GA-based Dynamical Correction of Dispersion Coefficients in Lagrangian Puff Model (Post-print)

Authors: MA Yuan-Wei, WANG De-Zhong, JI Zhi-Long

Date: 2023-06-18T00:00:00+00:00

Abstract

In atmospheric dispersion models of nuclear accidents, dispersion coefficients are typically obtained through tracer experiments and treated as constant across different atmospheric stability classifications. In reality, however, the atmospheric wind field is complex and unstable. Dispersion coefficients can vary even under the same atmospheric stability, introducing significant errors. According to regulations, the airborne concentration of nuclides around nuclear power plants should be monitored during an accident. This monitoring data can be used to dynamically correct dispersion coefficients, thereby minimizing errors. This inverse problem is nonlinear and sensitive to initial values. The ability of Genetic Algorithms (GA) to search for optimal solutions is well-suited for complex, high-dimensional scenarios. In this paper, GA is used to estimate the coefficients in conjunction with a Lagrangian dispersion model. Simulation results demonstrate that the GA scheme performs well when errors are large. However, when the correction process is applied to experimental data, the GA-estimated results become numerically unstable. The success rate of estimation is 5% lower than that without correction. Considering the continuity of dispersion coefficients, a Savitzky-Golay filter is employed to smooth the estimated parameters. The success rate of estimation increases to 75.86%. This method can improve the accuracy of atmospheric dispersion simulations.

Full Text

Preamble

ChinaXiv Cooperative Journal NUCLEAR SCIENCE AND TECHNIQUES 26, 030602 (2015)

GA-based dynamical correction of dispersion coefficients in Lagrangian puff model

MA Yuan-Wei (马元巍), WANG De-Zhong (王德忠),† and JI Zhi-Long (吉志龙)

School of Nuclear Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

(Received June 19, 2014; accepted in revised form September 20, 2014; published online June 20, 2015)

In atmospheric dispersion models for nuclear accidents, dispersion coefficients are typically obtained through tracer experiments and remain constant across different atmospheric stability classifications. However, actual atmospheric wind fields are complex and unstable, causing dispersion coefficients to vary even under the same stability conditions and introducing significant errors. According to regulations, radionuclide concentrations around nuclear power plants must be monitored during accidents. This monitoring data can be used to dynamically correct dispersion coefficients, thereby minimizing errors. This inverse problem is nonlinear and sensitive to initial values. Genetic algorithms (GA) are well-suited for searching optimal solutions in complex, high-dimensional scenarios. In this paper, coupled with a Lagrangian dispersion model, GA is employed to estimate these coefficients. Simulation results demonstrate that the GA scheme performs well when errors are large. However, when applied to experimental data, the GA-estimated results exhibit numerical instability, with a success rate 5% lower than estimation without correction. Considering the continuity of dispersion coefficients, a Savitzky-Golay filter is applied to smooth the estimated parameters, increasing the success rate to 75.86%. This method can effectively improve the accuracy of atmospheric dispersion simulations.

Keywords: Adaptive parameter, Genetic algorithm, Atmospheric dispersion, Nuclear accident

DOI: 10.13538/j.1001-8042/nst.26.030602

INTRODUCTION

Traditionally, radionuclide dispersion coefficients for nuclear accidents are preset values obtained from field campaign experiments, with differences between experimental and real-time meteorological conditions representing a major source of error. Most dispersion models employ the Pasquill-Gifford (PG) curve, established from Prairie-Grass-Field experiments [1], and similar dispersion coefficient systems to represent lateral and vertical plume spread rates. While these coefficient systems are generally used for conservative forecasting and recommended by regulations, validation through tracer experiments at nuclear power plants (NPPs) has shown that actual dispersion coefficients differ significantly from those fitted by experimental data [2]. During an NPP accident, radionuclide concentrations in air should be monitored [3], providing a viable means to dynamically correct empirical coefficients using monitored data. This correction relies solely on monitoring data rather than atmospheric stability classifications

or other information.

Haupt proposed coupling genetic algorithms (GA) with Gaussian plume and receptor models to achieve better estimation of pollutant dispersion [4]. Jeong applied similar techniques to analyze atmospheric dispersion of nuclides from an NPP [5, 6]. Allen used assimilation methods to estimate wind-field information in dispersion models [7]. However, most numerical simulations employed ideal models where the correlation coefficient between predicted and observed values was less than 0.5, and observations contained significant noise [8]. Such errors cannot be ignored in actual situations. These studies indicate that source term estimation and dynamic correction of dispersion parameters are effective methods for improving dispersion model accuracy, though the impact of dispersion coefficient errors on results requires further investigation.

To evaluate performance under diverse dispersion coefficient errors, a scheme based on GA was established and validated through numerical simulation and Kincaid experiment data [8].

