

## RESEARCH ARTICLE

# Robust pollution source parameter identification based on the artificial bee colony algorithm using a wireless sensor network

MengLi Cao <sup>\*</sup>, Xiong Hu

Logistics Engineering College, Shanghai Maritime University, Shanghai, China

<sup>\*</sup> [mlcao@shmtu.edu.cn](mailto:mlcao@shmtu.edu.cn) OPEN ACCESS

**Citation:** Cao M, Hu X (2020) Robust pollution source parameter identification based on the artificial bee colony algorithm using a wireless sensor network. PLoS ONE 15(5): e0232843. <https://doi.org/10.1371/journal.pone.0232843>

**Editor:** Jie Zhang, Newcastle University, UNITED KINGDOM

**Received:** January 13, 2020

**Accepted:** April 22, 2020

**Published:** May 15, 2020

**Copyright:** © 2020 Cao, Hu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All the simulation results are available from the Github database (<https://github.com/aoxfrank/SimulationData.git>).

**Funding:** This research was funded by the National Natural Science Foundation of China (<http://www.nsf.gov.cn/>) under No. 61801287. This research was also funded by Shanghai Universities' Young teachers' training support program of Shanghai Municipal Education Commission (<http://edu.sh.gov.cn/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Pollution source parameter identification (PSPI) is significant for pollution control, since it can provide important information and save a lot of time for subsequent pollution elimination works. For solving the PSPI problem, a large number of pollution sensor nodes can be rapidly deployed to cover a large area and form a wireless sensor network (WSN). Based on the measurements of WSN, least-squares estimation methods can solve the PSPI problem by searching for the solution that minimize the sum of squared measurement noises. They are independent of the measurement noise distribution, i.e., robust to the noise distribution. To search for the least-squares solution, population-based parallel search techniques usually can overcome the premature convergence problem, which can stagnate the single-point search algorithm. In this paper, we adapt the relatively newly presented artificial bee colony (ABC) algorithm to solve the WSN-based PSPI problem and verifies its feasibility and robustness. Extensive simulation results show that the ABC and the particle swarm optimization (PSO) algorithm obtained similar identification results in the same simulation scenario. Moreover, the ABC and the PSO achieved much better performance than a traditionally used single-point search algorithm, i.e., the trust-region reflective algorithm.

## 1. Introduction

Hazardous pollution can be released to the atmosphere and cause severe disease to human beings [1]. At the early stage of the releasing process, the released pollution forms a plume near the pollution source. A timely perception of the released pollution and identification of its source parameters can save a lot of time for the subsequent elimination of the pollution source. Feasible solutions to this problem include deploying an array of stationary pollution sensor nodes [2] or dispatching mobile robots [3, 4] to the concerned area. The deployed sensor nodes can construct a WSN in a self-organized manner, which is capable of rapidly covering a large area.

The problem of PSPI based on a WSN has been extensively studied. With respect to the underlying theory, the WSN-based PSPI methods can be categorized to two groups:

**Competing interests:** The authors have declared that no competing interests exist.

probabilistic methods and nonlinear least-squares (NLS) estimation methods. Probabilistic methods [5–7] output the estimated value with the highest probability of being the real value. Nehorai *et al.* [5] proposed a maximum likelihood estimation method to identify the source location, diffusion coefficient, and start time in a diffusion model. Keats *et al.* [6] proposed a maximum posterior probability method based on the Bayesian inference. Nofsinger *et al.* [7] proposed to implement the Ensemble Kalman Filtering algorithm to estimate the source location of a pollution plume. However, probabilistic methods usually assume that the measurement noises satisfy a specific distribution (e.g., the Gaussian distribution), which do not always hold in WSN-based PSPI applications.

Traditional NLS methods usually use single-point search algorithms to seek the solution that minimize the sum of squared measurement noises. Michaelides *et al.* [8] used the trust-region-reflective [9] (TRR) method to solve the NLS formulation of the WSN-based PSPI problem. Chen [10] proposed a  $l_2$ -regularized NLS formulation of the WSN-based PSPI problem, and then solved it with a greedy method. The NLS formulation [11] does not require any statistical assumptions about the measurement noises, and thus, is robust to non-Gaussian measurement noises [12]. Nevertheless, when used for solving the NLS formulation of the WSN-based PSPI problem, traditional single-point search methods are prone to be stagnated by the premature convergence problem, i.e., converge prematurely at solutions far from the true-values of the parameters. This problem can be avoided by using the population-based parallel search approaches, e.g., the particle swarm optimization (PSO) algorithm [13]. As a relatively new population-based parallel search algorithm, the artificial bee colony (ABC) algorithm presented by Karaboga *et al.* [14] has been widely applied in various applications. A comprehensive survey of the ABC and its applications can be found in [15]. Compared with other population-based algorithms, the ABC can achieve better or similar performance with the advantage of employing fewer control parameters [16]. Moreover, the ABC has been found to be able to produce very good results at a low computational cost [15], which is significant for saving time in the PSPI processes. In the area of WSN, the ABC has been implemented for node localization [17], node clustering [18], network routing [19], coverage optimization [20], and etc. However, to our knowledge, the ABC has not been adapted for solving the WSN-based PSPI problem yet.

In this paper, we adapt the ABC to solve the problem of WSN-based PSPI. To identify the real value of the unknown parameters, i.e., the source location, release rate, and the diffusion coefficient, an objective function of them is formed based on the NLS paradigm. Then, the ABC is used to search for the global minimum of the objective function. To evaluate the ABC, it was compared with the PSO [21], the Levenberg-Marquardt (LM) [22], and the TRR in extensive simulations. Simulation results show that the ABC and the PSO achieved similar estimation results in the same simulation scenario. Moreover, they both can achieve much better performance than traditional single-point search methods. This paper includes at least three contributions: (1) presents a method for adapting the ABC to solve the WSN-based PSPI problem, and verifies its feasibility and robustness; (2) verifies that the ABC outperforms the traditionally used single-point search methods; (3) reveals the influence of algorithm parameters on the identification performance of all tested algorithms.

The rest of this paper is organized as follows. Section 2 describes the pollution distribution model and concentration measurements of the WSN, and the scheme of all simulations. Section 3 details the ABC, PSO, LM, and TRR algorithms, Section 4 presents the simulation results and corresponding discussions. Section 5 concludes the whole paper and gives some future works.

## 2. Simulation setup

### 2.1 Pollution distribution and measurement

Within an 80-by-80 m<sup>2</sup> flat square area, we simulate a pollution source, which is releasing hazardous materials at a rate of  $q$  mL/s on location  $(x_s, y_s)$ , i.e., the release rate is  $q$  mL/s. The released materials are carried by an isotropic wind at a speed of  $v$  cm/s, forming a time-averaged pollution distribution. At the meantime,  $N$  sensor nodes are randomly deployed in the area at locations  $(x_i, y_i)$ ,  $i = 1, \dots, N$ . The sensor nodes are capable of forming a WSN in a self-organized manner, and measuring its surrounding pollution concentration, which is denoted as  $z_i$ ,  $i = 1, \dots, N$ .

