

LOKALIZACJA ŹRÓDŁA ZANIECZYSZCZEŃ ATMOSFERYCZNYCH ZA POMOCĄ ALGORYTMU ROJU CZĄSTECZEK

Anna Wawrzyńczak^{1,2}, Joanna Dańko², Mieczysław Borysiewicz¹

¹ Narodowe Centrum Badań Jądrowych, Świerk – Otwock

² Uniwersytet Przyrodniczo-Humanistyczny w Siedlcach

Streszczenie. W pracy przedstawiono zastosowanie algorytmu Roju Cząsteczek (ang. Particle Swarm Optimization, PSO) w lokalizacji źródła zanieczyszczeń atmosferycznych w oparciu o informacje o stężeniu uwolnionej substancji w atmosferze, zarejestrowane przez sieć detektorów rozmieszczonych na danym terenie. Algorytm PSO przeszukuje pięciowymiarową przestrzeń w celu znalezienia parametrów modelu dyspersji, takich jak lokalizacja źródła skażenia (współrzędne X,Y), poziom uwolnienia substancji (Q) oraz współczynniki dyfuzji atmosferycznej w kierunku pionowym i poziomym (σ_y , σ_z), dla których przewidywane przez zastosowany model dyspersji stężenia w miejscach położenia detektorów będą zgodne z stężeniami zarejestrowanymi. W rekonstrukcji zastosowano model smugowy Gaussa.

W pracy zaprezentowano przykładową konfigurację algorytmu PSO, której zadaniem jest w jak najkrótszym czasie wskazać najbardziej prawdopodobne parametry źródła skażenia. Efektywność algorytmu przetestowano w oparciu o wygenerowane dane syntetyczne.

Słowa kluczowe: optymalizacja, rój cząsteczek, dyspersja zanieczyszczeń w atmosferze, lokalizacja źródła zanieczyszczeń

WSTĘP

Przypadkowe i nieoczekiwane uwolnienia substancji niebezpiecznych do atmosfery mogą powodować nieodwracalne skutki w środowisku oraz zdrowiu żywych organizmów. W związku z tym istotne jest, aby szybko reagować na wszelkie awarie, których skutkiem jest emisja niebezpiecznych substancji do atmosfery.

Przykładem takiej katastrofy jest wydzielenie do atmosfery substancji TCDD w Seveso w roku 1976. Dnia 10 lipca odnotowano największe uwolnienie tej substancji. W następstwie tej awarii powstały regulacje prawne znane jako Dyrektywy Seveso II

Adres do korespondencji – Corresponding author: Anna Wawrzyńczak, Narodowe Centrum Badań Jądrowych, ul. A. Sołtana 7, 05-400 Otwock, Polska; Joanna Danko, e-mail: danko.joanna@gmail.com

[B. De Marchi i in. 1996]. Szybka reakcja na tego typu nieprzewidziane zdarzenia może znacznie zmniejszyć zagrożenia wynikające z awarii.

Mając określone położenie źródła uwolnienia zanieczyszczeń, kierunek wiatru oraz moc uwolnienia substancji z łatwością można obliczyć przewidywane stężenie uwolnionej substancji w dowolnym punkcie przestrzeni. Z kolei określenie źródła i ilości uwolnionej substancji w oparciu jedynie o punktowe jej stężenia nie jest zadaniem trywialnym. Zagadnienie sprowadzić można do zaproponowania modelu transportu skażeń w atmosferze, który przewiduje stężenia substancji w punktach położenia sensorów identyczne z zarejestrowanymi eksperymentalnie. Jednakże zaproponowanie takiego modelu wymaga sprecyzowania wielu parametrów wejściowych, z których najistotniejsze są lokalizacja źródła uwolnienia oraz moc uwolnienia.

W efekcie zagadnienie poszukiwania źródła uwolnienia substancji w oparciu o zarejestrowane przez sieć sensorów jej stężenia może być sprowadzone do problemu optymalnego dopasowania parametrów modelu dyspersji, tak aby jego wynik jak najlepiej zgadzał się z danymi rzeczywistymi.

Wszechstronne ujęcie problematyki modelowania rozprzestrzeniania się zanieczyszczeń w atmosferze zawarto w publikacji Markiewicza. Problem dyspersji gazów, oraz stosowane modele smugowe opisuje Michalewicz [2003]. Algorytmami najczęściej wykorzystywanymi w rekonstrukcji źródeł uwolnień zanieczyszczeń atmosferycznych są metody Monte Carlo oparte o łańcuchy Markowa (ang. *Markov Chain Monte Carlo*, MCMC) [Gilks i in. 1996] oraz Sekwencyjne metody Monte Carlo (ang. *Sequential Monte Carlo*, SMC) [Doucet i in. 2001]. Efektywność wykorzystania MCMC w problemie lokalizacji źródła uwolnienia niebezpiecznych substancji na podstawie wygenerowanych danych syntetycznych przedstawiono w opracowaniu [Borysiewicz i in. 2012 ab]. Implementacja metody SMC oraz jej przewaga w znajdowaniu współrzędnych źródła skażeń nad metodą MCMC została zobrazowana przez Wawrzyńczak i in. [2014 a]. Z kolei algorytm lokalizacji uwolnień zanieczyszczeń do atmosfery za pomocą algorytmu genetycznego (AG) zaprezentowano w pracy Wawrzyńczak i in. [2014 b].

W niniejszej pracy przedstawiono zastosowanie algorytmu Roju Cząstek (PSO), w ocenie lokalizacji źródła skażeń atmosferycznych. Algorytm Roju Cząstek (PSO) imituje ruch żywych stadnych organizmów (np. zachowanie ptaków) [Kennedy i Eberhard, 1995]. Zastosowanie PSO w przytoczonym problemie jest innowacyjnym podejściem. Algorytm ten nie był do tej pory wykorzystywany w problematyce rekonstrukcji rozprzestrzeniania się zanieczyszczeń. PSO jest stosunkowo nową metodą optymalizacji, lecz mimo to do tej pory znaleziono dla niego wiele zastosowań, dzięki czemu zyskał wielu zwolenników. Naśladowanie zachowania stada wykorzystano między innymi jako narzędzie skutecznie wspomagające optymalizację elementów konstrukcji [Foryś, 2008] przy modelowaniu krzywizny układu geometrycznego toru [Palikowska, 2013]. Swoje zastosowanie znalazł również w rozwiązywaniu zagadnień permutacyjnych [Chmiel i in. 2011] oraz w wielu problemach optymalizacyjnych [np. Filipowicz i Kwiecień, 2011].

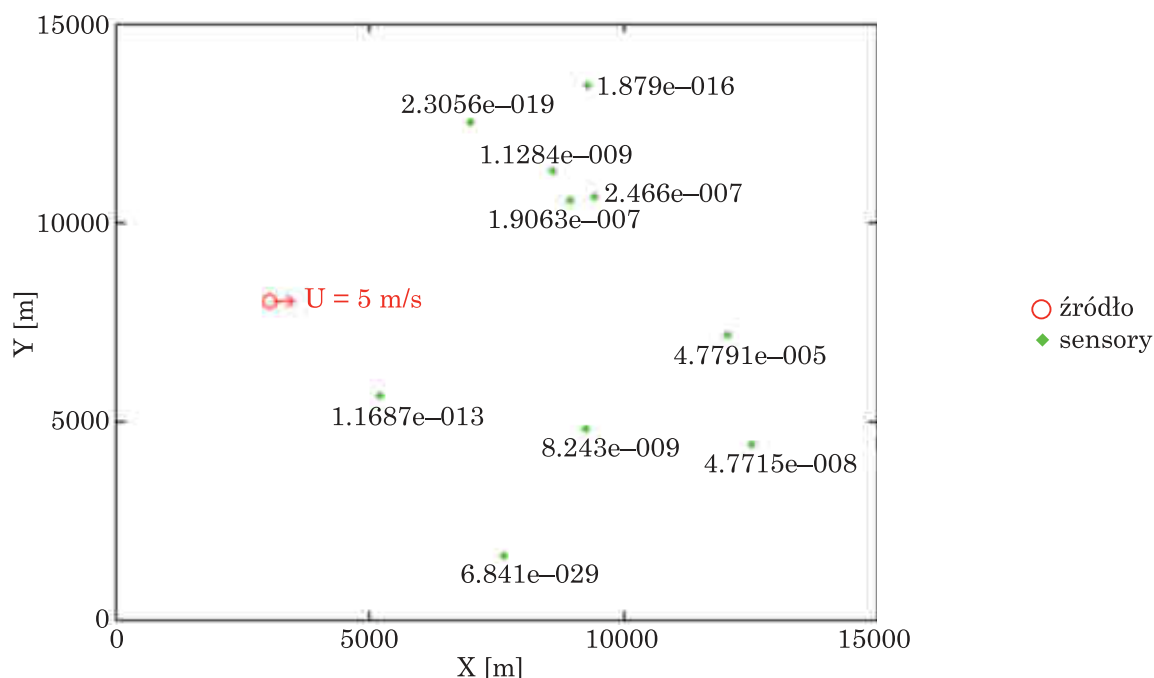
Algorytmy metaheurystyczne, aby znaleźć „optimum” muszą mieć określony system za pomocą którego będą mogły szukać celu. Każdy algorytm posiada parametry, których zadaniem jest je kontrolować. Wartości tych parametrów znacznie wpływają na wyniki zwracane przez algorytm, można je dowolnie modyfikować bądź rozszerzać zależnie od

potrzeb problemu [np. Reyes Sierra i Coello, 2005; Cai i in. 2009]. Najważniejszymi cechami jakimi powinien charakteryzować się model jest jakość oraz szybkość znajdowanego „optimum”. Głównymi i podstawowymi czynnikami wpływającymi na szybkość oraz jakość rozwiązania są wielkość populacji, rozmiar obszaru poszukiwań oraz przede wszystkim złożoność rozważanego problemu. Istotne jest, aby kompromis pomiędzy efektywnością a prędkością był zachowany.