II. METHOD

A. Lagrangian puff model

Parameter estimation requires a dispersion model suitable for complex meteorological conditions with high computational efficiency. Compared to Gaussian plume and particle models, the Lagrangian puff model better meets these requirements [9] and was therefore used as a forward model to calculate pollutant distributions.

The Lagrangian puff model assumes that sequential pollutant release can be treated as a series of Gaussian-shaped puffs at a fixed rate, where the concentration within each puff follows a Gaussian distribution. The total concentration at any location is the sum of contributions from all puffs at that point. The concentration distribution of a puff is described by Eq. (1):

$$C(x,y,z) = \frac{Q}{(\sqrt{2\pi})^3\sigma_x\sigma_y\sigma_z}\left[\exp\left(-\frac{(x-x_c)^2}{2\sigma_x^2}\right)\exp\left(-\frac{(y-y_c)^2}{2\sigma_y^2}\right)\exp\left(-\frac{(z-z_c)^2}{2\sigma_z^2}\right)+\exp\left(-\frac{(2z_{inv}-z_c)^2}{2\sigma_z^2}\right)\right.$$

where Q is the total concentration of the puff; (x_c, y_c, z_c) are coordinates of the puff center; z_{inv} is the height of the inversion temperature layer; (x, y, z) are coordinates of the observation point; and $(\sigma_x, \sigma_y, \sigma_z)$ are dispersion parameters of the puff.

Many empirical formulas compute dispersion parameters based on atmospheric stability and downwind distance, among which the PG curve, described in Eq. (2), is most commonly used:

$$\sigma_x = \sigma_y, \quad \sigma_y = p_y \cdot x^{q_y}, \quad \sigma_z = p_z \cdot x^{q_z}$$

where σ_x, σ_y and σ_z are computed according to downwind distance x, and the PG coefficients p_y, q_y, p_z and q_z are determined by wind-field and atmospheric stability.

B. PG coefficients correction

The PG curve is derived from Prairie Grass field experiments. When experimental conditions differ from actual meteorological conditions, errors are introduced into the model. To minimize PG error effects, one approach uses original PG coefficients as initial values and corrects them using observed values. This process can be regarded as an optimization problem that minimizes differences between calculated and observed values.

C. Genetic algorithm

The atmospheric dispersion inverse model is strongly nonlinear and sensitive to initial values. Tan used least squares methods to estimate dispersion coefficients, but most results were local optima [10], indicating the need for an effective global optimization algorithm—precisely what GA provides.

Invented by J. Holland in 1975, GA is an efficient, parallel global search method. The fitness function (f_t) is essential in GA, determining differences between observed and predicted values. Smaller f_t values indicate closer agreement between observations and predictions. Since observation weights differ due to varying observing conditions, an appropriate fitness function should satisfy two principles: (1) for observation station i, fitness function value should increase with differences between observed value O_i and calculated value C_i; and (2) observation stations should receive different weights based on data reliability.

Ji et al. compared four fitness function types and found that concentration-related weighting performs better than equal weighting [11]. The fitness function can be constructed as Eq. (3) based on the least squares method:

$$f_t = \frac{\sum_{i=1}^{N} O_i (C_i - O_i)^2}{\sum_{i=1}^{N} O_i}$$

where N is the number of observation stations; C_i is the predicted value at station i; and O_i is the observed value at station i. This fitness function modifies equal weighting by making weights proportional to observed concentration values.

III. NUMERICAL SIMULATIONS

The numerical stability of this estimation scheme was validated through simulations. White noise (serially uncorrelated random variables with zero mean and finite variance) was added to model-calculated observations. Based on these

modified observations, GA was used to estimate PG coefficients. Performance was evaluated by comparing "true" PG coefficients with estimated ones.

A. Setup of dispersion parameters

For simplified calculation in a Cartesian coordinate system, the release point was set at $(0, 0)$ with height 187 m, release rate of one unit, and SF6 as tracer gas. Wind was stable and unidirectional with atmospheric stability class D. The original PG coefficients, PG_{true}, were set as Eq. (4):

$$p_y^{true} = 1.503, \quad p_z^{true} = 0.151, \quad q_y^{true} = 0.833, \quad q_z^{true} = 1.219$$

Changes in PG coefficients were estimated based on observation values at each simulation step.

B. Configuration of observations

Signal-to-noise ratio (SNR) is defined as the ratio of signal power to noise power: $SNR = \sigma_{signal}/\sigma_{noise}$. Different SNR levels of 50, 25, 5, 2.5, and 0.25 were added to every sampling, including positive infinity. The noise is white noise with constant power spectral density. Based on modified observations, PG_{corr} was estimated to analyze error impact.