In real PSPI applications, the concentration measurements of the sensor nodes (i.e.,  $z_i$ ,  $i = 1, \dots, N$ ) are available. However, in computer simulations, the value of  $z_i$ ,  $i = 1, \dots, N$  should be obtained by adding an measurement error term  $e_i$ ,  $i = 1, \dots, N$  to the theoretical concentration  $c_i$ ,  $i = 1, \dots, N$ , as follows.

$$z_i = c_i + e_i, i = 1, \dots, N, \tag{1}$$

where  $c_i$ ,  $i = 1, \dots, N$  can be calculated by substituting  $(x_i, y_i)$ ,  $i = 1, \dots, N$  into a pollution concentration distribution model. Considering the fact that real pollution sensors are constrained by the detection range, the  $i$ -th sensor node is hibernated if  $z_i < lb$  or  $z_i > ub$ , where  $lb$  and  $ub$  are the lower and upper bounds of the sensor detection range, respectively.

To calculate  $c_i$ ,  $i = 1, \dots, N$ , we adopt the advection-diffusion model presented in [23], which has been widely used to describe the time-averaged concentration distribution in environments with homogenous airflow [24–27]. It takes the form

$$c_i(x_s, y_s, q, K) = \frac{q}{2\pi K d_i} \exp\left[-\frac{v(d_i - x_i + x_s)}{2K}\right], \tag{2}$$

where  $K$  is the turbulent diffusion coefficient,  $d_i$  is the distance between  $(x_i, y_i)$ ,  $i = 1, \dots, N$  and  $(x_s, y_s)$ . In our concerned WSN-based PSPI problem,  $(x_s, y_s)$ ,  $q$ , and  $K$  in Eq (2) are parameters to be identified, since their real values are previously unknown in the real PSPI process. However, to simulate the measurement process and generate the concentration measurements, the real values of  $(x_s, y_s)$ ,  $q$ , and  $K$  should be substitute into Eq (2) to calculate  $c_i$ ,  $i = 1, \dots, N$ , and in turn calculate  $z_i$ ,  $i = 1, \dots, N$  based on Eq (1).

### 2.2 Simulation scheme

To solve the PSPI problem, we follow the NLS estimation paradigm [11] to simultaneously identify the value of  $(x_s, y_s)$ ,  $q$ , and  $K$ . For ease of expression, the four unknown parameters  $(x_s, y_s)$ ,  $q$ , and  $K$  are integrated in a four-dimensional vector  $p = [x_s, y_s, q, K]$ . Then, the NLS estimation problem [11] can be represented as searching the minimum of

$$obj(p) = \sum_{i=1}^N [z_i - c_i(p)]^2. \tag{3}$$

By considering the value of  $fit(p) = 1/obj(p)$  as the fitness of  $p$ , the NLS estimation problem becomes an optimization problem. According the Eq (3), the value of  $obj(p)$  is positive. Thus, a solution that maximize the reciprocal of  $obj(p)$  (i.e.,  $fit(p) = 1/obj(p)$ ) can minimize  $obj(p)$ .

In order to solve the PSPI problem, the ABC is used to maximize the value of  $fit(p) = 1/obj(p)$ . The influence of ABC's tunable parameters on the estimation performance is assessed. Then, to comprehensively evaluate the ABC, it was compared with three counterpart algorithms: the PSO, LM, and TRR algorithms, in extensive simulations.

Last but not the least, the robustness in respect of non-Gaussian measurement noises were evaluated. Typically, the measurement noises satisfy Gaussian distribution, whereas measurement noises satisfying non-Gaussian distribution [12] also exist in many real PSPI applications. Thus, to test and compare their robustness in respect of non-Gaussian measurement noises, the ABC and the PSO were conducted based on a series of concentration measurements with contaminated Gaussian noises, which can be represented as a convex combination of a Gaussian noise and a uniform noise, as follows [12].

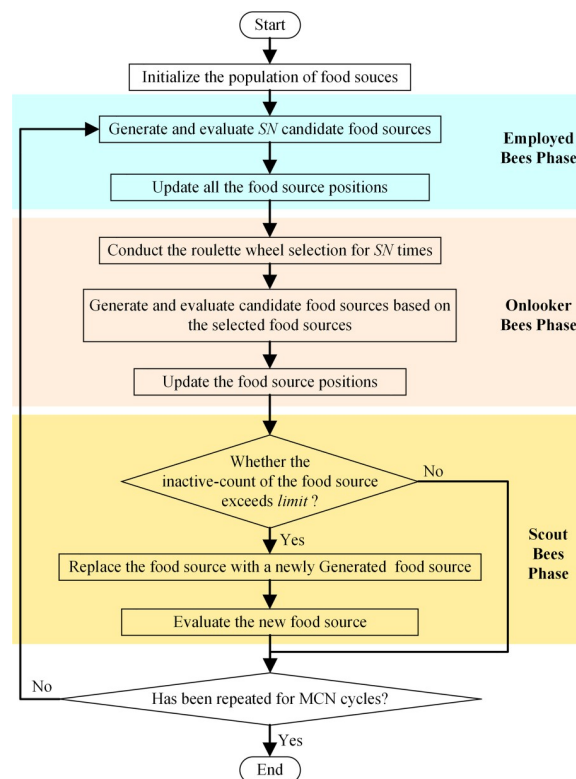
$$e_i \sim \lambda N(0, 10) + (1 - \lambda)U(0, 10), \lambda \in [0, 1]. \quad (4)$$

### 3. Search algorithms

#### 3.1 Artificial bee colony

The inner structure of the ABC [15] is shown in Fig 1.

In the ABC, a solution and its corresponding fitness value are called a food source and its corresponding nectar volume, respectively. For the optimization problem stated in section 2, a food source takes the form of the vector  $\mathbf{p}$  in Eq (3). The nectar volume of a food source is calculated by substituting  $\mathbf{p}$  into Eq (3). At the beginning, the ABC enters the initialization phase, a population of  $N$  food sources are randomly generated within the valid value ranges. Then, the ABC algorithm enters a cyclic state. Each cycle consists of the employed bees phase, the onlooker bees phase, the scout bees phase, and a phase for updating the best solution achieved so far.



