Source term Estimation of Key Radionuclides discharged from the Nuclear Power Plant into the Atmosphere based on Bayesian inference

Siho Jang [a], Juryong Park [a], Dakyoung Lee [a], Eung Soo Kim [a*]
[a]*Department of Nuclear Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, South Korea*
[*]*Corresponding author: kes7741@snu.ac.kr*

1. Introduction

In the case of a major nuclear accident, actual accident cases are extremely rare and data is limited. In this paper, we use data from the Shin-Kori nuclear power plant previously published in a paper to estimate 11 representative radionuclides as source terms. Various types of forward models, such as Gaussian puffs, can be used as atmospheric dispersion models, but we propose a source term estimation methodology based on the simplest yet powerful Gaussian plume model to reduce initial complexity [1].

Source term estimation is connected to solving an inverse problem, which requires the application of optimization methodology. Based on Bayesian inference, we have used Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ensemble Kalman Inversion (EKI) to find the state that minimizes the cost function [2, 3, 4]. Among these models, EKI has been selected as the most suitable means of estimating the source term, comparing computational simplicity, accuracy, and cost.

2. Methodology

We focus on Inverse modeling, which uses Bayesian inference to estimate the location and concentration of the source term [5, 6, 7]. The probabilistic estimation of the highest point is equivalent to finding the point where the cost function $\mathcal{J}$ is minimized, which is defined as follows:

$$\mathcal{J}(x) = \frac{1}{2}(x - x_a)^T S_a^{-1}(x - x_a) + \frac{1}{2}(y^{obs} - \mathcal{G}x_a)^T S_o^{-1}(y^{obs} - \mathcal{G}x_a) \quad (1)$$

where y^{obs} is the observation data, x and x_a are the state vector of the source term and the prior vector of the source term, S_a and S_o are the prior data error covariance, the observation data error covariance, respectively, and $\mathcal{G}$ denotes the forward operator mapping the state x to the observation y space, which is determined by the physics model.

To find the value of source term that minimizes the cost function, one can directly calculate the matrix for the given equation, solving $\nabla_x \mathcal{J}(x) = 0$.

$$x = x_a + S_a \mathcal{G}^T (\mathcal{G} S_a \mathcal{G}^T + S_o)^{-1}(y^{obs} - \mathcal{G}x_a) \quad (2)$$

We have selected the Gaussian plume model as the physics model and applied GA, PSO, and EKI methodologies to estimate the source term, and we will compare and analyze the performance of each result.

2.1 Gaussian Plume model

The Gaussian Plume model assumes a uniform meteorological field and normal distribution diffusion. The dispersion equation of the Gaussian Plume model is as follow [8]:

$$C(x,y,z) = \frac{Q}{2\pi\sigma_y\sigma_z u} exp\left(-\frac{y^2}{2\sigma_y^2}\right) \times \left[exp\left(-\frac{(z-H)^2}{2\sigma_z^2}\right) + exp\left(-\frac{(z+H)^2}{2\sigma_z^2}\right)\right] \quad (3)$$

where C denotes the concentration at a location separated by (x, y, z) from the source, Q is the emission rate of the pollutant from the source, u and H are the wind speed and the effective height of the plume, respectively, and σ_y and σ_z are the dispersion coefficients in y and z directions. The dispersion coefficient sets the diffusion coefficient in the above equation by utilizing the Pasquill-Gifford coefficients.

2.2 Genetic Algorithm

The key elements of GA include chromosome representation, selection, crossover, mutation, and fitness function computation. The optimization performance can vary depending on the specific application of each element [9]. The main algorithm flowchart is shown in Table I and in this paper, binary chromosomes, tournament selection, one-point crossover, and bit flip mutation are applied, with the least square method used as the fitness function.