W niniejszej pracy zaprezentowano zastosowanie PSO w lokalizacji źródła zanieczyszczeń atmosferycznych na podstawie zarejestrowanych stężeń z sieci czujników. Parametry algorytmu PSO oraz jego efektywność przetestowano w oparciu o wygenerowane dane syntetyczne.

SFORMUŁOWANIE PROBLEMU

Efektywność algorytmu Roju Cząsteczek w problemie lokalizacji źródła skażeń atmosferycznych była testowana w oparciu o wygenerowane na te potrzeby dane syntetyczne. Domena testowa, w której rozmieszczone były źródło skażenia oraz czujniki mierzące stężenia obejmowała obszar $15 \text{ km} \times 15 \text{ km}$ przedstawione na rysunku 1. Dane wejściowe algorytmu PSO to punktowe stężenia, zarejestrowane przez dziesięć losowo rozmieszczonych detektorów. Dane syntetyczne zostały wygenerowane przez model smugowy Gaussa [Turner, 1994]. W trakcie symulacji źródło skażenia umieszczone było w punkcie o współrzędnych $X = 3000 \text{ km}$ oraz $Y = 8000 \text{ km}$, na wysokości $Z = 50 \text{ m}$ nad ziemią.



Rys. 1. Lokalizacja źródła uwolnienia oraz sensorów w domenie testowej i zarejestrowane przez sensory stężenia przekazane do algorytmu PSO

Fig. 1. Distribution of the release source and the sensors in the test domain and the concentration registered by the sensors passed to the PSO algorithm

Poziom emisji substancji wynosił $Q \approx 5000$ g/s. Wiatr skierowany wzdłuż osi X miał prędkość $U=5$ m/s. Otrzymane w wyniku symulacji stężenia zostały zaburzone losowo maksymalnie o 20% zanim zostały przekazane do algorytmu rekonstrukcji. W ten sposób uwzględniono ewentualne turbulencje oraz straty na sensorach.

W procesie rekonstrukcji konieczne jest zastosowanie odpowiedniego modelu dyspersji skażeń w atmosferze by móc obliczyć stężenia C_i^M na 'i' tym sensorze dla danych parametrów modelu M w każdym kroku algorytmu. W tej pracy wykorzystano model smugowy Gaussa, który zapewnia kompromis między wiarygodnością zwracanych wyników, a szybkością działania przy tym nie wymaga sprecyzowania zbyt dużej ilości parametrów.

Przy jednorodnym wietrze stężenie substancji $C(\tilde{x}, \tilde{y}, z)$ w punkcie \tilde{x} metrów z wiatrem od źródła, \tilde{y} metrów od środka chmury oraz z metrów nad ziemią można obliczyć następująco:

$$C(\tilde{x}, \tilde{y}, z) = \frac{Q}{2\pi\sigma_y U} \exp\left[-\frac{1}{2}\left(\frac{\tilde{y}}{\sigma_y}\right)^2\right] \cdot \left\{ \exp\left[-\frac{1}{2}\left(\frac{z-H}{\sigma_y}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{z+H}{\sigma_y}\right)^2\right] \right\}, 1$$

Gdzie U to średnia prędkość wiatru w warstwie powietrza od $z = 0$ do $z = H$ w kierunku równoległym do osi X. Q to masa zanieczyszczeń wyrzucana ze źródła emisji wyrażana w jednostkach g/s, H to wysokość pozornego punktu emisji, wyrażona jako suma wysokości rzeczywistej komina (h) i tak zwanej wartości „wyniesienia” (Δh). W powyższym równaniu zmienne to współczynniki dyfuzji atmosferycznej odpowiednio w kierunku poziomym i pionowym. Dwa ostatnie parametry odnoszą się do klasy stabilności Pasquilla-Gifforda [Pasquill F., 1961]. Generując dane syntetyczne wybrana została klasa stabilności C, dla terenów miejskich. Odpowiednie formuły przedstawiają się następująco:

$$\sigma_y = 0.22 \cdot \tilde{x} \cdot \left(1 + \tilde{x} \cdot 4 \cdot 10^{-5}\right)^{0.5}, \sigma_z = 0.20 \cdot \tilde{x}, 2$$

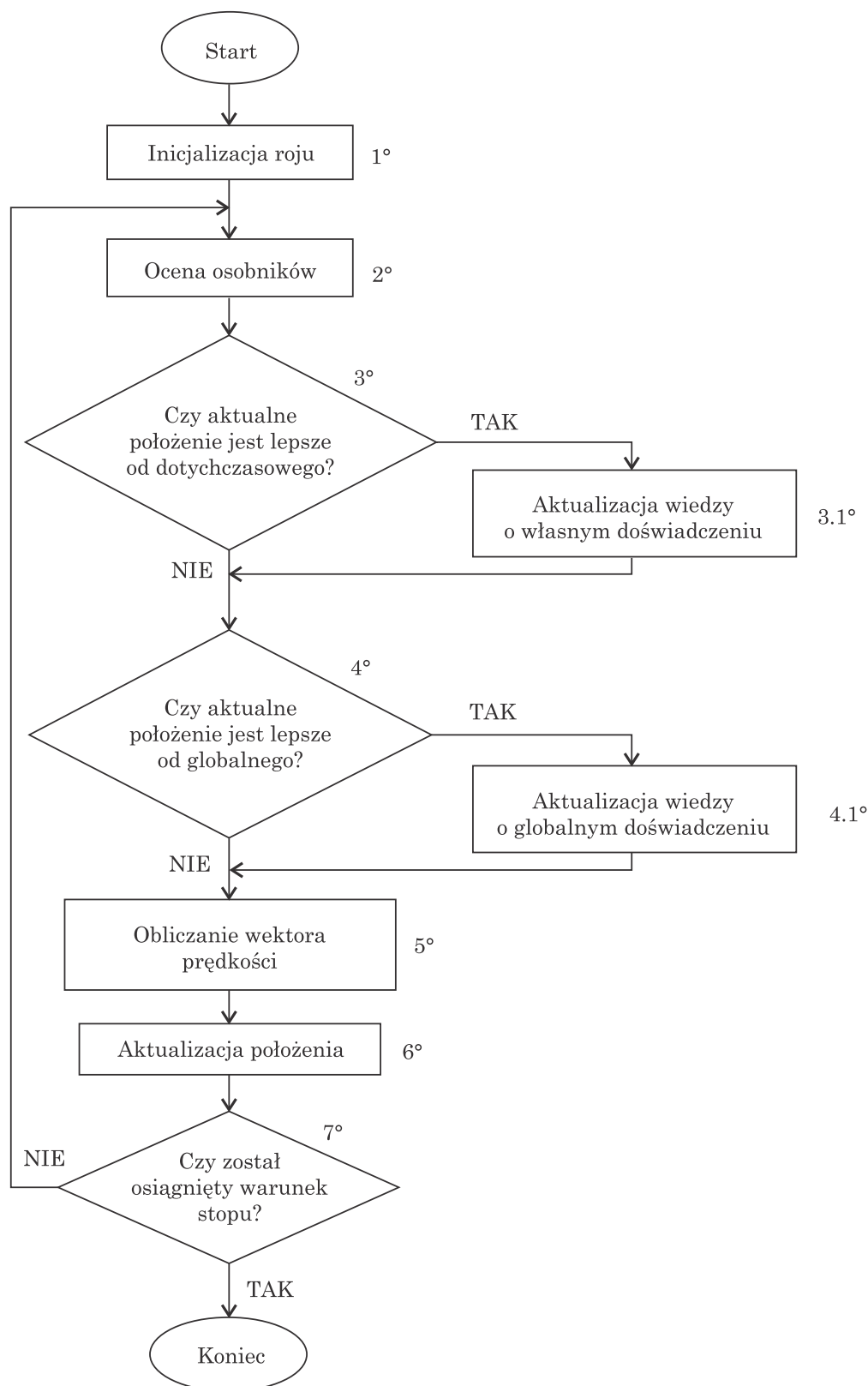
Jednakże w trakcie rekonstrukcji założono, że nie jest znane dokładne zachowanie smugi. W związku z tym w trakcie rekonstrukcji parametry zostały przyjęte jako:

$$\sigma_y = z_1 \cdot \tilde{x} \cdot \left(1 + \tilde{x} \cdot 4 \cdot 10^{-5}\right)^{0.5}, \sigma_z = z_2 \cdot \tilde{x}, 3$$

gdzie wartości z_1 oraz z_2 są próbkowane przez algorytm w dopuszczalnym fizycznie przedziale $\langle 0.001, 0.035 \rangle$.

Podsumowując, w niniejszym artykule przestrzeń poszukiwanych parametrów ma postać $M=(x,y,Q,z_1,z_2)$, gdzie x i y to współrzędne źródła w domenie, Q jest poziomem emisji substancji i z_1, z_2 to próbkowane parametry współczynników dyfuzji atmosferycznej. Ponadto w trakcie rekonstrukcji założono, że źródło skażenia znajduje się na wysokości 20 metrów (tj. 30 m niżej niż faktycznie).

ALGORYTM ROJU CZĄSTECZEK



Rys. 2. Schemat blokowy PSO
 Fig. 2. Flow chart of the PSO algorithm

Lokalizacja źródła skażenia w zdefiniowanym obszarze sprowadza się w tym przypadku do zaproponowania modelu dyspersji skażeń, którego przewidywane stężenia w punktach położenia sensorów będą najbardziej zgodne z rzeczywistymi. W efekcie zagadnienie sprowadza się do optymalnego doboru zestawów parametru modelu. Jako narzędzie optymalizacji wykorzystany został algorytm Roju Cząsteczek, którego schemat blokowy został przedstawiony na rysunku 2.