**Fig 1. The inner logic flowchart of the ABC.**

<https://doi.org/10.1371/journal.pone.0232843.g001>

In the employed bees phase, a new candidate food source  $\mathbf{p}'_i$  is generated by varying a randomly selected dimension of each food source in the population, as follows.

$$\begin{aligned} \mathbf{p}'_{ij} &= \mathbf{p}_{ij} + \phi_{ij}(\mathbf{p}_{ij} - \mathbf{p}_{kj}) \\ i &= 1, 2, \dots, N, j \in \{1, 2, 3, 4\}, \end{aligned} \quad (5)$$

where  $\phi_{ij}$  is a uniform random number in the range  $[-1, 1]$ ,  $k$  is randomly selected from the set  $[1, i - 1] \cup [i + 1, N]$ ,  $\mathbf{p}_{ij}$  ( $\mathbf{p}'_{ij}$  or  $\mathbf{p}_{kj}$ ) denotes the  $j$ -th element of  $\mathbf{p}_i$  ( $\mathbf{p}'_i$  or  $\mathbf{p}_k$ ). Then, the nectar volume of  $\mathbf{p}'_{ij}$  is evaluated.  $\mathbf{p}_i$  will be replaced by  $\mathbf{p}'_i$  in case that  $obj(\mathbf{p}'_i) < obj(\mathbf{p}_i)$ . Otherwise, the inactive-count of  $\mathbf{p}_i$ , which is denoted as  $l_i$ , will be self-increased to  $l_i + 1$ . The operations in the employed bees phase is inspired by the phenomenon that employed bees search for new food sources containing more nectar within the neighborhood of the food source in the memory [15].

In the onlooker bees phase, for each of the onlooker bees, a roulette wheel selection operation is conducted to select within the ABC population. Then, for the selected food source, a new candidate food source is generated and evaluated is conducted in the way like that in the employed bees phase. The operations in the onlooker bees phase is inspired by the phenomenon that onlooker bees probabilistically choose their food sources depending on the information provided by the employed bees [15].

In the scout bees phase, if the inactive-count of a food source exceeds a pre-defined “limit”, it will be replaced by a randomly generated food source. This operation is inspired by the phenomenon that scout bees randomly search for new food sources to replace the initially poor or fully exploited food source [15].

Finally, as the best solution achieved so far, the food source with the biggest fitness value in the population is updated.

### 3.2 Particle swarm optimization

The general structure of the PSO [21] is shown in Fig 2.

For using the PSO to minimize  $obj(\mathbf{p})$  in Eq (3), the particle (i.e., the solution in the PSO) takes the form of  $\mathbf{p}$ . After the PSO solutions are randomly initialized within the valid value ranges, all of the particles are cyclically updated.

In each cycle, the fitness of all the particles are calculated. Then, the historically best solutions (i.e., the particle-best solution  $\mathbf{pb}_i$ ,  $i = 1, 2, \dots, N$  and the global-best solution  $\mathbf{gb}$ ) achieved by the particle and the population are updated, respectively. Then, the  $i$ -th particle is updated

Initialization a population of particles

**Repeat**

Calculate the fitness of each particle in the population

Update the best solution achieved by the particle so far  
(i.e., the particle-best solution)

Update the best solution achieved by the population so far  
(i.e., the global-best solution)

Calculate the velocity of each particle in the population

Update the particles in the population

**Until** cycle == the Maximum Cycle Number (i.e.,  $MCN$ ).

**Fig 2. General structure of the PSO.**

<https://doi.org/10.1371/journal.pone.0232843.g002>

as follow.

$$\begin{aligned}
 \mathbf{v}_{ij}^k &= \omega \mathbf{v}_{ij}^{k-1} + \varphi_1 r_1 (\mathbf{p} \mathbf{b}_{ij}^{k-1} - \mathbf{p}_{ij}^{k-1}) + \varphi_2 r_2 (\mathbf{g} \mathbf{b}^{k-1} - \mathbf{p}_{ij}^{k-1}) \\
 \mathbf{p}_{ij}^k &= \mathbf{p}_{ij}^{k-1} + \mathbf{v}_{ij}^k \\
 i &= 1, 2, \dots, N, \quad j = 1, 2, 3, 4
 \end{aligned} \tag{6}$$

where  $\mathbf{v}_i$  is the velocity of the  $i$ -th particle,  $N$  is the number of particles,  $k$  denotes the cycle number, the second subscript  $j$  is the dimension index (i.e.,  $v_{ij}^k$  is the  $j$ -th dimension of  $\mathbf{v}_i$  at the  $k$ -th cycle),  $\omega$  is a nonnegative inertial factor,  $r_1$  and  $r_2$  are two random number within the range  $(0, 1)$ ,  $\varphi_1$  and  $\varphi_2$  are acceleration constants. [21]

### 3.3 Single-point search algorithms

The LM method and the TRR method are two single-point search methods that are traditionally used for solving the NLS problem. Both methods have the common point that only one single solution is cyclically evaluated and improved.

For the NLS problem of minimizing  $\|F(\mathbf{p})\|_2^2$  by varying the value of  $\mathbf{p}$ , the search direction of the LM method at the  $k$ -th cycle, denoted by  $\mathbf{d}^k$ , satisfies [22]

$$(J^T(\mathbf{p}^k)J(\mathbf{p}^k) + \mu^k I)\mathbf{d}^k = -J^T(\mathbf{p}^k)F(\mathbf{p}^k), \tag{7}$$

where  $J(\mathbf{p})$  is the *Jacobian* matrix of  $F(\mathbf{p})$ ,  $\mu^k$  is a constant coefficient. It is readily seen that, for the NLS problem concerned by this paper,  $F(\mathbf{p}) = z - c(\mathbf{p})$ .

The TRR method approximates the function  $obj(\mathbf{p})$  to be optimized by a much simpler function  $q(\mathbf{p})$  in the neighborhood of the current solution (i.e., the trusted region TR), and then solve the minimization problem [9]

$$\min_p \{q(\mathbf{p}), \mathbf{p} \in TR\}. \tag{8}$$

To determine the trusted region, the TRR method employs a preconditioned conjugate gradient process [28].

## 4. Results and discussion

As mentioned in section 2.2, a large number of simulations were conducted to tune the parameters of the ABC, PSO, LM, TRR methods, and then to compare their performance on solving the WSN-based PSPI problem.

### 4.1 Control parameter selection

A typical simulation scenario is generated to select the value of control parameters, which can influence the estimation errors. The real value of  $\mathbf{p}$  is set as [1000 cm, 4000 cm, 8000 mL/s, 5000 cm<sup>2</sup>/s]. To generate the initial populations for the ABC and the PSO, the valid value ranges were empirically set as follows. The lower bound of  $x_s, y_s, q, K$  were set as 0 cm, 0 cm, 800 mL/s, 4000 cm<sup>2</sup>/s, respectively. The upper bound of  $x_s, y_s, q, K$  were set as 8000 cm, 8000 cm, 16000 mL/s, 6000 cm<sup>2</sup>/s, respectively.

