

ZASTOSOWANIE ALGORYTMÓW GENETYCZNYCH DO WALIDACJI MODELI NUMERYCZNYCH TRANSPORTU CIEPŁA W GLEBIE*

K. Lamorski

Instytut Agrofizyki PAN, ul. Doświadczalna 4, 20-290 Lublin 27

S t r e s z c z e n i e: W pracy przedstawiono wyniki eksperymentów z zastosowaniem algorytmów genetycznych do zagadnień odwrotnych i walidacji modeli fizycznych transportu energii w glebie. Pokazano że algorytmy genetyczne są odpowiednią metodą obliczeniową do stosowania w opisywanych zagadnieniach.

S ł o w a k l u c z o w e: algorytmy genetyczne, walidacja, zagadnienie odwrotne, modelowanie.

WSTĘP

Modelowanie transportu wody i energii w glebie jest użytecznym narzędziem w badaniach agrofizycznych. Badania modelowe pozwalają poznać przebieg procesów transportu w ośrodku glebowym. Eksperymenty natomiast stanowią weryfikacyjny punkt odniesienia dla wyników modelowania. Często też w wyniku zjawisk obserwowanych w eksperymencie, których nie ma w wynikach symulacyjnych modyfikowany jest model opisujący dany obiekt. Eksperyment i opis fizyczny, modelowanie wzajemnie się uzupełniają, umożliwiając możliwie pełne poznanie danego obiektu fizycznego.

W przypadku podstawowych równań opisujących transport energii w glebie, zakładamy że właściwości termiczne ośrodka są reprezentowane przez współczynnik przewodnictwa cieplnego λ [W/mK] i objętościową pojemność cieplną C_v [J/m³ K]. Obie wielkości są w ogólności funkcjami współrzędnych przestrzennych oraz temperatury i wilgotności gleby w danym punkcie i wraz z warunkami brzegowymi determinują przebieg procesu transportu w glebie. W powyższym opisie

*Praca była częściowo finansowana przez Komitet Badań Naukowych w ramach projektu badawczego nr 6-P06B-043-20.

transportu energii nie bierzemy pod uwagę efektów energetycznych związanych z jednoczesnym transportem nieizotermicznym pary wodnej. Poniższe równanie opisuje w przyjętym modelu transport energii w glebie [2,3]:

$$C_v \frac{\partial T}{\partial t} = \nabla(\lambda \nabla T). \quad [1]$$

Do modelowania transportu energii w glebie niezbędna jest znajomość właściwości termicznych ośrodka. Ze względu na trudności pomiarowe czasami nie są znane parametry opisujące właściwości termiczne ośrodka. W takiej sytuacji można wyznaczyć przybliżone wartości parametrów termicznych ośrodka drogą pośrednią, poprzez walidację modelu numerycznego na podstawie eksperymentu pomiarowego wykonanego na opisywanym ośrodku.

W pierwszej fazie eksperymentu numerycznego poszukujemy wartości parametrów opisujących ośrodek na drodze porównań wartości temperatur generowanych przez model z temperaturami zmierzonymi w realnym ośrodku. Parametry ośrodka uzyskane tą drogą uważa się za poprawnie reprezentujące badany obiekt gdy dane pochodzące z obliczeń pokrywają się z danymi pomiarowymi.

W drugiej fazie na podstawie modelu po walidacji wykonywać można kolejne symulacje numeryczne zjawisk zachodzących w ośrodku.

Rozwiązanie zagadnienia odwrotnego wiąże się zazwyczaj z szukaniem ekstremum pewnej funkcji, której argumentami są wartości szukanych parametrów. Funkcja ta opisuje stopień zgodności przebiegów temperatury wyznaczonych przez model z przebiegami pochodzącymi z eksperymentu. W zależności od konkretnego zagadnienia funkcja ta będzie zależała od kilku zmiennych, reprezentujących poszczególne parametry modelu których wartości szukamy. Szukanie globalnego ekstremum funkcji zależnej od kilku zmiennych jest zagadnieniem trudnym do praktycznego rozwiązania. Nie istnieje możliwość stosowania metod analitycznych ponieważ wartości funkcji są wyliczane na podstawie danych wygenerowanych przez model każdorazowo dla różnych wartości jej argumentów, nie istnieje wobec tego analityczna postać funkcji. Pozostaje stosowanie komputerowych metod obliczeniowych. Istnieje kilka metod komputerowych stosowanych w zagadnieniach optymalizacji, są to między innymi:

- metoda gradientowa,
- metoda losowa,
- metoda symulowanego wyżarzania,
- algorytmy genetyczne.