Schemat blokowy (rys. 2) prezentuje koncepcje PSO widzianą oczami autorów [Kennedy i Eberhard, 1995]. Fundamentalnym elementem algorytmu PSO jest populacja osobników (zbiór cząstek) poruszających się w wielowymiarowej przestrzeni z prędkością określającą położenie danego osobnika w kolejnej iteracji. Ruch cząstek jest odcinkowo prostoliniowy. Położenie i -tego osobnika w punkcie czasowym t (iteracji) przedstawia się następująco:

$$x_i^{t+1} = x_i^t + v_i^t, 4$$

Formuła aktualizacji wektora przesunięcia i -tej cząstki w czasie t wyrażana jest poniższym wzorem:

$$v_i^{t+1} = c_1 r_1 v_i^t + c_2 r_2 (b_i^t - x_i^t) + c_3 r_3 (bg^t - x_i^t), 5$$

gdzie:

- c_1, c_2, c_3 – wagi intuicji, świadomości cząstki oraz myślenia społecznego,
- r_1, r_2, r_3 – liczby losowe rozkładu normalnego z przedziału $<0,1>$,
- b_i^t – najlepsze dotychczas znalezione przez cząstkę rozwiązanie,
- bg^t – najlepsze rozwiązanie znalezione przez rój.

Wartość prędkości (równanie 5, rys. 2, pkt 5^o) wyliczana jest we wszystkich iteracjach, dla każdego osobnika indywidualnie. Współczynniki c_1, c_2, c_3 to zmienne o stałej wartości. Losowe parametry r_1, r_2, r_3 mają na celu wprowadzenie różnorodności w roju i skutkują wybijaniem osobników z lokalnych optimów.

Po inicjalizacji roju (1^o) poszczególna cząstka poddawana jest ocenie (2^o), której zasadniczą intencją jest analiza jakości położenia danej jednostki. Skutkiem klasyfikacji jest rozróżnienie osobników i ukierunkowanie roju. Funkcja wiarygodności (przystosowania) określa jak dobrze osobnik jest dostosowany do środowiska i na jej podstawie wyznaczany jest cel, do którego mają dążyć cząstki. Rój podąża za bieżąco najlepszym położeniem modyfikując prędkość oraz kierunek „lotu”. Pojedyncze cząstki w poszczególnych iteracjach poddawane są weryfikacji (3^o i 4^o). Sprawdzana jest jakość aktualnego położenia z dotychczas najlepszym cząstki (3^o) oraz najlepszym w obszarze roju położeniem (4^o). Osobnik „podmienia” swoje informacje po natrafieniu na jakościowo lepsze dane (3.1^o i 4.1^o). Warunkiem przemieszczania się jednostek jest ewaluacja formuły prędkości (równanie 5), po której następuje przesunięcie się osobników o obliczony wektor (równanie 4). Czynności są cyklicznie powtarzane, aż do momentu spełnienia warunku stopu (7^o), którym w przedstawionym algorytmie jest wykonanie ustalonej liczby iteracji.

Każdy osobnik (cząstka) składa się z następujących elementów:

- aktualnego położenia (rozwiązanie),
- aktualnej prędkości (ograniczona z góry),

- dotychczas najlepszego znalezionej położenia (rozwiązanie),
- wartości oceny aktualnej

Dużym atutem algorytmu jest to, że operuje on tylko i wyłącznie na wartościach zmiennoprzecinkowych i nie wymaga konwersji na liczby binarne (jak w przypadku Algorytmu Genetycznego). Dzięki tej właściwości kod wykonuje się szybko, zmniejsza się ryzyko wystąpienia błędów, a dane są dokładniejsze bo nie muszą przechodzić przez szereg operacji, gdzie wraz ze wzrostem operacji istnieje możliwość wzrostu błędów obliczeń.

IMPLEMENTACJA PSO W PROBLEMIE LOKALIZACJI ŹRÓDŁA SKAŻEŃ

Opierając się na informacjach zarejestrowanych przez detektory, PSO powinien znaleźć parametry pasujące do atrybutów modelu dyspersji, czyli bliskie źródłowym, posługując się jedynie danymi z sensorów. Na potrzeby tego problemu model próbkuję przestrzeń pięciowymiarową. Odpowiednio każda cząsteczka przechowuje takie dane jak:

- x, y – współrzędne źródła uwolnienia substancji wyrażone w metrach,
- Q – emisja zanieczyszczenia w jednostce gram na sekundę,
- z_1, z_2 – współczynniki w parametrach.

W pierwszej iteracji każdy poszukiwany parametr modelu jest inicjalizowany losową liczbą rzeczywistą z przeszukiwanego przedziału kolejno: $x, y \in \langle 0, 15000 \rangle$, $Q \in \langle 1, 10000 \rangle$, $z_1, z_2 \in \langle 0.001, 0.35 \rangle$.

Przykładową postać cząstki prezentuje rysunek 3:

12903,89	2993,39	6996,28	0,3218	0,1988	-16980,08
↑	↑	↑	↑	↑	↑
Parametr X	Parametr Y	Parametr Q	Parametr Z1	Parametr Z2	Ocena osobnika

Rys. 3. Przykładowa postać cząstki

Fig. 3. An example of the particle representing the searched model's parameters

3.2 FUNKCJA WIARYGODNOŚCI

Ocena jakości osobników jest wykonywana dla każdej cząstki, w oparciu o wartość funkcji wiarygodności. W uproszczeniu, kryterium oceny polega na porównaniu danych wygenerowanych przez model z danymi syntetycznymi w każdym kroku czasowym algorytmu (rys. 2):

$$f(C_i^M, C_i^E) = \sum_{i=1}^N \left[\log(C_i^M) - \log(C_i^E) \right]^2, 6$$

gdzie:

C_i^M – wygenerowane z modelu dyspersji (modelu smugowego Gaussa) wartości stężeń na i -tym sensorze,

C_i^E – rzeczywiste stężenie zarejestrowane przez i -ty sensor,

N – liczba sensorów.

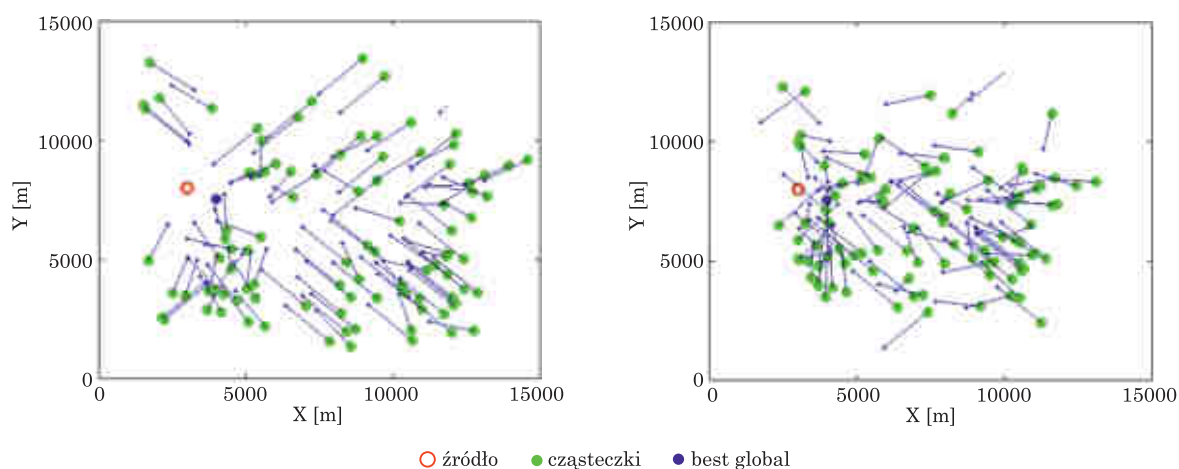
Poszukiwane parametry modelu (M) będą tym bliższe rzeczywistym, im jest wyższa wartość funkcji przystosowania.

WEKTOR PRZESUNIĘCIA

Główną operacją za pomocą której cząstki podejmują decyzję o zmianie położenia to formuła wektora przesunięcia, której ogólną postać przedstawiono w rozdziale 3. Na potrzeby rozwiązania problemu lokalizacji źródła emisji został on przekształcony do następującej postaci:

$$v_i^{t+1} = \omega_1 c_1 r_1 v_i^t + c_2 r_2 (b_i^t - x_i^t) + c_3 r_3 (bg^t - x_i^t), 7$$

Różnica polega na wzbogaceniu formuły o parametr bezwładności (ω_i), o przypisywanych wartościach „0” lub „1”. Zadaniem tej zmiennej było „dezaktywowanie” pierwszego z trzech wektorów poprzez przybieranie wartości $\omega=0$ w przypadku, gdy osobnik w aktualnej iteracji obrał „właściwy kierunek”. Czyli, jeśli jego obecne położenie stało się jego dotychczasowo najlepszym (zaktualizował b_i^t) bądź stał się „liderem” roju (zaktualizował bg^t), czyli nastąpiła poprawa wartości funkcji celu. Czynność ta powodowała kierowanie się osobnika w stronę najlepszego położenia bez brania pod uwagę modułu odpowiedzialnego za losowy kierunek. Cząstka mogła przez jedną iterację kierować się w stronę „najlepszego” osobnika, bądź nie ruszać się w przypadku gdy stała się „najlepszą”. Wartość $\omega_i=1$ zezwala na swobodny ruch i przyśpieszenie cząstki. Parametr bezwładności pod koniec każdej iteracji powraca do wartości początkowej ($\omega_i=1$) w celu uniknięcia „wpadania” cząstki w lokalne maksima lub zwiększania się liczby nieruszających się „liderów” wraz z kolejnymi powtórzeniami.



Rys. 4. Przesunięcia cząstek w drugiej i w trzeciej iteracji w płaszczyźnie XY

Fig. 4. The particles displacements in the seconds and third iteration in the XY surface

Rysunek 4 przedstawia ruch cząstek w drugiej i w trzeciej iteracji w płaszczyźnie XY. Niebieskie linie reprezentują długość wektora przesunięcia, u większości osobników została osiągnięta prędkość maksymalna. Zielone punkty to osobniki, natomiast niebieski „grot” na drugim końcu wektora to położenie cząstki w kolejnej iteracji.