The wind velocity, which can be measured in real applications, was set as 3 m/s. In order to keep enough sensor nodes at work, the value of  $lb$ , and  $ub$  were set as 50 ppm and 1000 ppm, respectively. The concentration measurement noises were sampled from the Gaussian distribution  $N(0,10)$ .

**4.1.1 The ABC algorithm.** In the ABC, there are three control parameters: the number of food sources  $SN$ ,  $limit$  in the scout bees phase, and the maximum cycle number ( $MCN$ ).  $SN$  is



also the population size of the ABC. Empirically, like other population-based parallel-search methods, larger population size means higher search efficiency, while larger value of *MCN* means longer and more sufficient search. However, the increasing of *SN* and *MCN* both bring about computational cost. The computational complexity of the ABC, which is not the main concern of this paper, has been studied in [15]. Therefore, considering the computational resource of our experimental computer (i.e., Intel i7-7500U CPU and 8GB RAM), the value of *SN* and *MCN* was set as 100 and 500, respectively.

To evaluate the influence of *limit* in our problem, the WSN was randomly deployed for 100 times. For each of the random node deployment schemes, the value of *limit* was set from 200 to 1600 in interval of 200. Boxplots of the estimation errors with respect to different *limit* values are shown in Fig 3. According to Fig 3, the distinctions between the results obtained with different *limit* values are minor. In all the six groups of trials, the medians of the estimation errors are close to zero. It seems that, for solving the WSN-based PSPI problem, the scout bees phase can hardly influence the search ability of the ABC. This is probably because the exploration strengths in the employed bees phase and the onlooker bees phase of the ABC are strong enough to maintain the ABC population diversity in our problem.

**4.1.2 Counterpart algorithms.** The PSO, LM, and TRR algorithms are selected as the counterpart algorithms, which can also be used to solve the WSN-based PSPI problem concerned in this paper, to be compared with the ABC. It is necessary to tune the parameters of the counterpart algorithms too.

For rigorous comparison, the population size and *MCN* of the PSO were set as the same as those selected for the ABC, i.e.,  $N = 100$ ,  $MCN = 500$ . Moreover, there are three control parameters: the acceleration constants  $\varphi_1$  and  $\varphi_2$ , and the inertial factor  $\omega$ . Empirically, setting the value of the acceleration constants to 2, i.e.,  $\varphi_1 = \varphi_2 = 2$ , can achieve good performance [21]. Thus, only the value of inertial factor  $\omega$  was tuned. To evaluate the influence of inertial factor  $\omega$  in our problem, the WSN was randomly deployed for 100 times. For each of the random node deployment schemes, the value of  $\omega$  was set from 0.175 to 0.875 in interval of 0.175. Boxplot of the estimation errors in the five groups are shown in Fig 4.

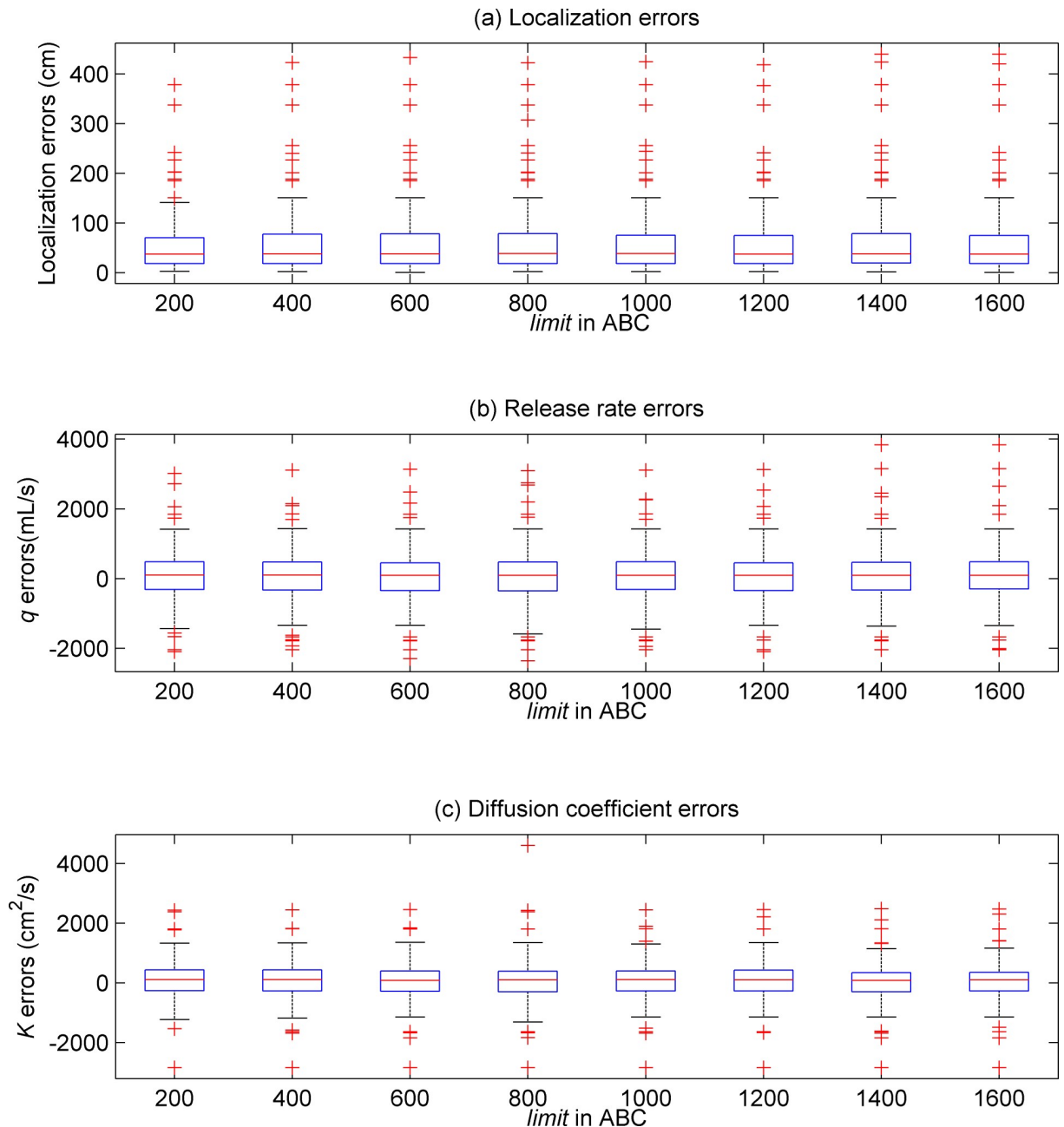
As shown in Fig 4, when the value of  $\omega$  equals 0.7, the *IQR* of the boxplot is smallest, and that the median of the boxplot is closest to zero. Thus, it can be concluded that the PSO obtained generally smaller estimation errors than in other cases. The most possible reason is that the small value of  $\omega$  brings about small diversity of the particles.

