the better performance of the source parameters determination method is. When two error correction coefficients were added to the Gaussian dispersion model, the accuracy of source strength and downwind distance estimation is increased. Other different cost functions were also applied to identify the source parameters. Finally, a new forward dispersion model based on radial basis function neural network and Gaussian model (Gaussian–RBF network) was presented and then it was applied to determine the leakage source parameters. The results showed that the performance of optimization method based on Gaussian–RBF network model is significantly improved, especially for location estimation. Therefore, the optimization method with a good selection of forward dispersion model and cost function will obtain a satisfactory estimation result.

1. Introduction

The occurrence of dangerous gas emission harms human health and pollutes the environment in many cases. For example, in CO_2 capture and sequestration (CCS) case (Zhang et al., 2006; Yu et al.,

* Corresponding author. State Key Laboratory of Multiphase Flow in Power Engineering, Xi'an Jiaotong University, No. 28 Xianning West Road, Xi'an 710049, PR China. Tel.: +86 29 82660689.
E-mail addresses: zhangzx@mail.xjtu.edu.cn, zaoxiaoz@hotmail.com (Z. Zhang).

1352-2310/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.atmosenv.2013.09.012

2010; Ma et al., 2013), CO_2 leakage risk is one of the most vital parts that investigators focus on. Many methods have been proposed to determine the leakage source parameters, which can generally be divided into the direct way and indirect way. In the direct way, a number of point sensors or some mobile instruments are used to detect abnormal leakage signals and then the source parameters are determined according to the information given from the instruments. Since the measurement instruments will be moved around the monitor area, or alternatively the sensors will be distributed widely in a very large area, the cost is high and only limited information can be obtained. Moreover, it is usually impossible to obtain the leakage source strength. Therefore, another category of methods was presented, where the common practice is to solve an inverse problem. Three inverse solution methods have been investigated, including direct inverse solution (Enting, 2002; Yang et al., 2007), stochastic approximation (Cannon and Yin, 1988; Heemink and Segers, 2002; Keats et al., 2007), and optimization method. Because the source term estimation is always an ill-pose inverse problem (Victor, 2000), it is not suitable to solve the atmospheric dispersion model directly to obtain the source parameters. Stochastic approximation is a new method based on the Bayesian inferences theory, from which the probability distribution results under some confidence levels may be more reasonable because the errors of instruments and the dispersion model really exist. But this method always needs a time-consuming Markov chain Monte Carlo (MCMC) process (Keats et al., 2007). The optimization method utilizes the measurement results and a numerical model to obtain simulation results, which maximally match the measurement information. Many investigations have proven that the common optimization method coupled with Gaussian dispersion model can determine the source term accurately and rapidly (Haupt, 2005; Chow et al., 2006; Allen et al., 2007; Long et al., 2010; Selvaraju and Pushpavanam, 2010).

Haupt et al. (2009) have applied the genetic algorithms (GA) method to address air quality problems. They identified and characterized a source of contaminant despite having only imprecise knowledge of the source location, emission rate, and time of release, and it was also tested using synthetic identical twin experiments. The GA coupled model works rather well in spite of many problems. Khlaifi et al. (2009) also solved the inverse problem for quantifying SO_2 pollutant source with GA coupled with a direct model of diffusion (Pasquill's Gaussian model). Long et al. (2010) combined GA with a gradient descent algorithm to find the combination of source location, source height, source strength, surface wind direction, surface wind speed, and time of release. Thomson et al. (2007) adopted a random search algorithm, simulated annealing (SA), to locate a known gas source in a desert. Addepalli et al. (2009) characterized the source parameters of atmospheric releases by using quasi-random sampling and regularized gradient optimization. The solution methodology consists of quasi-Monte Carlo (QMC) sampling of the model parameter space and the subsequent application of gradient optimization. Konda et al. (2010) presented a grid-based estimation method of solving a convex optimization problem to identify multiple source terms.

In this paper, the different optimization methods will be tested for source parameters determination and then the estimation performance will be separately rated with skill scores. Moreover, the dependence of optimization methods on initial values, sensor number, sensor distribution form, cost function and forward dispersion model will be analyzed and improved. A new forward dispersion model based on radial basis function neural network and Gaussian model (Gaussian–RBF network) will be also discussed.

2. Basic mathematic model

2.1. Cost function of the problem

In order to estimate the leakage source parameters, the cost function of the optimization problem should be built, as Eq. (1) shows.

$$\begin{aligned} &\min \quad f = \sum_{i=1}^{N} \left[C_{\text{mea},\,i} - C_{\text{model},\,i}(Q,\,x,\,y,\,z)\right]^2 \\ &\text{s.t.} \quad Q > 0;\ x > 0;\ -\infty < y < \infty;\ z > 0. \end{aligned} \tag{1}$$

where C_{mea} (g m^{-3}) is the measurement concentration at sensor i; $C_{\text{model,i}}$ is the result from dispersion model, which is depend on the leakage rate Q (g s^{-1}), downwind distance x (m), crosswind distance y (m) and height z (m) above ground and environment conditions; N is the number of sensors. Therefore, it is a constrained minimization problem.

Usually, there are three important models to simulate the gas dispersion in atmosphere, which are the Gaussian model based on semi-analytical solution, the computational fluid dynamics (CFD) model based on N–S equations and the Lagrangian statistical (LS) model based on Markov process. The CFD and LS dispersion models are not suitable in fast source identification process because they will consume more time than the Gaussian model to finish one forward simulation process, The Gaussian model has advantages like simple implementation and low time cost with sufficient accuracy. Therefore, it has been adopted as the forward dispersion model in many applications.

The Gaussian plume model is a simple mathematical model, which is typically applied for point source emitters. The expression for a continuous point emission source is

$$\begin{aligned} C(x,\,y,\,z) = \frac{Q_c}{2\pi u\sigma_y\sigma_z}\Bigg\{&\exp\left(\frac{-(z-h-z_0)^2}{2\sigma_z^2}\right) \\ &+\exp\left(\frac{-(z+h+z_0)^2}{2\sigma_z^2}\right)\Bigg\}\left\{\exp\left(\frac{-(y)^2}{2\sigma_y^2}\right)\right\} \end{aligned} \tag{2}$$

where $C(x,y,z)$ is the concentration of the emission at any position (x,y,z). u is the wind speed (m s^{-1}). h is effective stack height (m), which is the sum of stack height, plume rise and deposition height; z_0 is the roughness height of surface; σ_y (m) and σ_z (m) are the standard deviations of a statistically normal plume in the lateral and vertical dimensions, respectively. According to the results by Briggs (1973), the standard deviations (σ_y and σ_z) depend on atmospheric stability and downwind distance. Pasquill's stability categories are adopted to classify atmospheric stability (Hanna et al., 1982).

The objective of the optimization is to reduce the deviations of the prediction results with the mathematical model from measurements to a minimum. Source location and strength are two parameters to be determined. The genetic algorithm (GA), simulated annealing (SA), pattern search (PS), and Nelder–Mead Simplex method (N–M Simplex) are applied in this paper. Among these methods, GA and SA are heuristic intelligent search methods to obtain global optimization values while PS and N–M simplex are easier to obtain local optimization values. Therefore, the hybrid optimization methods will also be discussed in our paper. Because the results from local optimization methods are usually largely dependent on the initial values while heuristic intelligent global optimization methods consume more time, the global search methods are used to obtain the initial values first and then the