Table I: GA Algorithm

Main Flow
(Input the data and initialization)
Set population size and maximum number of iterations
Generate initial population
Compute fitness value for each chromosome
While number of iterations is less than maximum iterations:
Select pair of chromosomes based on fitness
Apply crossover
Apply mutation

Replace old population with new population
Increment iteration counter

Return best solution

2.3 Particle Swarm Optimization

The fundamental principles of PSO can be explained using the following equation [10]:

$$v_{j,g}^{(t+1)} = w \times v_{j,g}^{(t)} + c_1 \times r_1 \times \left(pbes\ t_{j,g} - x_{j,g}^{(t)}\right) + c_2 \times r_2 \times \left(gbes\ t_{j,g} - x_{j,g}^{(t)}\right) \quad (4)$$

$$x_{j,g}^{(t+1)} = x_{j,g}^{(t)} + v_{j,g}^{(t+1)} \quad (5)$$

with $j = 1,2, \dots, n$ and $g = 1, 2, \dots, m$,
where n is the number of particles in the swarm, m is the number of components for the vector v_j and x_j, t is the number of iterations(generations), $v_{j,g}^{(t)}$ denotes the gth component of the velocity of particle j at iteration t ($v_g^{min} \le v_{j,g}^{(t)} \le v_g^{max}$), w is the inertia weight factor, r_1 and r_2 are random numbers uniformly distributed in the range $(0,1)$, and c_1 and c_2 are the acceleration coefficients. Also, $x_{j,g}^{(t)}$ denotes the gth component of the position of particle j at iteration t, and $pbes\ t_j$ and $gbes\ t_j$ are the personal best position of particle j and the global best position of the swarm, respectively.

The main algorithm flowchart of PSO is shown in Table II. We use typical numerical values for these coefficients, setting the minimum inertia weight factor to 0.75 and the acceleration coefficients c1 and c2 to 1.2.

Table II: PSO Algorithm

Main Flow
(Input the data and initialization)
Set swarm size and number of dimensions for each particle
Generate position and velocity randomly
Evaluate fitness
Set personal best and its fitness for each particle
Set global best and its fitness
Update personal best and global best if necessary for each particle
While number of iterations is less than maximum iterations:
Update velocity and position for each particle
Update personal best and global best if necessary
Increment iteration counter
Return global best solution

2.4 Ensemble Kalman Inversion

With EKI, the complicated $\mathcal{G}$ operator in Equation (2) are simplified by replacing them with covariance operations as follow:

The number of Ensemble $J \to \infty$:

$$C_n^{xg} \equiv \frac{1}{J}\sum_{j=1}^{J}\left(x_{n-1}^{(j)} - m_J\right)\left(\mathcal{G}x_{n-1}^{(j)} - \mathcal{G}m_J\right)^T \approx S_a\mathcal{G}^T \quad (6)$$

$$C_n^{gg} \equiv \frac{1}{J}\sum_{j=1}^{J}\left(\mathcal{G}x_{n-1}^{(j)} - \mathcal{G}m_J\right)\left(\mathcal{G}x_{n-1}^{(j)} - \mathcal{G}m_J\right)^T \approx \mathcal{G}S_a\mathcal{G}^T \quad (7)$$

with $m_J \equiv \frac{1}{J}\sum_{j=1}^{J} x_0^{(j)}$,
where n is the number of iterations.

Consequently, EKI simplifies the computation and solves the inverse problem by requiring only forward calculations. The main algorithm flowchart of EKI is shown in Table III [11]. We will use EKI to estimate the concentration of a Gaussian plume model through a forward model and Ensemble operations.

Table III: EKI Algorithm

Main Flow
(Input the state and initialization)
Set ensemble size and maximum number of iterations
Calculate the initial mean and covariance
While number of iterations is less than maximum iterations:
(Predict the state)
Select the ensemble randomly based on covariance
Estimate the predicted state
Recalculate the mean and covariance
(Update the state)
Transform the ensemble into observation space
Recalculate the mean in observation space
Get the Kalman gain
Update the state
Recalculate the mean and covariance
Increment iteration counter
Return the mean and covariance

3. Nuclear Severe Accident Applications

In the case of a severe accident, assuming the location of the occurrence is highly likely to be known from a macroscopic spatial perspective, we fix the source location and simulated multiple radionuclide occurrences while calculating the ability to estimate the source term with an increasing number of observation stations. And we expand the simulation up to a maximum of 11 radionuclides.