Jak wykazują doświadczenia do najbardziej skutecznych metod optymalizacji funkcji wielowymiarowych należą algorytmy genetyczne.

MATERIAŁ I METODA

Koncepcja algorytmów genetycznych została po raz pierwszy zaprezentowana w latach 70-tych. Rozwijana była początkowo jako formalizm matematyczny służący opisowi zjawisk ewolucyjnych w badaniach biologicznych. Szybko zaczęto stosować sam formalizm algorytmów genetycznych do rozwiązywania zadań optymalizacyjnych w praktycznych zagadnieniach z różnych dziedzin nauki i techniki [1].

Algorytmy genetyczne są metodą optymalizacji, oznacza to w najogólniejszym sensie, że służą szukaniu ekstremum funkcji. Jeżeli rzeczywiste zagadnienie którego rozwiązania poszukujemy wiąże się z poszukiwaniem minimum funkcji to trzeba je przeformułować tak aby algorytmy genetyczne operowały na funkcji która ma maksimum tam gdzie oryginalna funkcja miała minimum. Podstawowym krokiem w kierunku rozwiązania zagadnienia z zastosowaniem metod algorytmów genetycznych jest odpowiednie zbudowanie funkcji celu.

Funkcja celu osiąga maksimum dla pewnych wartości swoich zmiennych. Te wartości zmiennych odpowiadające maksimum funkcji są poszukiwanymi wartościami parametrów rozwiązywanego zagadnienia. Wartości parametrów od których zależy funkcja celu są w algorytmach genetycznych kodowane w postaci ciągów binarnych. Konkretna postać kodowania nie jest istotna dla opisu działania algorytmów genetycznych. Istotne jest to, że takie kodowanie jest określone i że istnieje kodowanie odwrotne pozwalające na odtworzenie z zakodowanej binarnej postaci ciągów, odpowiadających im wartości parametrów opisujących badane zagadnienie.

Kolejną istotną cechą algorytmów genetycznych jest to, że operują one jednocześnie na całym zbiorze parametrów w postaci zakodowanej. Zbiór ten jest nazywany populacją, ciągi binarne zakodowanych parametrów problemu nazywają się w nomenklaturze algorytmów genetycznych osobnikami populacji.

Pierwszym krokiem w algorytmie genetycznym wykonywanym jednorazowo na początku obliczeń jest zainicjowanie wartościami losowymi populacji. W tym celu należy wylosować wartości parametrów i zakodować je dla każdego osobnika populacji. W wyniku tej operacji otrzymamy zbiór różnych ciągów binarnych reprezentujących poszczególnych osobników. Z każdym osobnikiem skojarzony jest pełen zestaw szukanych parametrów.

Następnie wykonywane są sekwencyjnie operacje reprodukcji, krzyżowania i mutacji populacji. Operacja reprodukcji polega na wygenerowaniu na podstawie

istniejącej populacji, populacji nowych osobników poprzez powielanie osobników ze starej populacji. Procentowy udział starych osobników w nowej populacji, jest proporcjonalny do wartości funkcji celu odpowiadającej wartościom parametrów reprezentowanym przez ciągi binarne poszczególnych osobników. W wyniku reprodukcji powstaje populacja równoliczna z wyjściową, składająca się z takich samych osobników lecz różniąca się ilością poszczególnych osobników w populacji. Ta część algorytmu genetycznego jest deterministyczna i nie ma w niej elementów probabilistycznych.

Na wygenerowanej w procesie reprodukcji grupie osobników jest następnie wykonywana operacja krzyżowania. Polega ona na losowym wyborze par osobników w ramach populacji. Następnie dla każdej z par losowane jest miejsce w ciągu binarnym reprezentującym danego osobnika. Następnie zawartości bitów starszych niż wylosowana pozycja w ciągach binarnych są wymieniane między osobnikami w ramach z każdej z par w populacji. W wyniku tej losowej operacji otrzymujemy nową populację która jest już szukaną populacją potomną w danym kroku iteracji.