For the two typical single-point search algorithms, i.e., the LM and the TRR algorithms, the influence of start point on the estimation performance was evaluated. Compared with the real value to be estimated, i.e., [1000 cm, 4000 cm, 8000 mL/s, 5000 cm<sup>2</sup>/s], the four start points in Table 1 are obviously different. The fourth start point is closest to the real values, whereas the first one has the biggest gap with the real values.

Each of the four start points shown in Table 1 were tested for 200 times. The boxplots of the corresponding estimation errors are shown in Fig 5. In each subfigure of Fig 5, the left four and right four boxplots correspond to TRR and LM, respectively.

On one hand, the estimation errors in Fig 5 reveal that the estimation performance of LM and TRR were both obviously influenced by the start point. By utilizing the fourth start point, the LM and TRR both can achieve the smallest estimation errors. Possible reason is that, among all the start points, the fourth start point is closest to the real values.

On the other hand, it can be conceptually considered that the TRR is more likely to obtain small estimation errors than the LM algorithm. The TRR obtained a lot of small estimation errors. By contrast, the LM converged far from zero in most trails, although the results of the LM are more aggregated than those of the TRR.



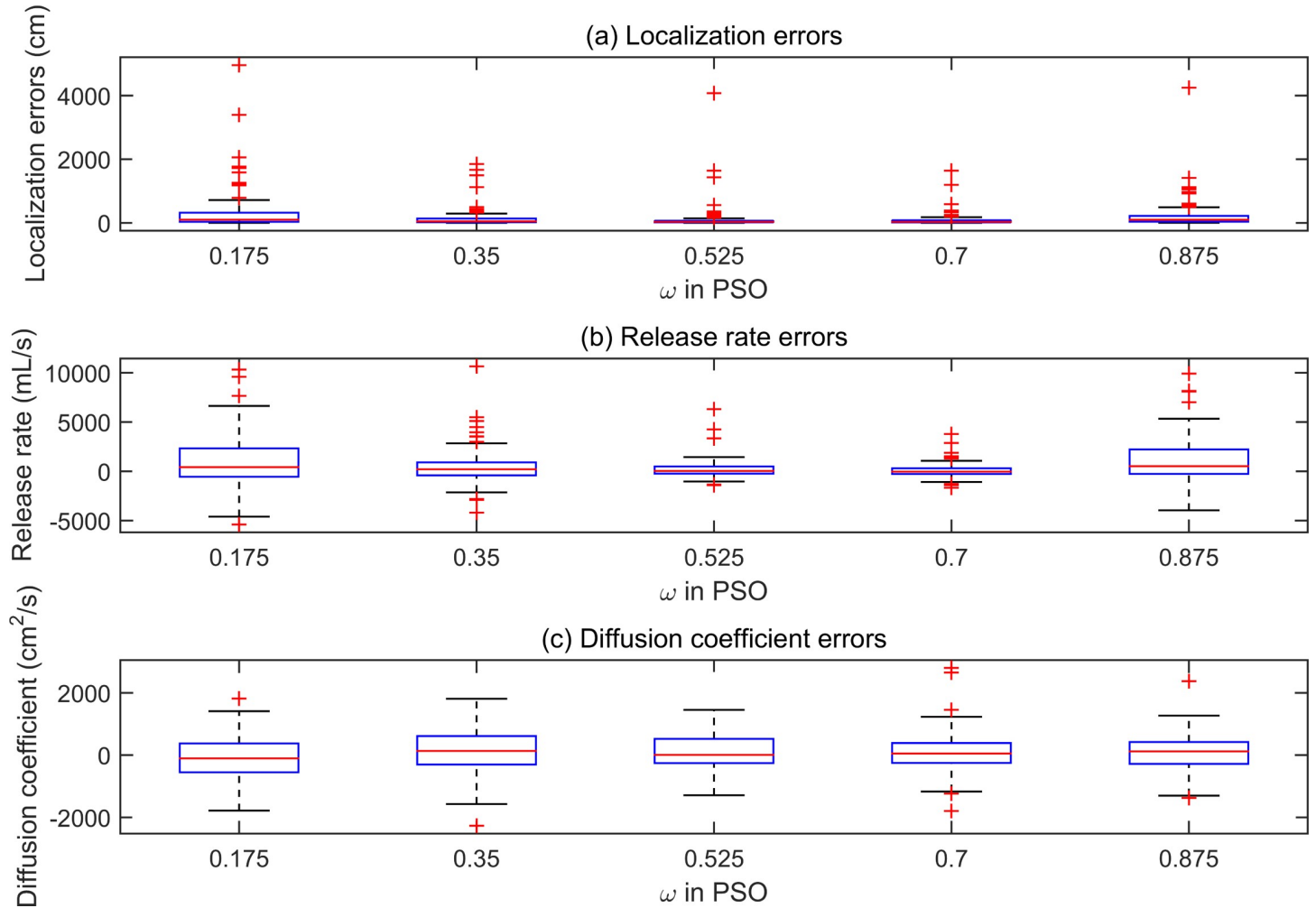
**Fig 3. Boxplots of the estimation errors in eight groups of the simulation trials with different values of "limit" in the ABC.**

<https://doi.org/10.1371/journal.pone.0232843.g003>

## 4.2 Performance comparison

**4.2.1 PSPI with Gaussian measurement noises.** On the basis of the parameter selection results in section 4.1, the WSN was randomly deployed for 100 times to compare the ABC with the PSO, TRR methods. For each of the random node deployment schemes, each of the ABC, PSO, TRR methods was tried once. For both ABC and the PSO, the population size and MCN were set as 100 and 500, respectively. The "limit" value of ABC was set as 1000. The inertial factor of PSO was set as 0.7. For the TRR, the start point was set as the fourth start point in





**Fig 4. Boxplots of the estimation errors in five groups of simulation trials with different values of inertial factor in the PSO.**