3.1 Source term Information

Given the lack of detailed information on the actual source term for a severe nuclear accident in a nuclear power plant, we use the source term information derived from the PSA Level 2. We utilize information on 11 representative radionuclides for three source term categories (STC-3) for the Shin-Kori nuclear power plant, as shown in Table IV [1]. Different isotopes can be distributed and included in each radionuclide group. However, in the current scope of research, it is limited to cover all radioactive isotopes, which total more than 60. Therefore, representative isotopes have been selected for each radionuclide group. Among the 11 radionuclides selected, excluding Mo-103 and Sb-129, most of the radionuclides were discussed in previous studies [12], [13]. In this study, to maintain the complexity of solution conditions, these two isotopes have been included without exclusion.

Table IV: STC-3 (Shin-Kori, Loss of offsite power)

Nuclide	Half-life	Radioactivity (Bq/s)
Kr-88	**2.84hr**	**1.90387731e+13**
Xe-133	**5.2475d**	**7.19994213e+13**
I-131	**8.0207d**	**1.04716435e+12**
Cs-137	**30.1671yr**	**8.67141204e+10**
Te-132	**3.204d**	**1.46342593e+12**
Sr-89	**50.57d**	**3.02453704e+10**
Mo-103	**65.94d**	**2.38444444e+11**
Ba-140	**12.752d**	**1.13539352e+10**
La-140	**1.6781d**	**3.26791667e+10**
Ce-144	**284.91d**	**2.34791667e+10**
Sb-129	**4.39d**	**4.00798611e+11**

3.2 Simulation Settings

The simulation domain has been set to be 10km in the x-direction, 1km in the y-direction, and 100m in the z-direction. Wind direction is assumed to be 1.5m/s in the x-direction, and Pasquill-Gifford coefficients are taken into account.

4. Results and Discussion

Performance evaluation of various optimization methodologies has been carried out based on the analysis of simulation results.

4.1 Single Source and Multiple Radionuclides

In order to investigate scenarios where multiple radionuclides can be released at a single source, we have conducted an analysis by increasing the number of radionuclides from one to eleven using the STC-3 information, as mentioned above, and analyzed the estimation of the source term depending on the number of observation stations. Unlike Case1, the location of the source is fixed, considering that in a severe accident, the approximate location of the accident could be known from a macroscopic perspective of meteorology.

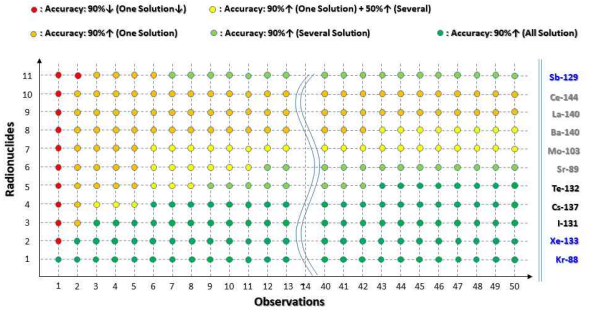


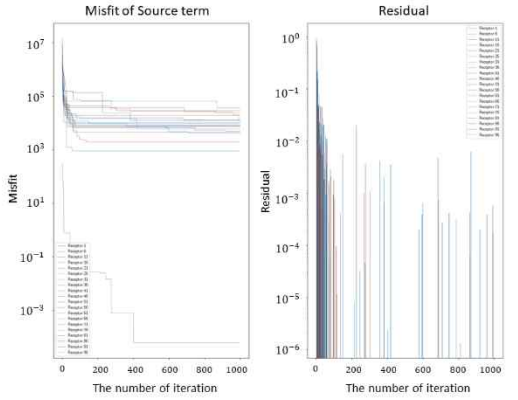
Fig. 1. Solvable Condition Diagram: This graph summarizes the results of analyzing the solvability of solutions while increasing the number of radionuclides from one Kr-88 to 11, including Sb-129, and the number of observations from one to 100. The solvability may vary depending on factors such as the half-life and concentration of each radionuclide, but this graph provides an overall trend for reference.