Initial parameter values were set as Eq. (6), with each parameter containing added noise:

$$p_y = 1.533, \quad p_z = 0.255, \quad q_y = 0.712, \quad q_z = 1.401$$

GA with the fitness function corrected these PG_{pre} to achieve PG_{corr}. Smaller differences between PG_{corr} and PG_{true} indicate better fitness function performance.

C. Result of numerical experiment

The error between PG_{corr} and PG_{true} is expressed by Eq. (7):

$$\sigma_{pq} = \sqrt{(\log_2 p_y^{corr} - \log_2 p_y^{true})^2 + (q_y^{corr} - q_y^{true})^2 + (\log_2 p_z^{corr} - \log_2 p_z^{true})^2 + (q_z^{corr} - q_z^{true})^2}$$

As shown in Fig. 1 [Figure 1: see original paper], the estimation parameter error increased with $1/SNR$. The variance of white noise is $1/SNR$, and at $1/SNR = 1$, $\sigma_{pq} < 1.5$, meaning PG_{corr} is very close to PG_{true}. This indicates good robustness of the scheme.

IV. EXPERIMENT VALIDATION

A. Kincaid tracer experiment

The Kincaid field experiment was conducted as part of the EPRI Plume Model Validation and Development Project. A comprehensive experimental campaign was carried out in 1980 and 1981. The Kincaid NPP is located in Illinois, USA (39.59°N, 89.49°W) at an elevation of approximately 180 m; terrain effects are therefore not considered in the model. The NPP has a 187 m stack with 9 m diameter. In the experiment, SF6 was released from the stack, with tracer releases starting several hours before sampling. The experimental campaign included approximately 350 hours of tracer experiments.

The 1284 arc-max concentration data points, typically classified as "high quality" and used for model validation and development, were not screened for this estimation. In a real accident, insufficient information exists to judge data effectiveness, so all data were used, introducing uncertainty into the estimation.

B. Validation method

Twenty experimental cases were conducted, each lasting 3 to 9 hours. Over 100 observation stations recorded SF6 concentrations in air during the release period. Statistical metrics—FA2, FA5, FB, NMSE, CORR, and BIAS—were used to measure differences between predicted and observed values:

$$FA2 = \text{fraction of data satisfying } 0.5 < C_o/C_p < 2$$

$$FA5 = \text{fraction of data satisfying } 0.2 < C_o/C_p < 5$$

$$FB = \frac{\bar{C_o} - \bar{C_p}}{0.5(\bar{C_o} + \bar{C_p})}$$

$$NMSE = \frac{(C_o - C_p)^2}{\bar{C_o}\bar{C_p}}$$

$$CORR = \frac{(C_o - \bar{C_o})(C_p - \bar{C_p})}{\sigma_o \sigma_p}$$

$$BIAS = \bar{C_o} - \bar{C_p}$$

where p denotes model prediction; o denotes observation; overbar denotes averaged dataset; and σ is the mean square error over the dataset.

V. RESULTS AND DISCUSSION

A. Results of GA estimation

Results are shown in Table 1 . In the "Original data" row, the correlation coefficient (CORR) is 0.20, indicating weak positive relationship between predicted and observed data; FA2 (0.11) and FA5 (0.26) values indicate large differences. Therefore, estimating PG coefficients is necessary to improve model quality. Using the GA method (the "GA" row) increases FA values while decreasing BIAS and FB values, demonstrating that dynamic correction improves model accuracy and efficiency even with large dispersion coefficient errors.

B. GA Prediction

The prediction process uses parameters fitted from historical data. For each case, the steps are: (1) use standard PG parameters as prior values and correct parameters using GA at every step; (2) calculate prediction values based on the last step's corrected parameters; (3) calculate prediction values based on standard PG parameters; and (4) compare predictions between steps 2 and 3 to evaluate GA parameter prediction effectiveness.

Results are given in Table 1. Most data worsened because dispersion parameter calculations in one step were not related to the next step. The GA result is only numerically optimal, not physically optimal, ignoring the continuous property of PG coefficients.

C. Smoothing filter

Although statistical values are much better than in the original model, GA-estimated results are numerically unstable, with some estimated values greatly deviating from true values. The estimation success rate was 5% lower than without correction. To reduce GA estimation fluctuations, dispersion parameters were assumed to change continuously as shown in Eq. (9):

$$p_y[t+1] = p_y[t]+\eta, \quad q_y[t+1] = q_y[t]+\eta, \quad p_z[t+1] = p_z[t]+\eta, \quad q_z[t+1] = q_z[t]+\eta$$

where η is white noise representing time-dependent changes in dispersion coefficients; t is the current time step and $t+1$ is the next time step.

Thus, before use in prediction, GA-estimated parameters were smoothed using a Savitzky-Golay filter. This method's main advantage is preserving distribution features such as relative maxima, minima, and width, which are typically "flattened" by other adjacent averaging techniques. Information in estimated dispersion coefficients is retained after Savitzky-Golay smoothing.