<https://doi.org/10.1371/journal.pone.0232843.g004>

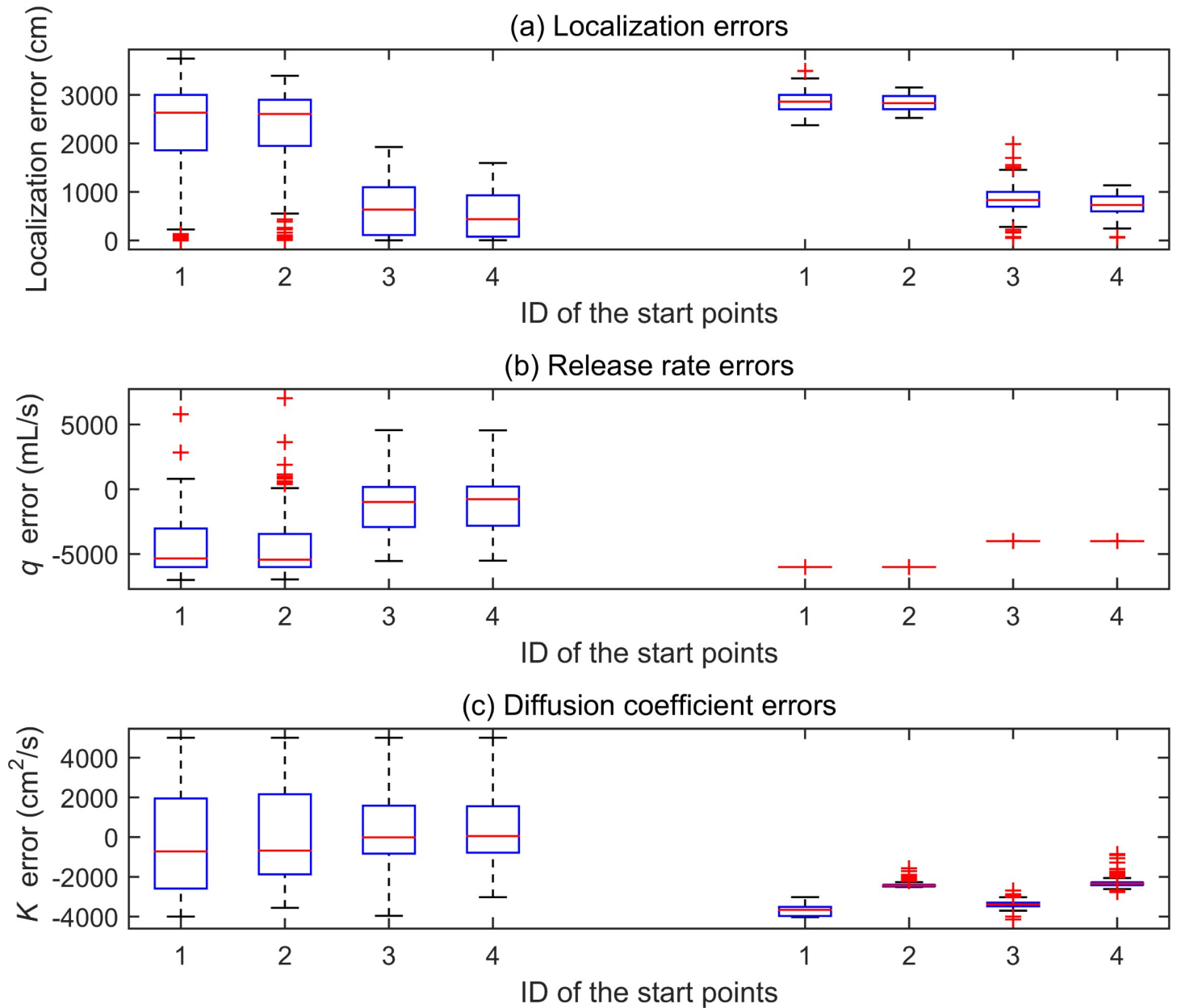
**Table 1.** Other simulation setups are the same as those in section 4.1. The boxplots of the estimation errors are shown in Fig 6.

From Fig 6, we can see that the ABC and the PSO achieved similar estimation errors. This similarity is mainly because of the same measurement scheme and the similar search ability of the ABC and the PSO. Moreover, the IQR and estimation biases obtained by the TRR algorithm are much bigger than those of the ABC and the PSO. Note that the start point for the TRR algorithm has been intentionally set within the neighborhood of the real values (i.e., the fourth start point in Table 1), which is not applicable in real PSPI problems.

**Table 1. Start points for LM and TRR.**

ID	$x_s$ (cm)	$y_s$ (cm)	$q$ (mL/s)	$K$ (cm <sup>2</sup> /s)
1	4000	4000	2000	1000
2	4000	4000	2000	2500
3	2000	4000	4000	1000
4	2000	4000	4000	2500

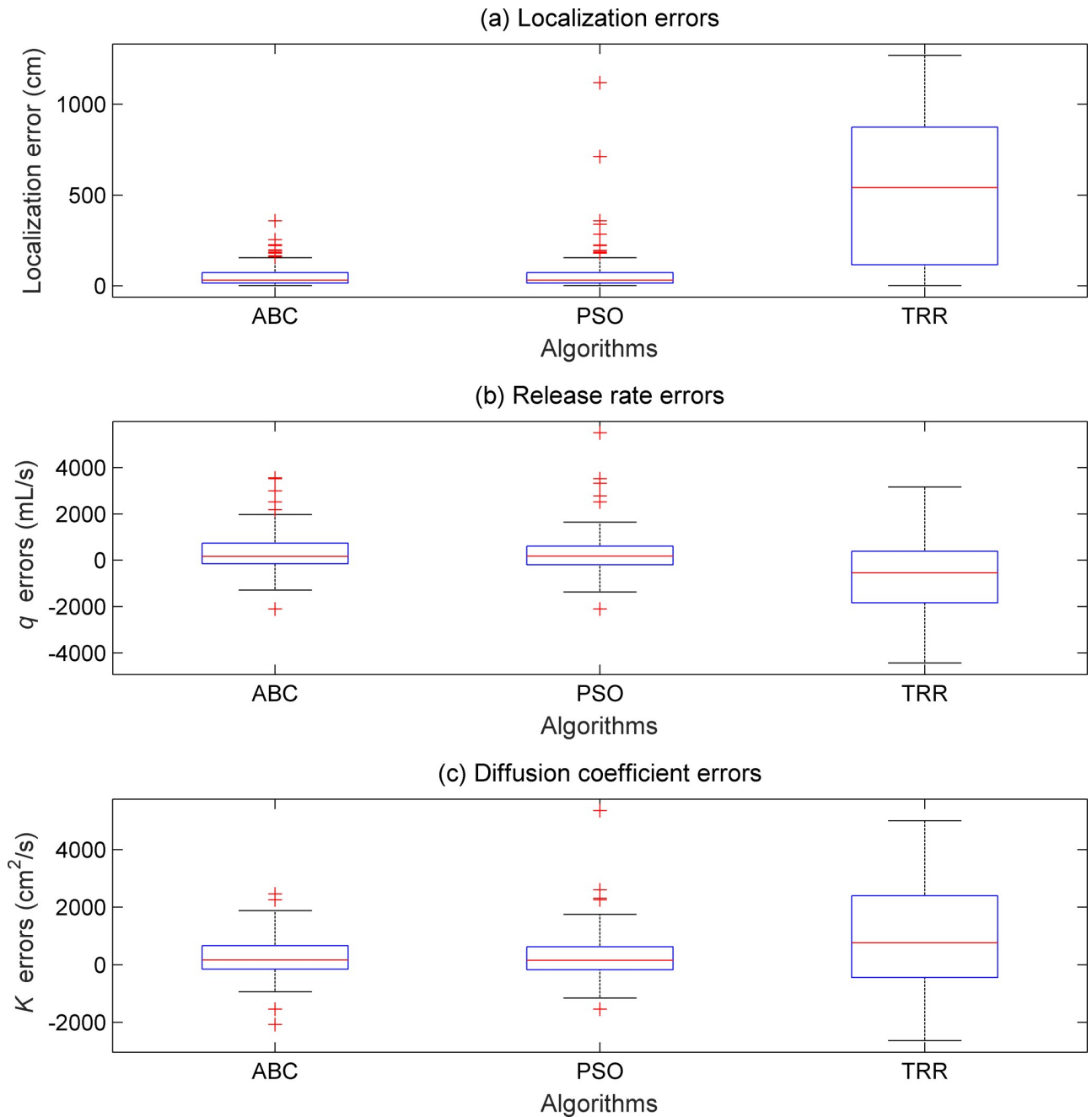
<https://doi.org/10.1371/journal.pone.0232843.t001>