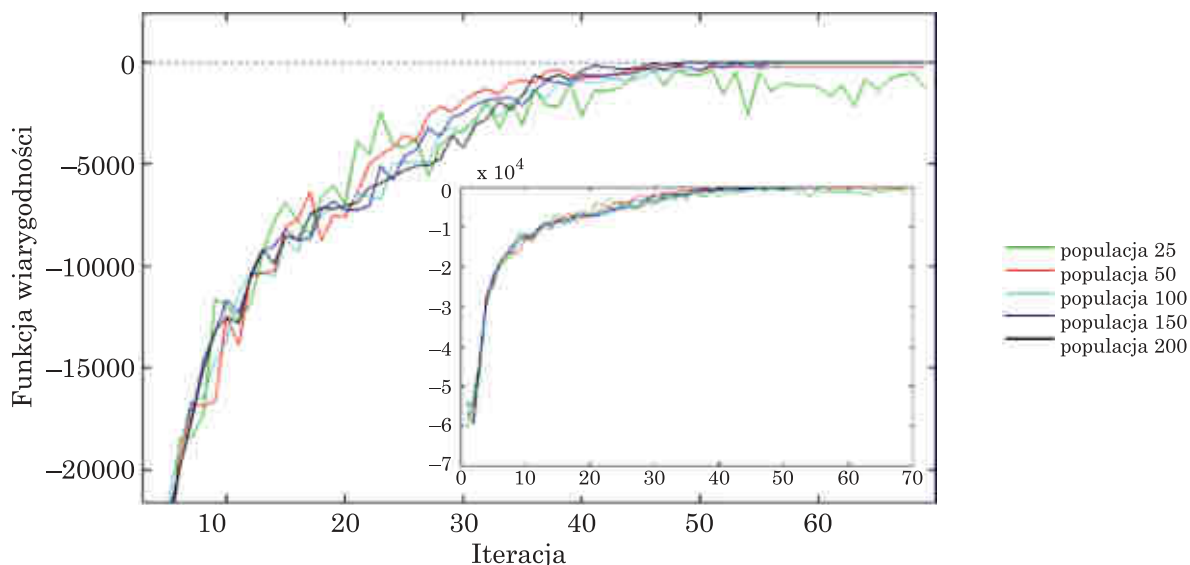
PARAMETR SĄSIEDZTWA

W celu uniknięcia efektu eksplozji cząstek (wychodzenie osobników poza przestrzeń poszukiwań) autorzy algorytmu przyjęli ograniczenia na prędkość osobników [Kennedy i Eberhard, 1995]. Zastosowany parametr sąsiedztwa dotyczy tego właśnie parametru.

W przedstawionej implementacji założono 10% wartości górnej granicy próbkowania dla każdego parametru, np. dla współrzędnej x wartość $v_{max} = 1500$ m. To rozwiązanie skutecznie znajduje optimum i jednocześnie daje elastyczność w przypadku zmiany wielkości domeny. Jednakże, posiada wadę w postaci nie zbiegania się roju. Duży przeskok powodował „gubienie się” pojedynczych cząsteczek w przestrzeni, wydłużając czas zbiegania się cząstek, lecz nie pogarszając w żaden sposób jakości wyniku. W celu poprawienia zbieżności algorytmu wprowadzono parametr sąsiedztwa, którego zadaniem jest „monitorowanie” ilości osobników znajdujących się na miejscu optimum globalnego. Założono, że jeśli 90% osobników znajdzie się w sąsiedztwie bg (z ang. *best global*) to v_{max} pozostałych 10% cząstek staje się równa wartości różnicy najlepszego rozwiązania i położenia „zagubionego” osobnika ($bg^t - x_i^t$). Skutkiem takich modyfikacji jest algorytm znajdujący optimum i nie mający problemu ze zbieżnością.

TESTOWANIE ALGORYTMU

Opisany algorytm PSO został zaimplementowany i przetestowany w środowisku MATLAB. W analizowanym zagadnieniu funkcja oceny (równanie 6) osiąga wartości z przedziału $-\infty, 0$. Kwantyfikatorem poprawności działania algorytmu jest funkcja przystosowania dążąca do '0'. Przedstawiony na rysunku 5 wykres funkcji wiarygodności w kolejnych iteracjach pokazuje, że różnice we wzroście funkcji oceny w każdym z zaprezentowanych przypadków są niewielkie. Główną cechą przedstawionego przebiegu jest zależność między wartością funkcji, a liczbą iteracji algorytmu PSO. Wraz ze zwiększaniem się ilości iteracji, wartość funkcji oceny dąży do '0', co świadczy o poprawnym działaniu algorytmu. Można również zaobserwować, że przy większej ilości osobników przebieg funkcji oceny jest bardziej stabilny. Rozbieżność pomiędzy rezultatami uzyskanymi wykorzystując 100 czy 200 osobników nie jest na tyle duża, aby można było uznać za konieczne zwiększenie liczby osobników. Trzeba pamiętać, że zwiększenie rozmiaru populacji wydłuża znacząco czas pracy algorytmu, który ma istotne znaczenie w zastosowaniu praktycznym. W związku z tym w przedstawionej rekonstrukcji zastosowano populację liczącą 100 cząstek.

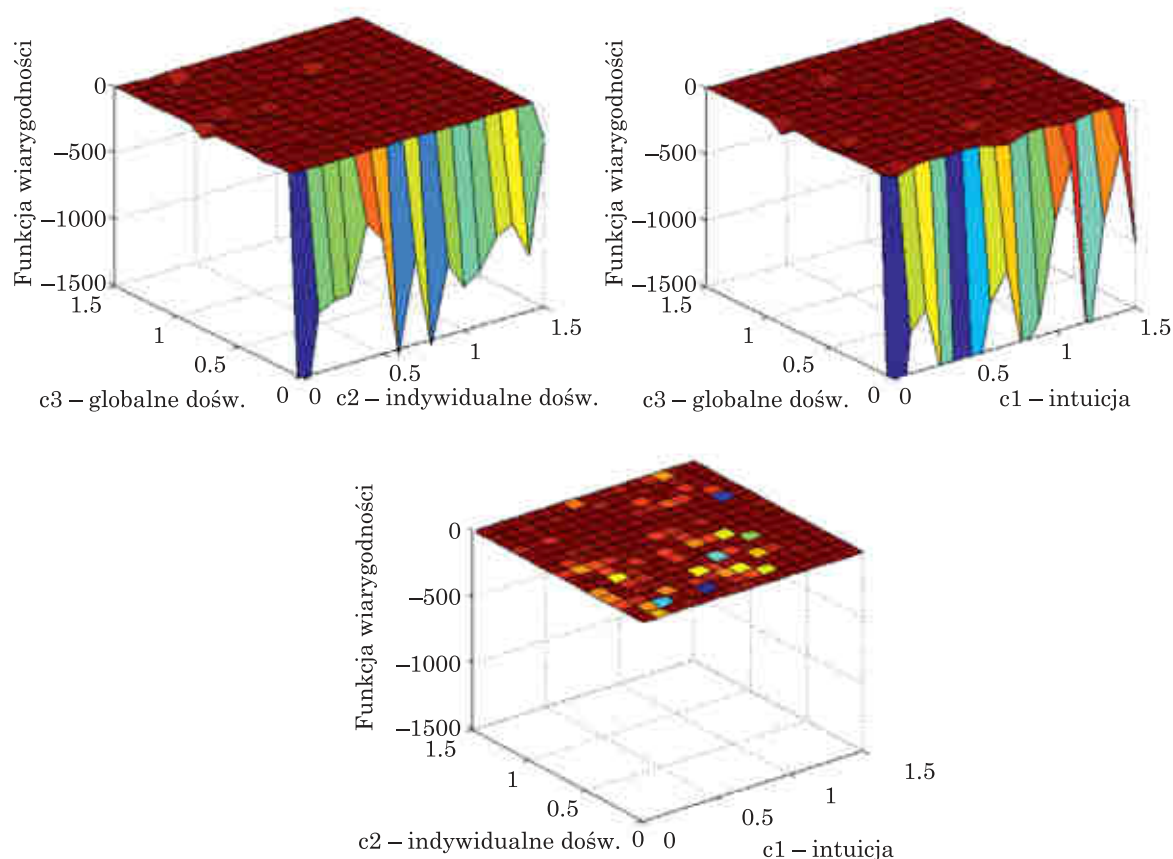


Rys. 5. Średnia wartość funkcji wiarygodności z 50 uruchomień algorytmu w kolejnych iteracjach algorytmu PSO, dla populacji składającej się odpowiednio z 25,50,100,150 oraz 200 osobników

Fig. 5. The average objective function values obtained during the 50 algorithm PSO runs for 25, 50, 100, 150 and 200 particles in the population

Rysunek 6 obrazuje zależność osiągniętej wartości funkcji wiarygodności względem zmieniających się wag. Poszczególne wykresy znajdujące się na rysunku 5 przedstawiają pięć średnich wartości funkcji oceny z pięćdziesięciu uruchomień dla każdego przypadku. Ze względu na brak możliwości przedstawienia zmieniającej się funkcji oceny na podstawie trzech wag na raz, zmuszeni byliśmy pojedyncze wagi zatrzymać na wcześniej ustalonych wartościach. Wszystkie wartości „unieruchomionych” parametrów były ustalone na podstawie testów i finalnie najlepszą okazała się konfiguracja: $c1=0.8$, $c2=1.3$, $c3=1.5$. Pierwszy wykres (z lewej) ilustruje związek wzrostu wartości funkcji oceny względem wagi $c3$ oraz $c1$, próbkowanych na przedziałach $\langle 0.0, 1.5 \rangle$, odpowiednio ustalono $c2=1.3$. Środkowy wykres rysunku 6 odzwierciedla zmiany wartości funkcji wiarygodności w stosunku do wag $c3$ i $c2$, próbkowanych na tych samych przedziałach, gdzie $c1=0.8$ oraz ostatni wykres przedstawia analogiczną zależność względem wagi $c1$ oraz $c2$ z zamrożonym $c3=1.5$. Ostatni wykres pokazuje najważniejszą właściwość tych trzech wag. Okazuje się bowiem, że istotnym kryterium poprawnego działania algorytmu jest przypisanie do parametru $c3$ wartości większej od zera, jest to waga odpowiedzialna za ukierunkowanie cząstek na „najlepsze” globalnie położenie (rozdział 3), realizację głównej idei algorytmu.