As shown in Fig. 1, our simulation results indicate that up to a certain number of radionuclides, a complete estimation of the source term is possible, but beyond that number, even with an increasing number of observation stations, a complete estimation may be difficult, and partial estimates are only possible for radionuclides with high concentration or short half-life. Moreover, we found that more observation stations than the number of desired solutions may be needed, and that even with many observation stations, the search for solutions could be limited. The use of elimination techniques can improve partial estimation, but this approach also incurs errors that arise from the eliminated uncertainties, and it may be possible to achieve complete estimation for cases where partial estimation was previously done.

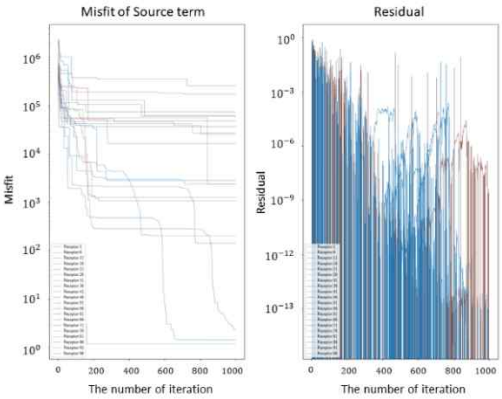
Based on the Solvable Condition Diagram derived above, a significant change in finding the solution is observed when there are 5 radionuclides. We will mainly focus on the analysis of 5 radionuclides in the following analysis, when the number of receptors increases at a condition of STC-3.

Differences in the specific performance of each method have been examined in more detail. First, regarding iteration(generation), the convergence speed of EKI has been found to be the fastest, followed by GA and PSO, as shown in Fig. 2. EKI required tens of iterations to converge to 1e-6, while GA required tens to hundreds of iterations, and PSO required hundreds to thousands of iterations. However, it should be noted that convergence speed of GA and PSO may be influenced by how the coefficients are set. While it is necessary to analyze in detail the convergence speed of GA and PSO as a function of changes in each coefficient, it is not within the scope of this paper. Based on the coefficients

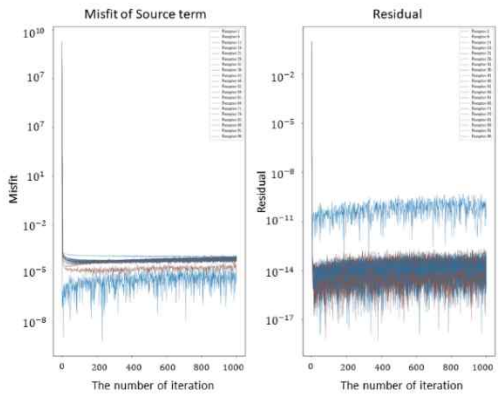
commonly used in the literature, we have selected the coefficients for the GA and PSO simulations.



(a)



(b)



(c)

Fig. 2. (a) Convergence analysis graph of GA as a function of generation (iteration), (b) Convergence analysis graph of PSO as a function of generation (iteration), (c) Convergence analysis graph of EKI as a function of iteration: The graphs in (a) and (b) may appear somewhat discontinuous, but this is because there are cases where the global best does not change even after the iteration, and therefore it should be noted that the y-axis is in log scale.

Next, regarding ensemble(population), as shown in Fig. 3, an increase in ensemble size has led to increased calculation stability. It also has resulted in a decrease in the number of iterations required.