**Fig 5. Boxplots of the estimation errors obtained by TRR and LM with respect to four different initial points.** In each subfigure, the left four boxplots correspond with TRR, whereas the others correspond with LM.

<https://doi.org/10.1371/journal.pone.0232843.g005>

**4.2.2 PSPI with non-Gaussian measurement noises.** The above simulations are based on the assumption that the concentration measurement noises of the WSN satisfy the Gaussian distribution. However, this assumption is not valid in many applications with external interference factors. Thus, to test the algorithms' robustness in respect of non-Gaussian measurement noises, measurement noises sampled from the contaminated Gaussian distribution function in Eq (4) were added into the concentration measurements of the WSN. The value of  $\lambda$  in Eq (4) were varied from 1 to 0 in intervals of 0.25. For each value of  $\lambda$ , the ABC and the PSO were both tried for 200 times. The boxplots of these simulations are shown in Fig 7. For better

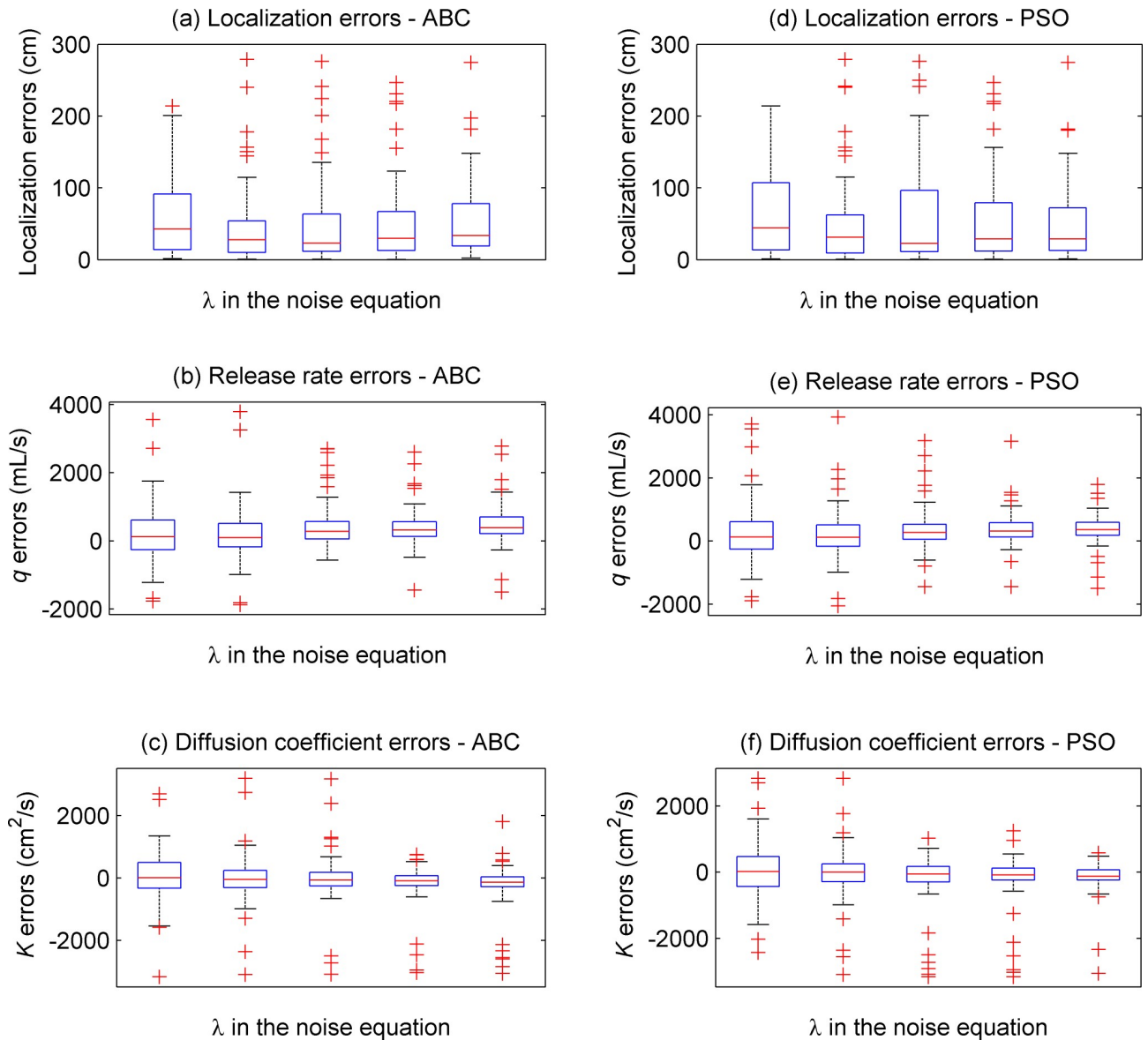


**Fig 6. Boxplots of the estimation errors obtained by the ABC, PSO, and TRR.**

<https://doi.org/10.1371/journal.pone.0232843.g006>

illustration, the display data ranges of Fig 7(A) and 7(D) were restrained as (0, 300) cm, which means a few outliers are not displayed in Fig 7(A) and 7(D).

As shown in the left four boxplots, the ABC obtained similar estimation errors when the value of  $\lambda$  varied. These estimation errors are generally close to zero. According to Eq (4), a smaller value of  $\lambda$  means that the Gaussian component takes less effect in the measurement noises. Therefore, the estimation performance obtained by the ABC does not rely on the distribution of measurement noise. In other words, the ABC is robust when it is used for solving the



**Fig 7. Boxplots of the estimation errors obtained by the PSO and ABC based on non-Gaussian measurement noises.** In each subfigure, the left four boxplots correspond with the ABC, whereas others correspond with the PSO.