Nieczytelność pierwszych dwóch wykresów jest skutkiem osiągniętej bardzo niskiej, w stosunku do reszty wyników, wartości funkcji dopasowania dla $c3=0$. Ostatni wykres ($c1, c2$), pomimo braku skrajnych wyników, zaprezentowaliśmy w tej samej skali co pozostałe dwa, aby przedstawić istotność parametru świadomości globalnej. W celu poprawy czytelności zmniejszyliśmy przedziały wyświetlanych danych z $\langle 0, 1.5 \rangle$ do $\langle 0.5, 1.5 \rangle$ eliminując w ten sposób skrajne wartości.

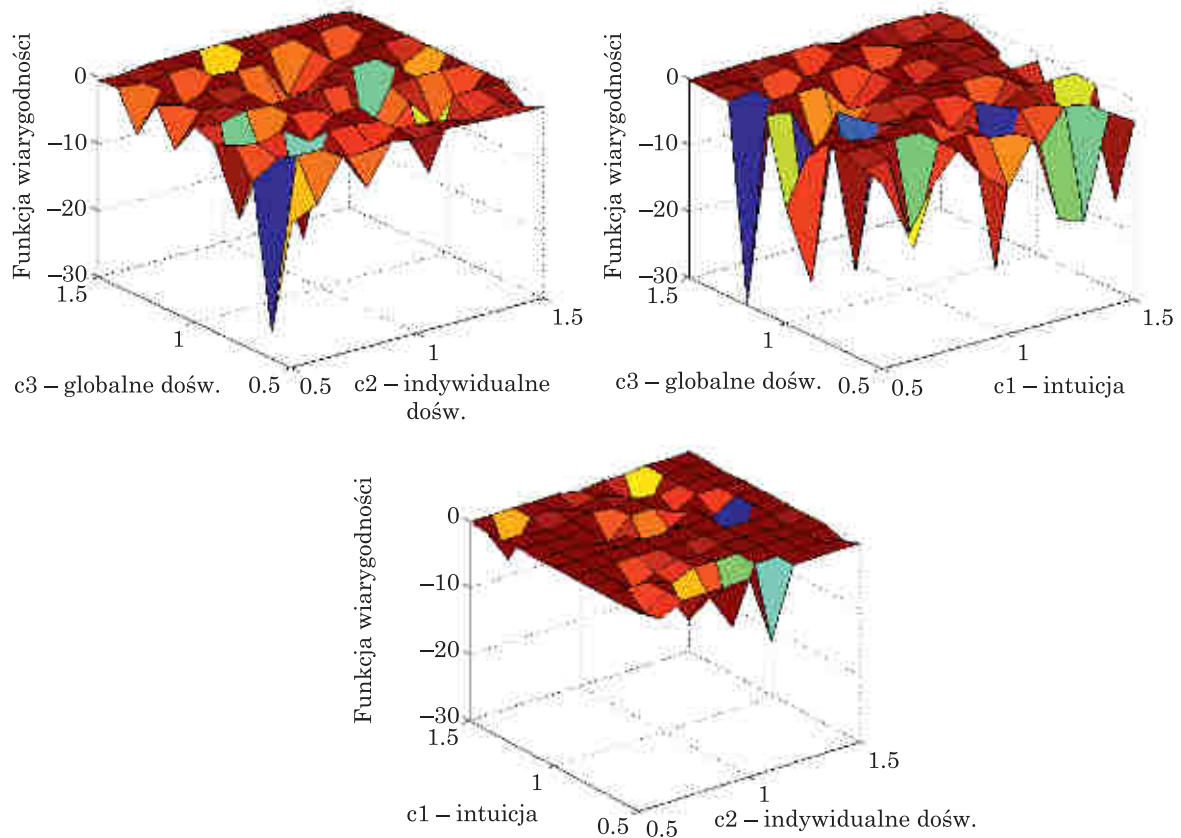


Rys. 6. Wykresy zależności wartości funkcji oceny względem wag świadomości. Na wszystkich wykresach zastosowano tę samą skalę $\langle -1500;0 \rangle$, $\langle 0.0;1.5 \rangle$, $\langle 0.0;1.5 \rangle$

Fig. 6. The dependences of the objective function value versus the weighs. All figures use the same scale $\langle -1500;0 \rangle$, $\langle 0.0;1.5 \rangle$, $\langle 0.0;1.5 \rangle$

W tym przypadku (rys. 7) poza tym, że osiągnięte wartości funkcji wiarygodności są bardzo bliskie zero można łatwo zaobserwować, że wartość funkcji oceny jest najbliższa optimum w przypadku, gdy waga c_3 jest równa 1.5. Dzieje się tak w dwóch pierwszych przypadkach (c_3, c_1 oraz c_3, c_2). Tą tezę wspiera ostatni, trzeci wykres ilustrujący największą stabilność algorytmu (na sztywno ustawione $c_3=1.5$). Ponadto, można zauważyć, że funkcja oceny osiąga optimum w przypadku kombinacji $c_2=\langle 1.1; 1.5 \rangle$ $c_1=0.8$ lub 1.3. Dowodzi to słuszności wyboru następującej konfiguracji algorytmu $c_1=0.8$, $c_2=1.3$, $c_3=1.5$.

Podsumowując, przeprowadzone testy wykazały, że algorytm jest wystarczająco silny na potrzeby rozważanego zagadnienia. Nie jest on zbyt wrażliwy na nieznaczne zmiany parametrów, w większości przypadków funkcja celu waha się w przedziale $\langle -15;0 \rangle$. Inna konfiguracja wag może spowodować wydłużenie się czasu poszukiwania optimum i w niewielu wypadkach spowodować błędne wyniki.

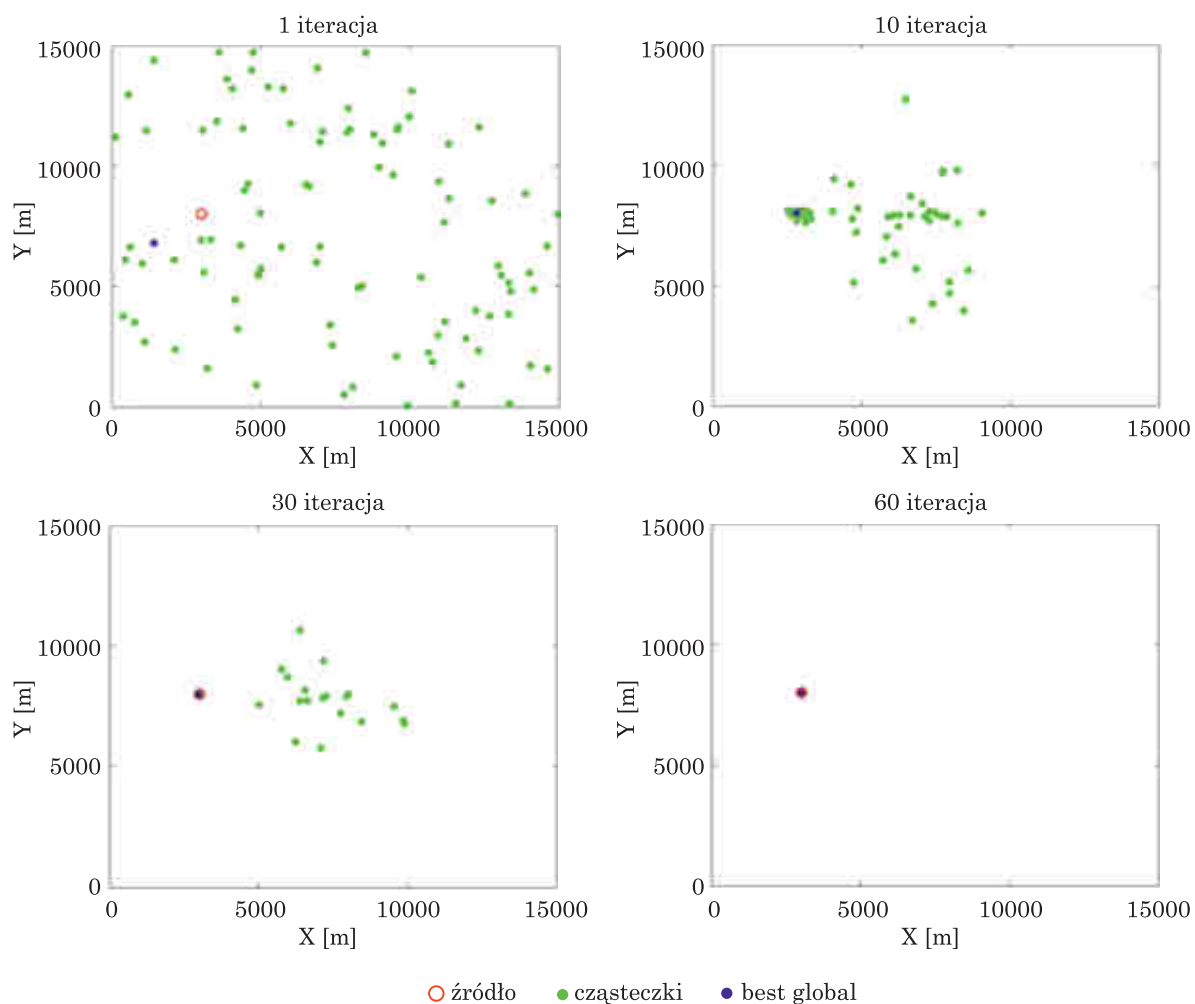


Rys. 7. Wykresy zależności funkcji oceny względem wag świadomości w skali $\langle -30;0 \rangle$, $\langle 0.5;1.5 \rangle$, $\langle 0.5;1.5 \rangle$