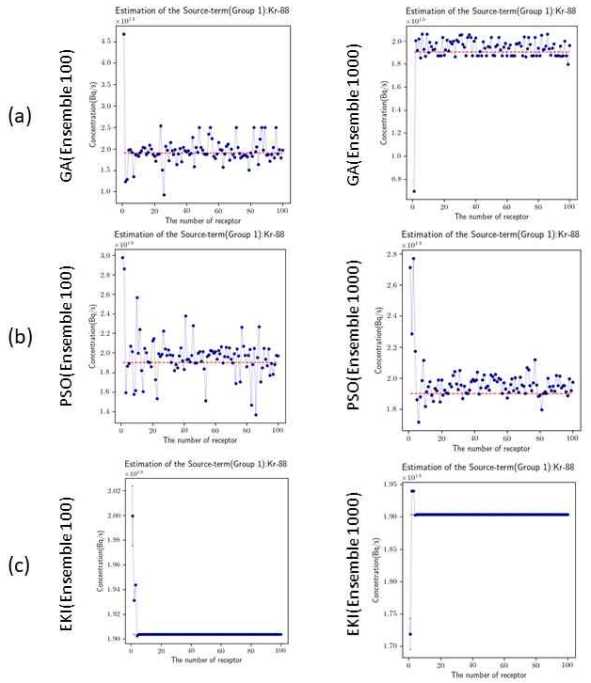
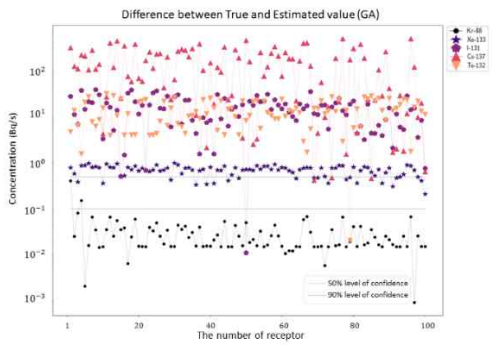
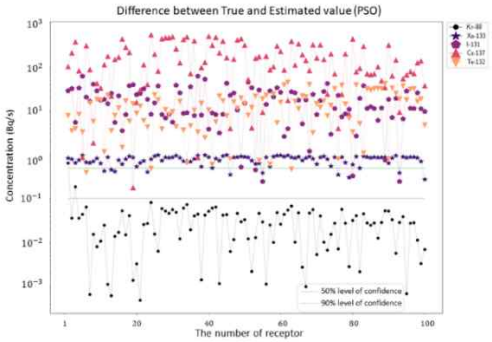


Fig. 3. (a) Changes in GA simulation results with increasing population size, (b) Changes in PSO simulation results with increasing population size, (c) Changes in EKI simulation results with increasing ensemble size: Overall, it can be seen that the stability of the calculations increases as the number of populations (ensembles) increases.

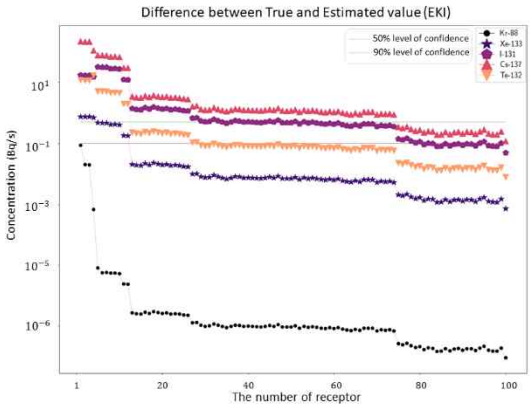
Regarding accuracy, GA shows an error rate of 1e-2, PSO shows 1e-3, and EKI show 1e-5, as shown in Fig. 4. Therefore, it can be seen that the estimation accuracy is in the order of EKI, PSO, and GA.



(a)



(b)



(c)

Fig. 4. (a) Error rate between estimated and true values using GA with identical initial boundary conditions, (b) Error rate between estimated and true values using PSO with identical initial boundary conditions, (c) Error rate between the estimated value using EKI and the true value, when the initial prior information is uniformly assigned to each radionuclide.

Overall, considering the convergence speed, accuracy, and computation time, EKI has been found to perform better than GA or PSO.

4.2 GPU Parallelization

We have modified the EKI code that shows the overall best performance to enable GPU parallel computation using the Cupy package provided by Python for more convenient GPU utilization. As a result, we confirmed that GPU parallelization can reduce computation time for large-scale operations. In particular, as shown in Fig. 5, when we plotted the product of Ensemble and Iteration that satisfied a specific convergence condition (Residual of Cost function reaching 10e-6) against the corresponding computation time, we have found that GPU computation is faster for a certain number of Ensemble and Iteration or more, while CPU computation is faster for less than that number. This suggests that the order of computation time increase due to parallelization is significantly lower and a difference of tens to hundreds of times in computation speed can occur as the scale gets larger.