<https://doi.org/10.1371/journal.pone.0232843.g007>

WSN-based PSPI problem. Moreover, the right four boxplots are similar as the left ones, which reveals that the PSO also achieved similar estimation results as those obtained by the ABC. This similarity corresponds with that demonstrated in Fig 6.

### 5. Conclusions

The ABC has been adapted for solving the WSN-based PSPI problem, in which the parameter values of a previously unknown pollution source should be identified based on the measurements of a wireless sensor network. According to the simulation results, the location, release rate, and diffusion coefficient of the pollution source can be simultaneously identified with considerable accuracy by using the ABC. Moreover, for solving the WSN-based PSPI problem,

the ABC and the PSO obtained similar identification results in the same simulation scenario, although the ABC employs less tunable parameters than the PSO. As two intelligent optimization algorithms, the ABC and the PSO both obtained generally more accurate results than the TRR, which is a typical single-point search algorithm. The above conclusions are based on the result of computer simulations. Future works may cover collecting concentration measurements with a real WSN, and using them to evaluate the performance of the ABC and its counterpart algorithms.

## Author Contributions

**Funding acquisition:** MengLi Cao.

**Investigation:** MengLi Cao.

**Methodology:** MengLi Cao.

**Resources:** Xiong Hu.

**Supervision:** Xiong Hu.

**Writing – original draft:** MengLi Cao.

**Writing – review & editing:** MengLi Cao.

## References

1. Liu C, Tsow F, Zou Y, Tao N. Particle Pollution Estimation Based on Image Analysis. *PLoS ONE*. 2016; 11(2):1–14.
2. Cao M-L, Meng Q-H, Jing Y-Q, Wang J-Y, Zeng M. Distributed Sequential Location Estimation of a Gas Source via Convex Combination in WSNs. *IEEE Transactions on Instrumentation and Measurement*. 2016; 65(6):1484–94.
3. Li J-G, Meng Q-H, Wang Y, Zeng M. Odor source localization using a mobile robot in outdoor airflow environments with a particle filter algorithm. *Autonomous Robots*. 2011; 30(3):281–92.
4. Cao M-L, Meng Q-H, Wang J-Y, Luo B, Jing Y-Q, Ma S-G. Learning to rapidly re-contact the lost plume in chemical plume tracing. *Sensors*. 2015; 15(4):7512–36. <https://doi.org/10.3390/s150407512> PMID: 25825974
5. Nehorai A, Porat B, Paldi E. Detection and localization of vapor-emitting sources. *IEEE Trans Signal Process*. 1995; 43(1):243–52.
6. Zhao T, Nehorai A. Distributed Sequential Bayesian Estimation of a Diffusive Source in Wireless Sensor Networks. *IEEE Transactions on Signal Processing*. 2007; 55(4):1511–24.
7. Nofsinger GT, Smith KW, editors. Plume source detection using a process query system. *Defense and Security*; 2004.
8. Michaelides MP, Panayiotou CG, editors. Plume Source Position Estimation Using Sensor Networks. *IEEE International Symposium on, Mediterrean Conference on Control and Automation*; 2005 27–29 June 2005.
9. Conn AR, Gould NIM, Toint PL. Trust-region methods. *MPS-SIAM Series on Optimization*, SIAM. 2000; 53(2):119–36.
10. Chen H-M. Greedy methods in plume detection, localization and tracking. *Greedy Algorithms*. Vienna: Intech; 2008. p. 305–24.
11. Gelb A. *Nonlinear Estimation. Applied optimal estimation*. Cambridge, MA: MIT press; 1974. p. 180–228.
12. Liu Y, Hu YH, Pan Q. Distributed, Robust Acoustic Source Localization in a Wireless Sensor Network. *IEEE Transactions on Signal Processing*. 2012; 60(8):4350–9.
13. Liu Y, Zhou W-T, Yan L, Pan Q, Cheng Y-M, editors. A novel PSO based acoustic source localization algorithm in Wireless Sensor Network. 2010 8th World Congress on Intelligent Control and Automation, WCICA 2010; 2010: Institute of Electrical and Electronics Engineers Inc.
14. Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization*. 2007; 39(3):459–71.

15. Karaboga D, Gorkemli B, Ozturk C, Karaboga N. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Artificial Intelligence Review*. 2012; 34:1–37.
16. Karaboga D, Akay B. A comparative study of Artificial Bee Colony algorithm. *Applied Mathematics and Computation*. 2009; 214(1):108–32.
17. Wang H-C, Wang Y-C, Tsae M-S, editors. Performance comparisons of Genetic Algorithm and Artificial Bee Colony Algorithm applications for localization in Wireless sensor networks. 2010 International Conference on System Science and Engineering, ICSSE 2010; 2010: IEEE Computer Society.
18. Mann PS, Singh S. Improved artificial bee colony metaheuristic for energy-efficient clustering in wireless sensor networks. *Artificial Intelligence Review*. 2019; 51(3):329–54.
19. Kumar R, Kumar D. Multi-objective fractional artificial bee colony algorithm to energy aware routing protocol in wireless sensor network. *Wireless Networks*. 2016; 22(5):1461–74.
20. Xu X, Zhou G, Chen T. Research on coverage optimisation of wireless sensor networks based on an artificial bee colony algorithm. *International Journal of Wireless and Mobile Computing*. 2015; 9(2):199–204.
21. Yang W, Li Q. Survey on partial swarm optimization algorithm. *Engineering Science*. 2004; 6(5):87–94.
22. Yamashita N, Fukushima M. On the Rate of Convergence of the Levenberg-Marquardt Method. *Computing Supplementa*. 152001. p. 239–49.
23. Roberts OFT. The Theoretical Scattering of Smoke in a Turbulent Atmosphere. *Proceedings of the Royal Society A Mathematical Physical & Engineering Sciences*. 1923; 104(104):640–54.
24. Farrell JA, Murlis J, Long X, Li W, Cardé RT. Filament-Based Atmospheric Dispersion Model to Achieve Short Time-Scale Structure odor plumes. *Environmental Fluid Mechanics*. 2002; 2(1):143–69.
25. Fukazawa Y, Ishida H, editors. Estimating gas-source location in outdoor environment using mobile robot equipped with gas sensors and anemometer. *IEEE Sensors Conference*; 2009 25–28 Oct. 2009; Christchurch, New Zealand: Institute of Electrical and Electronics Engineers Inc.
26. Singh MP, Yadav AK. Mathematical model for atmospheric dispersion in low winds with eddy diffusivities as linear functions of downwind distance. 1996; 30(7):1137–45.
27. Sharan M, Yadav AKJAE. Simulation of diffusion experiments under light wind, stable conditions by a variable K -theory model. 1998; 32(20):3481–92.
28. Steihaug T. The Conjugate Gradient Method and Trust Regions in Large Scale Optimization. *Siam Journal on Numerical Analysis*. 1983; 20(3):626–37.