Fig. 7. The dependences of the objective function value versus the weighs. All figures use the same scale $\langle -30;0 \rangle$, $\langle 0.5;1.5 \rangle$, $\langle 0.5;1.5 \rangle$

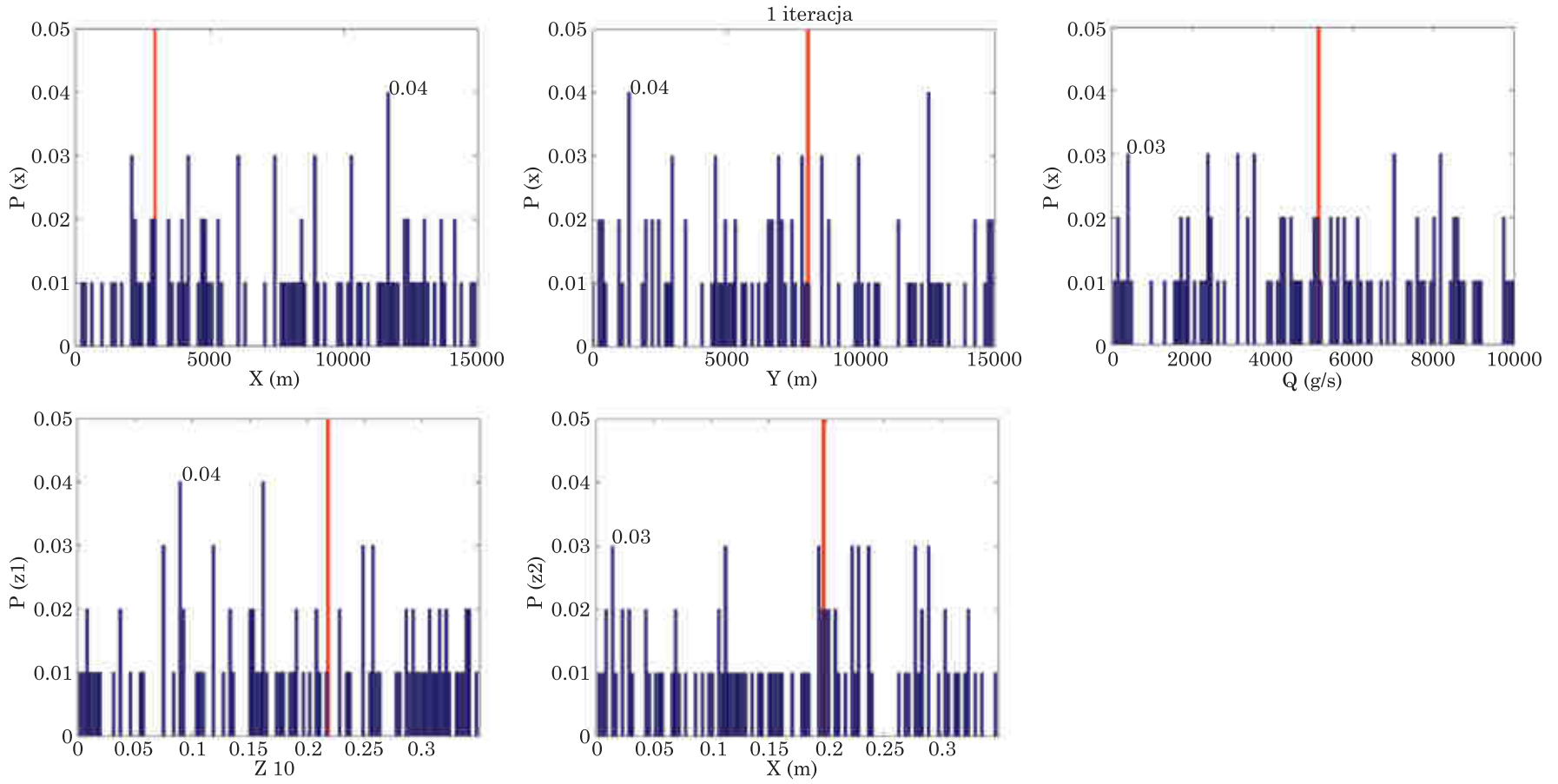
W efekcie przeprowadzonych testów w trakcie rekonstrukcji źródła uwolnienia zastosowano poniższą konfigurację parametrów algorytmu PSO:

- Liczba osobników: 100 cząstek,
- Wagi odpowiadające za komunikację osobników: $c1=0.8$, $c2=1.3$, $c3=1.5$.
- Parametr sąsiedztwa: 90%.
- Liczba iteracji: 60.



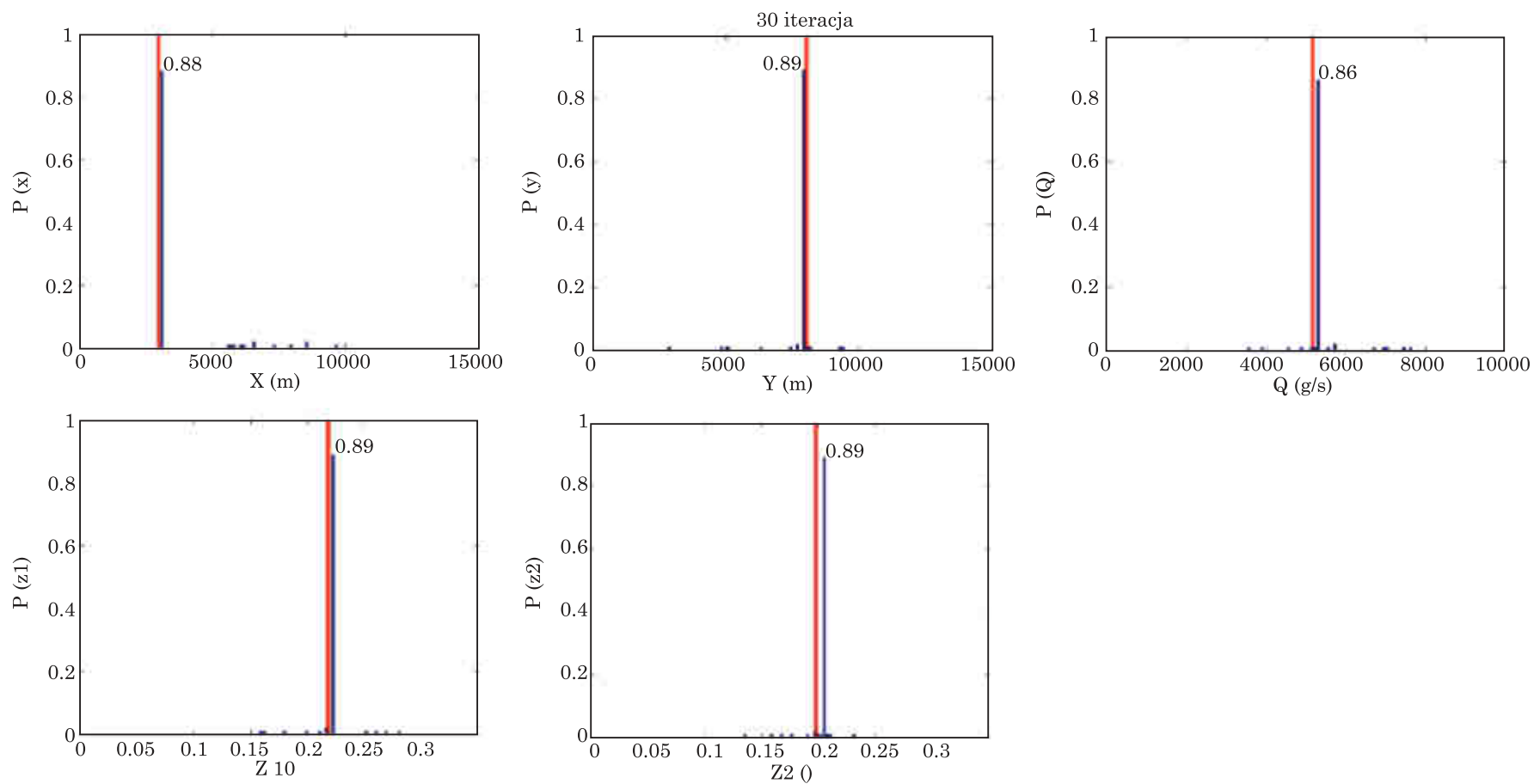
Rys. 8. Rozkład oszacowanych wartości parametrów x i y w 1, 10, 30 i 60 iteracji algorytmu PSO. ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)

Fig. 8. The distribution of the xandy coordinates estimated during the PSO runs in 1st, 10th, 30th and 60th iteration ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)