Opcjonalną operacją jaką można jeszcze wykonać na populacji jest jej mutacja. Operacja mutacji polega na zupełnie przypadkowej zamianie losowo wybranych bitów w nielicznych losowo wybranych osobnikach populacji wygenerowanej w procesie reprodukcji i krzyżowania.

Proces ewolucji populacji jest przeprowadzany tak długo aż zostanie znaleziony osobnik, który maksymalizuje funkcję celu. W praktyce stosowane jest również inne kryterium przzerwania obliczeń, polega ono na sprawdzeniu czy zmienia się dostosowanie (wartość funkcji celu) najlepiej przystosowanego osobnika w populacji. Jeżeli wartość funkcji celu dla najlepszego osobnika w kolejnych populacjach nie zmienia się oznacza to że algorytm znalazł maksimum funkcji celu.

Algorytmy genetyczne zyskały szerokie zastosowanie dzięki ich następującym cechom:

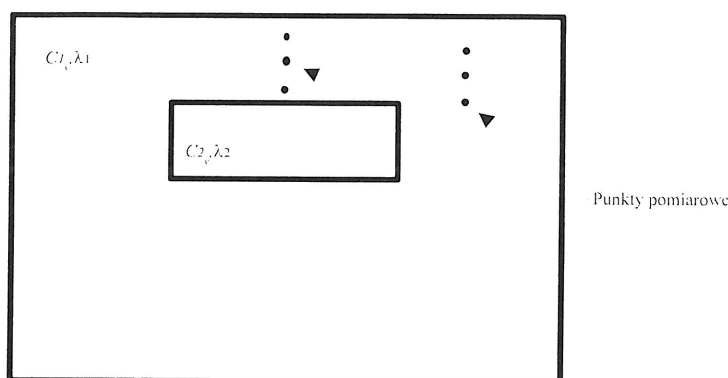
- korzystają tylko z funkcji celu, nie jest potrzebna znajomość pochodnych funkcji celu;
- prowadzą obliczenia jednocześnie dla całego zespołu (populacji) wartości szukanych parametrów;
- nie przetwarzają bezpośrednio parametrów zadania lecz operują na jej zakodowanej formie;
- stosują losowe a nie deterministyczne reguły wyboru parametrów optymalnych.

Dzięki tym cechom algorytmów genetycznych są one bardziej odporne od innych metod optymalizacji na utknięcie w maksimum lokalnym funkcji celu. Algorytmy genetyczne prowadzą do znalezienia maksimum globalnego funkcji celu.

Dzięki prowadzeniu obliczeń dla całej populacji wartości parametrów można algorytmy genetyczne łatwo zaimplementować i efektywnie wykorzystywać w wieloprocesorowym środowisku komputerowym.

Do wykorzystania algorytmów genetycznych jest konieczna tylko znajomość wartości funkcji celu. Pozwala to na stosowanie tych metod do rozwiązywania szerokiej gamy zagadnień. Wartość funkcji celu nie musi być znana w postaci wyrażenia algebraicznego, można ją obliczyć np. na podstawie wyników wygenerowanych przez inny program (np. symulacyjny).

Weźmy pod uwagę jednorodny ośrodek glebowy w którym umieszczono zaburzenie o parametrach termicznych różnych od parametrów termicznych ośrodka. Taką geometrię ośrodka przedstawia rys. 1.



Rys. 1. Schemat obszaru symulowanego zagadnienia.

Fig. 1. Scheme of simulated area.

Przy zadanych warunkach brzegowych temperatura w punktach pomiarowych będzie zależała tylko od właściwości termicznych ośrodka (C_{1v}, λ_1) i wtrącenia (C_{2v}, λ_2).

W badanym zagadnieniu praktycznym program symulacyjny obliczał wartości temperatury w wybranych punktach profilu glebowego. Znane były z eksperymentu wartości temperatur zmierzone w tych punktach. Nie były znane właściwości przewodnictwa cieplnego i pojemności cieplnej ciała umieszczonego w jednorodnym ośrodku glebowym jak i samego ośrodka. Aby rozwiązać problem znalezienia wartości szukanych parametrów z wykorzystaniem algorytmów genetycznych, należy zbudować funkcję celu. Miarą dopasowania danych wyliczonych przez model do danych eksperymentalnych może być średnie odchylenie standardowe wyliczane wg następującej formuły:

$$S_1 = \sum_{i=1}^N (T_i^p - T_i^s)^2 \quad [2]$$

gdzie: T_i^p jest temperaturą zmierzoną w i -tym punkcie profilu glebowego a T_i^s jest temperaturą wyliczoną w tym samym punkcie. Im mniejsza będzie wartość S_1 tym dane wyliczane na podstawie modelu będą bliższe danym pomiarowym. Niestety znalezienie rozwiązania tego problemu polega na minimalizacji funkcji S_1 , algorytmy genetyczne natomiast wymagają takiego sformułowania problemu by rozwiązanie optymalne wiązało się z poszukiwaniem maksimum funkcji. Wobec tego zamiast minimalizować funkcję S_1 można maksymalizować funkcję następującej postaci:

$$S_2 = C - S_1 \quad [3]$$

która osiąga maksimum dla tych wartości zmiennych dla których funkcja S_1 ma minimum. Wielkość C jest stałą. Argumentami funkcji S_2 która jest szukaną funkcją celu są wartości parametrów $\lambda_1, C_{v1}, \lambda_2, C_{v2}$ opisujących właściwości termiczne ośrodka glebowego. Wartości tych parametrów będą kodowane w algorytmie genetycznym jako pojedyncze ciągi binarne.

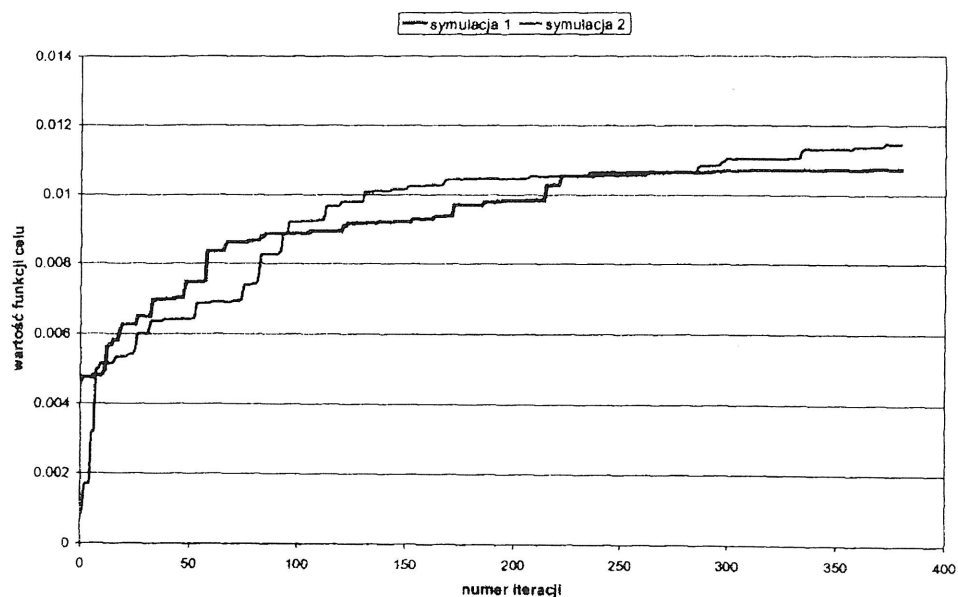
W programie rozwiązującym opisany problem zastosowana została w części obliczeniowej biblioteka GALIB [4] zawierająca zaimplementowane procedury algorytmów genetycznych.

WYNIKI

W celu sprawdzenia poprawności działania zaprezentowanej powyżej metody przeprowadzone zostały obliczenia. Przeprowadzono obliczenia dwukrotnie dla różnych losowo wybieranych wartości początkowych szukanych parametrów termicznych gleby i wtrącenia. Otrzymano zbieżne wyniki wyliczonych tą drogą wartości szukanych parametrów.