Using case 50780 as an example, smoothing filter results are shown in Fig. 2 [Figure 2: see original paper]. PG coefficients become more stable than before,

indicating that nuclide distribution changes smoothly step by step. Figure 3 [Figure 3: see original paper] shows statistical values for the three schemes. The "GA smooth" scheme effectively improves results while inheriting excellent characteristics of the "GA prediction" scheme, balancing stability and efficiency to perform best among the three.

Statistical values for all cases are shown in the "GA Smooth" row of Table 1. NMSE, FA2, FA5, and FB values are slightly worse than "GA" and "GA Prediction" but still much better than "Original data." The success rate equals that of "GA," indicating removal of numerical instability. The "GA Smooth" scheme estimates dispersion coefficients efficiently.

D. Correlation test

To analyze factors determining estimation accuracy, correlation tests were performed between FA5 values and other statistics. Results in Table 2 indicate: (1) FA5 has a positive relationship with the number of observation stations—more observations yield more accurate predictions; and (2) the number of computation steps is not related to FA5 because the GA method does not consider historical data.

VI. CONCLUSION

Traditional dispersion coefficients remain constant within the same atmospheric stability class. A GA scheme was established to dynamically correct dispersion coefficients in the Lagrangian puff model, minimizing errors. For large dispersion coefficient errors, numerical simulation results show that the GA scheme performs well. After applying the correction procedure to Kincaid experiment data, the dispersion model's forecasting capability improved significantly. Considering the continuity of dispersion coefficients, a Savitzky-Golay filter was used to smooth estimated parameters, increasing the estimation success rate by 5%. The GA method coupled with Savitzky-Golay filter can efficiently improve atmospheric dispersion simulation accuracy.

REFERENCES

[1] Davidson G. A modified Pasquill-Gifford power law dispersion coefficients. JAPCA J Air Waste Ma, 1990, 40: 1146–1147. DOI: 10.1080/10473289.1990.10466761

[2] Leroy C, Maro D, Hébert D, et al. A study of the atmospheric dispersion of a high release of krypton-85 above a complex coastal terrain, comparison with the predictions of Gaussian models (Briggs, Doury, ADMS4). J Environ Radioactiv, 2010, 101: 937–944. DOI: 10.1016/j.jenvrad.2010.06.011

[3] Xiang Y Y, Wang K, Zhang Y, et al. Radioactivity monitoring in environmental water and air around QNPP. Nucl Sci Tech, 2007, 18: 316–320. DOI: 10.1016/S1001-8042(07)60069-4

[4] Haupt S E. A demonstration of coupled receptor/dispersion modeling with a genetic algorithm. Atmos Environ, 2005, 39: 7181–7189. DOI: 10.1016/j.atmosenv.2005.08.027

[5] Jeong H J, Han M H, Hwang W T, et al. Application of data assimilation to improve the forecasting capability of an atmospheric dispersion model for a radioactive plume. Ann Nucl Energy, 2008, 35: 838–844. DOI: 10.1016/j.anucene.2007.09.008

[6] Jeong H J, Kim E H, Suh K S, et al. Determination of the source rate released into the environment from a nuclear power plant. Radiat Prot Dosim, 2005, 113: 308–313. DOI: 10.1093/rpd/nch460

[7] Allen C T, Young G S, Haupt S E. Improving pollutant source characterization by better estimating wind direction with a genetic algorithm. Atmos Environ, 2007, 41: 2283–2289. DOI: 10.1016/j.atmosenv.2006.11.007

[8] Korsakissok I and Mallet V. Comparative study of Gaussian dispersion formulas within the polyphemus platform: evaluation with Prairie Grass and Kincaid experiments. J Appl Meteorol Clim, 2009, 48: 2459–2473. DOI: 10.1175/2009JAMC2160.1

[9] Jung Y R, Park W G, Park O H. Pollution dispersion analysis using the puff model with numerical flow field data. Mech Res Commun, 2003, 30: 277–286. DOI: 10.1016/S0093-6413(03)00024-7

[10] Tan W J, Wang D Z, Ma Y W, et al. Diffusion coefficient adaptive correction in Lagrangian puff model. Atom Energy Sci Technol, 2014, 48: 571–576. DOI: 10.7538/yzk.2014.48.03.0571

[11] Ji Z L, Ma Y W, Wang D Z. Study on fitness functions of genetic algorithm for dynamically correcting nuclide atmospheric diffusion model. Nucl Tech, 2014, 37: 040503. (in Chinese) DOI: 10.11889/j.0253-3219.2014.hjs.37.040503

Note: Figure translations are in progress. See original paper for figures.