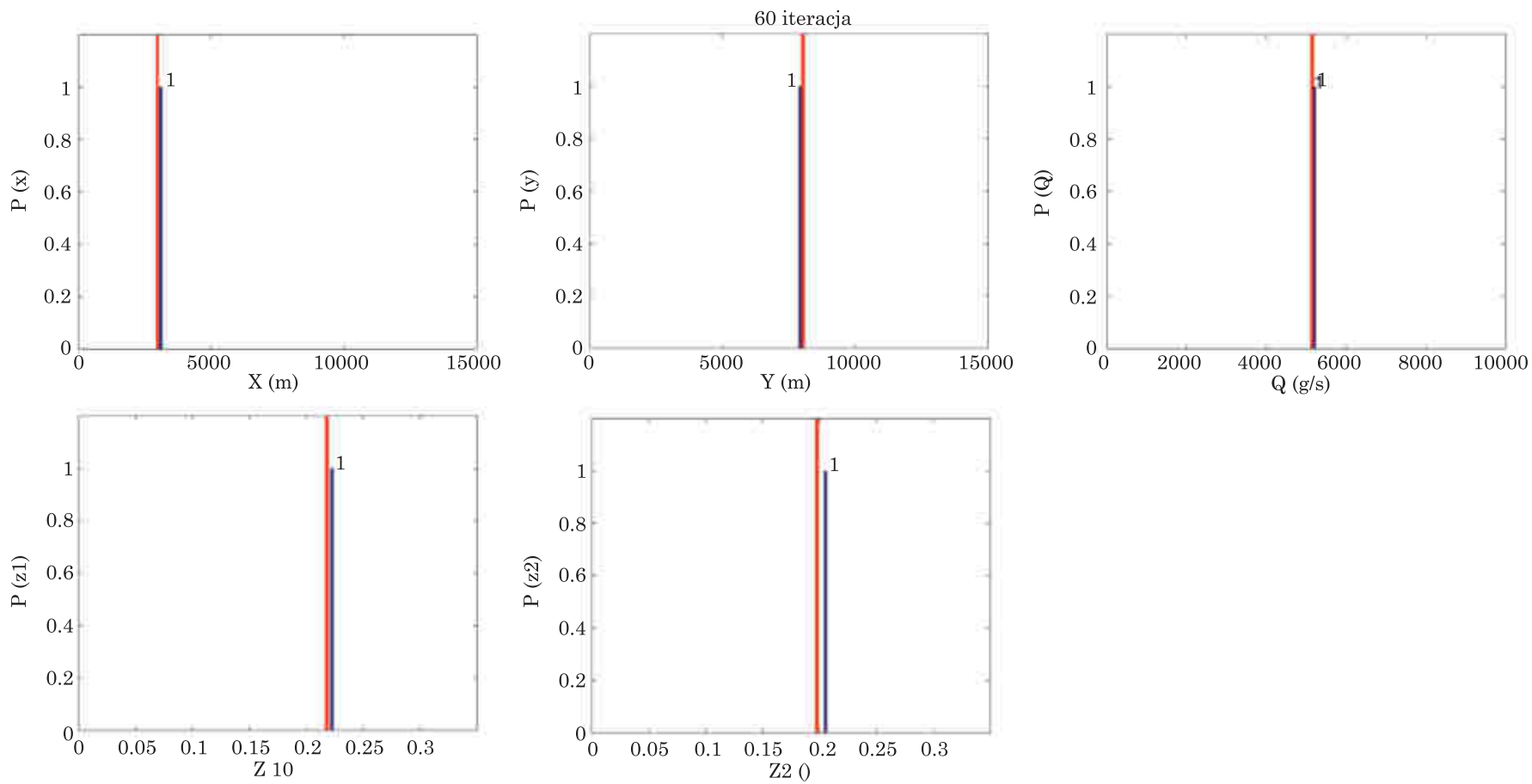
Rys. 9. Rozkład prawdopodobieństw wartości poszukiwanych parametrów modelu w 1. iteracji algorytmu PSO ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)

Fig. 9. Probability distributions of the models parameters x , y , Q , z_1 and z_2 in the generation in the PSO 1st iteration ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)



Rys. 10. Rozkład prawdopodobieństw wartości poszukiwanych parametrów modelu w 30. iteracji algorytmu PSO ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)

Fig. 10. Probability distributions of the models parameters x, y, Q, z1 and z2 in the last generation in PSO 30th, iteration ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)



Rys. 11. Rozkład prawdopodobieństw wartości poszukiwanych parametrów modelu w 60. iteracji algorytmu PSO ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)

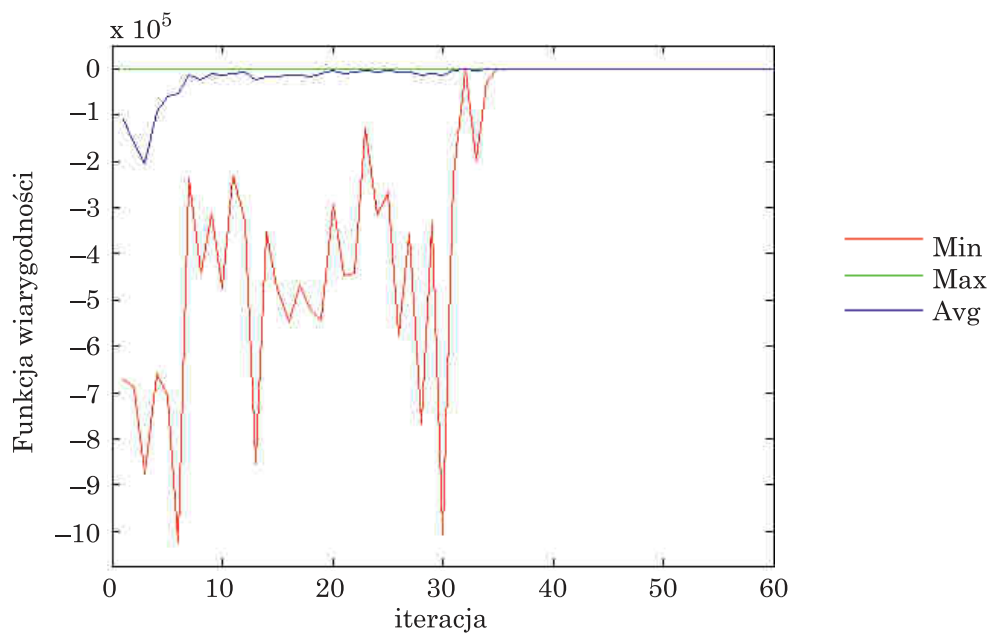
Fig. 11. Probability distributions of the models parameters x , y , Q , z_1 and z_2 in the last generations in the PSO 60th, iteration ($c_1=0.8$, $c_2=1.3$, $c_3=1.5$)

Uzyskane wyniki dla wytypowanej konfiguracji PSO prezentują rysunki 8, 9, 10, 11. Rysunek 8 obrazuje działanie algorytmu na płaszczyźnie x,y w 1, 10, 30 i 60. iteracji. Widocznie jest, że w pierwszej iteracji przestrzeń poszukiwań jest równomiernie wypełniona przez cząstki. Położenie źródła skażenia oznaczone jest kolorem czerwonym, natomiast położenie cząstek kolorem zielonym. Pojedynczy niebieski punkt to według funkcji wiarygodności, położenie cząstki spełniającej wymagania globalnego optimum, w danej iteracji.

Analizując kolejne iteracje łatwo spostrzec, iż rój skutecznie skupia się w kierunku „lidera”. Płynnie dążąc w kierunku celu, algorytm realizuje wyznaczone zadanie i w 60. iteracji znajduje prawidłowo lokalizację źródła uwolnienia.

W celu lepszego zobrazowania wyników rekonstrukcji z wykorzystaniem algorytmu PSO w kolejnych iteracjach stworzono histogramy wartości poszczególnych parametrów w oparciu o położenie cząstek z danej populacji. Jako najbardziej prawdopodobną wartość parametru przyjęto środkową wartość słupka histogramu o najwyższym prawdopodobieństwie. Z kolei jako błąd oszacowania połowę szerokości słupka histogramu. Analizując Rysunki 9, 10 oraz 11 oczywistym jest polepszanie się jakości położenia cząstek w 1, 30 i 60. generacji. Docelowe wartości poszukiwanych parametrów przyjęte w trakcie generowania danych syntetycznych oznaczone zostały na histogramach czerwoną pionową linią.

Rysunek 9 przedstawia rozkłady uzyskane w pierwszej iteracji, w której następową inicjalizacja roju. Jak widać cząstki rozłożone są równomiernie po całej domenie. W kolejnych iteracjach algorytmu PSO cząstki poprawiały odpowiednio swoje położenie i już w 30. iteracji (rysunek 10) widać, że większość roju skupiona jest w miejscach docelowych, oznaczonych przez czerwone linie. Rozkład prawdopodobieństw poszczególnych parametrów modelu zmienia się wraz ze wzrostem iteracji i w 30. iteracji największe prawdopodobieństwo dla parametru x wynosi $P(x=3079.4\pm 136)=0.85$, dla y $P(y=7990.9\pm 136)=0.89$, siła uwolnienia $P(Q=5263.1\pm 90)=0.88$, oraz współczynniki dyfuzji $P(z1=0.2230\pm 0.006)=0.92$, $P(z2=0.2055\pm 0.006)=0.90$. Ostatecznie (rys. 11) wartość prawdopodobieństwa w ostatniej iteracji, dla każdego parametru wynosi 1 (100%), co oznacza, że wszystkie cząstki znalazły się w tym samym położeniu w przestrzeni poszukiwanych parametrów. Finalnie, algorytm PSO określił jako najbardziej prawdopodobne następujące wartości poszukiwanych parametrów $x=3068.2\pm 1500m$, $y=7992.1\pm 1500m$, $Q=5188.0\pm 1000g/s$ $z1=0.2226\pm 0.035$, $z2=0.2052\pm 0.035$. Natomiast wartość funkcji wiarygodności wyniosła $f(x,y,Q,z1,z2)=-0.2100$. Jak widać wszystkie parametry zostały odnaleziona prawidłowo. Przebieg wartości funkcji wiarygodności dla powyższego przykładu prezentuje rysunek 12.