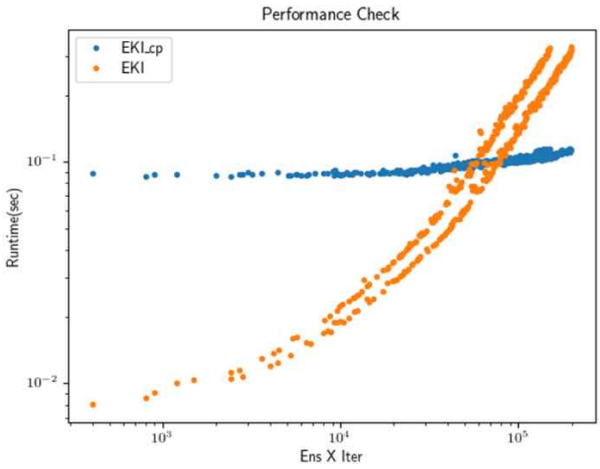


Fig. 5. Comparison of EKI computation using CPU and GPU: Considering that EKI converges within 10 iterations, GPU computation speed exceeds CPU computation speed in the range of tens of thousands to hundreds of thousands and in large-scale computations, there can be a difference of tens to hundreds of times.

5. Conclusions

Our study shows that the probability of finding a solution increases when the number of receptor points is greater than the number of sources, although this is not always true under specific conditions. Additionally, estimation of source locations is easier for radionuclides with higher radioactivity or shorter half-life.

These results can be used to inform the development of more accurate and efficient monitoring systems and to better understand the behavior of radioactive materials in the environment.

Acknowledgement

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety(KoFONS) using the financial resource granted by the Nuclear Safety and Security Commission(NSSC) of the Republic of Korea. (No. 2105010-0222-WT112)

REFERENCES

[1] Ju Young Jeon, Tae-Eun Kwon, and Jai-Ki Lee. "A Study on the Effect of Containment Filtered Venting System to Off-site under Severe Accident." Journal of Radiation Protection and Research 40.4, 244-251, 2015.
[2] Chang Liu, et al. "Gas diffusion model based on an improved Gaussian plume model for inverse calculations of the source strength." Journal of Loss Prevention in the Process Industries 75, 104677, 2022.
[3] X. L. Zhang, et al. "Modified ensemble Kalman filter for nuclear accident atmospheric dispersion: Prediction improved and source estimated." Journal of hazardous materials 280, 143-155, 2014.

[4] X. L. Zhang, et al. "Iterative ensemble Kalman filter for atmospheric dispersion in nuclear accidents: An application to Kincaid tracer experiment." Journal of hazardous materials 297, 329-339, 2015.
[5] Ondřej Tichý, et al. "Bayesian inverse modeling and source location of an unintended 131 I release in Europe in the fall of 2011." Atmospheric Chemistry and Physics 17.20, 12677-12696, 2017.
[6] Donald D. Lucas, et al. "Bayesian inverse modeling of the atmospheric transport and emissions of a controlled tracer release from a nuclear power plant." Atmospheric Chemistry and Physics 17.22, 13521-13543, 2017.
[7] Olivier Saunier, et al. "An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations." Atmospheric Chemistry and Physics 13.22, 11403-11421, 2013.
[8] Bo Cao, et al. "Development and uncertainty analysis of radionuclide atmospheric dispersion modeling codes based on Gaussian plume model." Energy 194, 116925, 2020.
[9] Sourabh Katoch, Sumit Singh Chauhan, and Vijay Kumar. "A review on genetic algorithm: past, present, and future." Multimedia Tools and Applications 80, 8091-8126, 2021.
[10] Dongshu Wang, Dapei Tan, and Lei Liu. "Particle swarm optimization algorithm: an overview." Soft computing 22, 387-408, 2018.
[11] Marco Iglesias, and Yuchen Yang. "Adaptive regularisation for ensemble Kalman inversion." Inverse Problems 37.2, 025008, 2021.
[12] Ling, Yongsheng, et al. "Multi-nuclide source term estimation method for severe nuclear accidents from sequential gamma dose rate based on a recurrent neural network." Journal of Hazardous Materials 414, 125546, 2021.
[13] Ling, Yongsheng, et al. "Inversion Method for Multiple Nuclide Source Terms in Nuclear Accidents Based on Deep Learning Fusion Model." Atmosphere 14.1, 148, 2023.