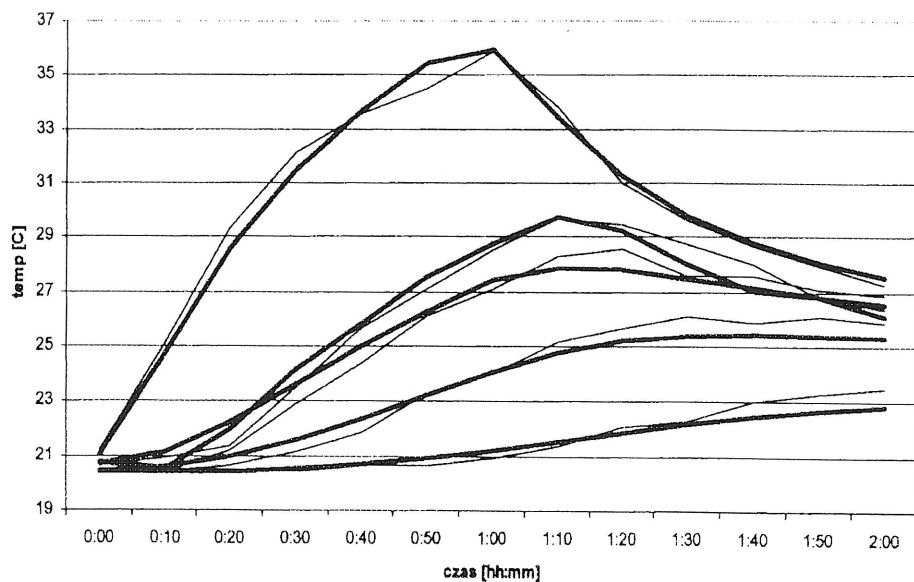
Rysunek 2 przedstawia jak w trakcie obliczeń zmieniała się wartość funkcji celu dla osobnika najlepiej dopasowanego w każdej z populacji.

Wartość średnią wartości parametrów termicznych ośrodka wyznaczono jako średnie wartości pochodzące z poszczególnych serii obliczeniowych. Rysunek przedstawia porównanie danych pochodzących z eksperymentu z danymi wyliczonymi przez model na podstawie parametrów ośrodka pochodzących z wyliczeń przy pomocy algorytmów genetycznych. Linie pogrubione reprezentują wartości temperatury pochodzące z symulacji, linie cienkie przedstawiają wyniki pochodzące z pomiarów. Punkty pomiarowe w których mierzona była temperatura przedstawiona na rys. 3 umieszczone były na głębokości 1 cm, 3 cm, 5 cm oraz 7 cm pod powierzchnią gleby.



Rys. 2. Zmiany wartości funkcji celu dla najlepiej przystosowanego osobnika w trakcie ewolucji populacji.

Fig. 2. Changes of objective function of the best individual due to population evolution.



Rys. 3. Porównanie temperatur wyliczonych i zmierzonych w wybranych punktach profilu glebowego.

Fig. 3. Comparison of evaluated and measured temperatures in selected points of soil profile.

WNIOSKI

W przypadku opierania się w procesie walidacji modelu numerycznego o jeden eksperyment istnieje realna groźba, że wyznaczone parametry opisujące model nie będą rzeczywistymi parametrami poprawnie opisującymi obiekt badań. Aby się tego ustrzec należy w procesie walidacji dysponować możliwie szeroką gamą danych eksperymentalnych dla danego obiektu. Im więcej parametrów musimy wyznaczyć w procesie walidacji tym liczniejsze muszą być dane pomiarowe.

Eksperymenty pokazały użyteczność zastosowań algorytmów genetycznych do zagadnień optymalizacyjnych związanych z walidacją modeli transportu ciepła w ośrodku glebowym.

LITERATURA

1. **Goldberg D.E.:** Algorytmy genetyczne i ich zastosowania. WNT, Warszawa, Polska, 1998.
2. **Hillel D.:** Environmental Soil Physics. Academic Press, London, San Diego, USA, 1998.
3. **Kutilek M., Nielsen D.R.:** Soil Hydrology. Catena Verlag, Cremlingen-Berlin, Germany, 1994.
4. **Wall M.:** GALib: A C++ Library of Genetic Algorithm Components. <http://lancet.mit.edu/ga/>.

USE OF GENETIC ALGORITHMS IN REVERSE PROBLEMS ARISING
FROM SOIL HEAT TRANSPORT MODELS

K. Lamorski

Institute of Agrophysics PAS, ul. Doświadczalna 4, 20-290 Lublin, Poland

A b s t r a c t: The aim of this is to show usefulness of genetic algorithms, as a method of optimization used in problems arising from validation of heat soil transport models and reverse problems. It is shown that genetic algorithms may be used in such problems, they are effective and computationally efficient. Results of sample calculations are presented.

K e y w o r d s: genetic algorithms, validation, reverse problems, modelling.