Rys. 12. Przebieg wartości funkcji wiarygodności w kolejnych iteracjach dla parametrów $c1=0.8$, $c2=1.3$, $c3=1.5$, Populacja=100 osobników

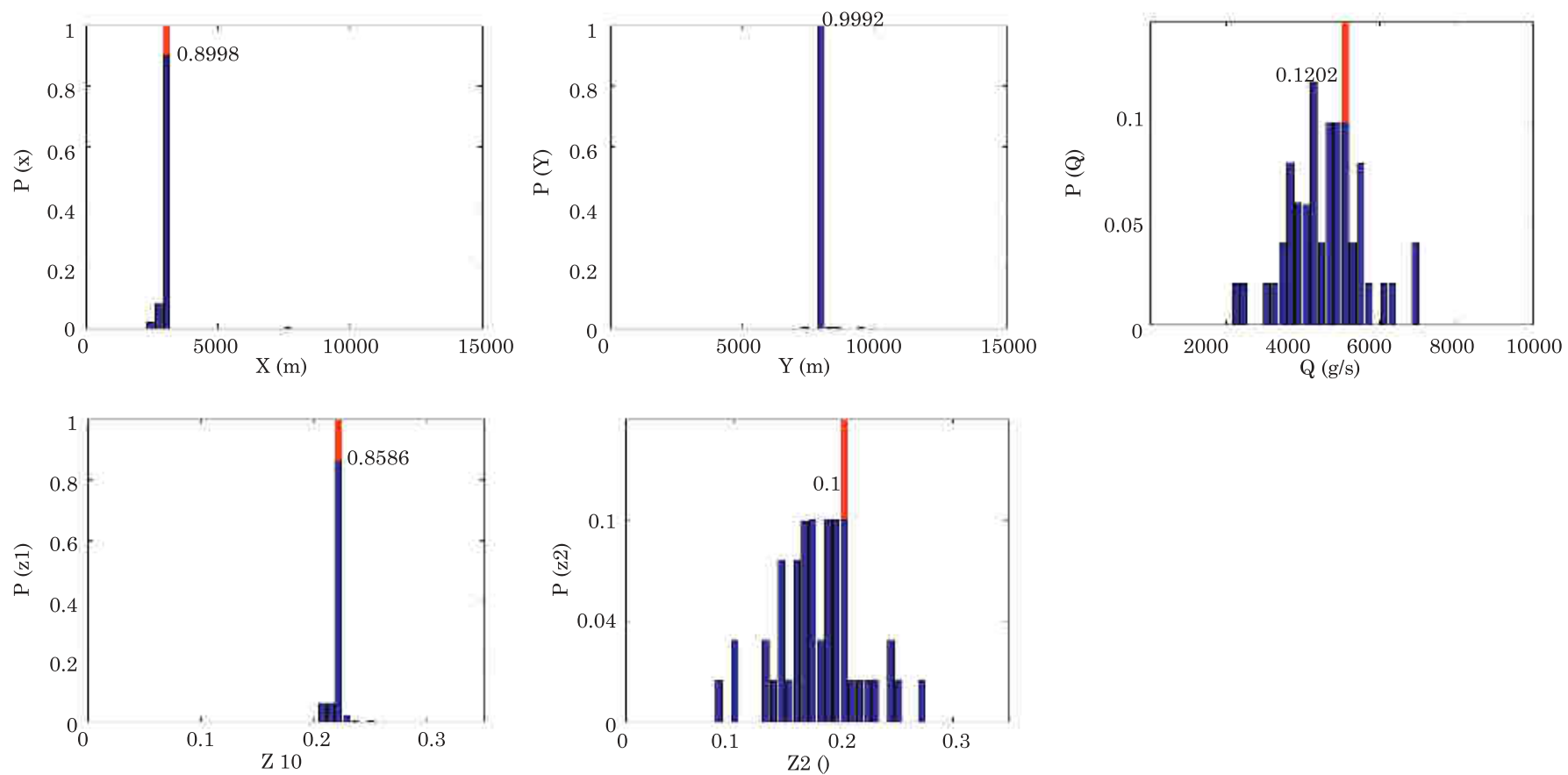
Fig. 12. The average objective function values obtained in the subsequent iterations for parameters: $c1=0.8$, $c2=1.3$, $c3=1.5$, Population size=100 particles

Symulacja obrazująca jedno wybrane uruchomienie algorytmu może nie być wystarczającym dowodem na poprawność działania modelu. W związku z tym przedstawiony algorytm został uruchomiony niezależnie 50 razy. Rysunek 13 ilustruje rozkład prawdopodobieństw oszacowanych wartości poszukiwanych parametrów uśrednionych w oparciu o wyniki z 50. niezależnych uruchomień. Wartości docelowe oznaczone zostały czerwoną pionową linią.

Powyższe histogramy przedstawiają uruchomienie modelu dla identycznej konfiguracji parametrów co w wcześniejszych testach. Oczywistym jest fakt, że średnie wartości algorytmu są gorsze niż podane z jednego wybranego przykładu. Jednakże przedstawione wykresy skutecznie wspierają lokalizację współrzędnych źródła $P(x=3260.9\pm 136)=0.88$, $P(y=8043.5\pm 136)=0.99$ oraz współczynnika dyfuzji atmosferycznej w kierunku poziomym $P(z1=0.2283\pm 0.006)=0.72$. Pozostałe dwa parametry nie są tak sprawnie zgrupowane jak poprzednie, aczkolwiek największe prawdopodobieństwo występuje blisko wartości poszukiwanej $P(z2=0.1978\pm 0.006)=0.10$, $P(Q=4202.9\pm 90)=0.12$.

PODSUMOWANIE

Algorytm Roju Cząsteczek jest stosunkowo nową metodą optymalizacji opartą na zachowaniu stadnych organizmów. Do tej pory był wykorzystywany w wielu zagadnieniach optymalizacji i cieszy się rosnącym zainteresowaniem. Dzięki swojej prostocie daje możliwość wielu modyfikacji zależnie od problemu. Według naszej opinii możemy śmiało stwierdzić, że w przypadku lokalizacji źródła skażeń atmosferycznych również



Rys. 13. Rozkład prawdopodobieństw wartości poszukiwanych parametrów modelu M dla 50 uruchomień modelu w ostatniej 60. iteracji
 Fig. 13. Probability distributions of the models parameters M for the 50th generation in the PSO last, 60th iteration

okazał się niemierniej skuteczny. Posiadając jedynie wiedzę na temat kierunku wiatru i danych z sieci detektorów, wykorzystując odpowiedni model dyspersji podołał zadaniu i skutecznie znalazł wszystkie pięć poszukiwanych parametrów.

PODZIĘKOWANIA

Niniejszy artykuł powstał w ramach programu Welcome Fundacji na Rzecz Nauki Polskiej, finansowanego z Funduszy Unii Europejskiej w ramach Programu Operacyjnego Innowacyjna Gospodarka na lata 2007–2013.

PIŚMIENNICTWO

- Borysiewicz M., Wawrzyńczak A., Kopka P., 2012a. Stochastic algorithm for estimation of the model's unknown parameters via Bayesian inference, *Proceedings of the Federated Conference on Computer Science and Information Systems*, 501–508, IEEE Press.
- Borysiewicz M., Wawrzyńczak A., Kopka P., 2012b. Bayesian-Based Methods for the Estimation of the Unknown Model's Parameters in the Case of the Localization of the Atmospheric Contamination Source, *Foundations of Computing and Decision Sciences*, 37, 4, 253–270.
- Cai X., Cui Z., Zeng J., Tan Y., 2009. Individual Parameter Selection Strategy for Particle Swarm Optimization, *Taiyuan University of Science and Technology*.
- Chmiel W., Kadłuczka P., Packanik G., 2011. Zastosowanie algorytmów rojowych w rozwiązywaniu zagadnień permutacyjnych, *Automatyka / Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie*, T. 15, z. 2, 117–126.
- Filipowicz B, Kwiecień J., 2011. Algorytmy stadne w problemach optymalizacji, *Przegląd Elektrotechniczny*, AGH Akademia Górniczo-Hutnicza w Krakowie, R. 90 NR 6/2014.
- Foryś P., 2008. Zastosowanie metody roju cząstek w optymalnym projektowaniu elementów konstrukcji, *Politechnika Krakowska*, 4-M.
- Gilks W., Richardson S., Spiegelhalter D., 1996. *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.
- Kennedy J., Eberhart R., 1995. *Particle Swarm Optimization*, Purdue School of Engineering and Technology, Washington.
- Marchi B., Funtowicz S., Ravetz J., 1996. *The Long Road to Recovery: Community Responses to Industrial Disaster, Seveso: A paradoxical classic disaster*, United Nations University, Tokyo, Paris, New York, 1, 4, 307.
- Markiewicz M. T., 2004. *Podstawy modelowania rozprzestrzeniania się zanieczyszczeń w powietrzu atmosferycznym*, Oficyna Wydawnicza Politechniki Warszawskiej.
- Michalczyk J. K., 2003. *Transport gazowych zanieczyszczeń w powietrzu – symulacje numeryczne w skali lokalnej*. Politechnika Lubelska.
- Palikowska K., 2013. Modelowanie krzywizny układu geometrycznego toru z wykorzystaniem algorytmu PSO, *Wydział Transportu / Politechnika Gdańska*, Zeszyt 98, 509-519.
- Pasquill, F., 1961. The estimation of the dispersion of windborne material, *The Meteorological Magazine*, vol 90, No. 1063, pp 33-49.
- Reyes Sierra M., Coello Coello A. C., 2005. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and ρ -Dominance, Mexico.
- Turner D. Bruce, 1994. *Workbook of Atmospheric Dispersion Estimates*, Lewis Publishers, USA.

Wawrzynczak A., Kopka P., Borysiewicz M., 2014. Sequential Monte Carlo in Bayesian assessment of contaminant source localization based on the distributed sensors measurements, Lecture Notes in Computer Sciences 8385, PPAM 2013, Part II, ch. 38.

Wawrzynczak A., Jaroszyński M., Borysiewicz M., 2014. Data-driven Genetic Algorithm in Bayesian estimation of the abrupt atmospheric contamination, Federated Conference on Computer Science and Information Systems (FedCSIS), Warsaw.

THE LOCALIZATION OF ATMOSPHERIC POLLUTION SOURCE BY MEANS OF PARTICLE SWARM OPTIMIZATION

Abstract. In this paper, we applied the Particle Swarm Optimization (PSO) to the problem of the atmospheric pollution source localization. Localization is performed based on the information about concentration of the released substance in the atmosphere, recorded by a network of detectors deployed in the area. PSO algorithm searches the five-dimensional space in order to find the dispersion model parameters. The parameters are: the contaminant source coordinates (X , Y), release strength (Q) and the atmospheric transport dispersion coefficients (σ_y , σ_z). In the reconstruction, we employed the fast-running Gaussian plume dispersion model. We presented a sample configuration of the PSO algorithm, which task is to indicate the most likely parameters of contamination source in the shortest time. The efficiency of the algorithm was tested on the basis of the generated synthetic data.

Keywords: Particle Swarm Optimization, atmospheric transport pollution dispersion, contaminant source localization.

Zaakceptowano do druku – Accepted for print: 20.02.2014