

Mieczysław M. Kania*, Albert Kubzdela*

ANALIZA PRZEMIESZCZEŃ PALA W SPRĘŻYSTYM OŚRODKU TRÓJWARSTWOWYM PRZY UŻYCIU FUNKCJI MARS

1. Wprowadzenie

Zależności w geotechnice mają na ogół złożony, wielowymiarowy charakter, wynikający z udziału w procesach geotechnicznych szeregu różnych uwarunkowań geometrycznych, materiałowych, czasoprzestrzennych, klimatycznych czy przyrodniczych. Uzyskanie rozwiązań analitycznych dla zagadnień wieloczynnikowych, przy trudnej do określenia teoretycznego istocie zależności między czynnikami składowymi danego procesu, jest dla wielu zadań brzegowych niemożliwe. Alternatywą są wówczas uogólnienia ustalane najczęściej na podstawie uproszczonych doświadczeń w skali modelowej lub w skali naturalnej. Doświadczenia fizyczne w geotechnice są jednak zbyt czasochłonne i kosztowne, aby mogły spełniać rolę narzędzia do poszukiwania zależności funkcyjnych dla wielu zmiennych badanego procesu. Poszukiwanie i badanie charakteru złożonych zależności wielu zmiennych można wówczas prowadzić na drodze symulacji numerycznych, z wykorzystaniem nowoczesnych technik modelowania komputerowego (np. metoda elementów skończonych) i technik wielowymiarowej analizy statystycznej [6].

W artykule przedstawiono przykład rozwiązania zadania oszacowania wpływu różnych czynników geometrycznych i materiałowych na pionowe przemieszczenia głowicy pojedynczego pala fundamentowego, pograżonego w trójwarstwowym ośrodku sprężystym, z jedną warstwą znacznie bardziej odkształcalną od pozostałych. Przemieszczenia obliczone za pomocą metody elementów skończonych, potraktowano jak dane doświadczalne i poddano analizie przy użyciu metody *MARS* (jest to akronim od pełnej nazwy Multivariate Adaptive Regression Splines, co można przetłumaczyć jako Wieloraka Adaptacyjna Regresja Sklejana — przez analogię do polskiego tłumaczenia nazwy funkcji „spline” jako funkcje „sklejane”).

* Zakład Geotechniki i Geologii Inżynierskiej, Instytut Inżynierii Lądowej, Wydział Budownictwa i Inżynierii Środowiska, Politechnika Poznańska, Poznań

2. Metoda *MARS*

Liniowy model regresji wielorakiej pozwala na zbadanie i opisanie liniowej zależności zmiennej zależnej od ustalonej liczby zmiennych niezależnych w oparciu o dany zbiór obserwacji. Wynikiem jest zależność opisana formułą liniową:

$$y = b_0 + \sum_{i=1}^n b_i x_i \quad (1)$$

gdzie b_0, \dots, b_n są parametrami modelu.

Model regresji powierzchni odpowiedzi, jako przykład ogólnego modelu liniowego (GLM), pozwala na opisanie również wielu krzywoliniowych zależności pomiędzy zmiennymi. W modelu kwadratowym powierzchni odpowiedzi, badaniu podlega nie tylko wpływ na zmienną zależną efektów głównych dla wielkości wejściowych (X_1, \dots, X_k), lecz także ich dwuczynnikowej interakcji ($X_1 X_2, X_1 X_3, \dots, X_{k-1} X_k$), oraz wyrazów kwadratowych (X_{12}, \dots, X_{k2}) (patrz [2]). Równanie regresji dla modelu kwadratowej regresji powierzchni odpowiedzi dla dwóch zmiennych objaśniających ciągłych X_1, X_2 miałoby postać:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2 + b_5 x_1 x_2 \quad (2)$$

Metoda *MARS* (Multivariate Adaptive Regression Splines) jest jednym z wielu narzędzi statystycznej eksploracji danych (*data mining*) [3], służącym między innymi do rozwiązywania zagadnień regresyjnych. Jako model nieparametryczny, nie wymaga założeń dotyczących kształtu zależności funkcyjnej między zmiennymi zależnymi a niezależnymi. Zależność ta jest ustalana jako kombinacja liniowa funkcji bazowych, wyłącznie na podstawie informacji zawartych w danych wejściowych. Dane dzielone są na obszary w których niezależnie definiowane są osobne funkcje regresyjne. Narzędzie to jest szczególnie cenne przy większej liczbie wymiarów i dużych zbiorach danych, o złożonych zależnościach między nimi. Funkcja regresji definiowana jest jako funkcja sklejana powstała z kombinacji liniowej funkcji bazowych. Funkcjami bazowym są dwustronne, obcięte funkcje liniowe, wyprowadzane w oparciu o metodę najmniejszych kwadratów. Zależność dana jest ogólnym równaniem:

$$f(X) = b_0 + \sum_{i=1}^n b_i x_i(X) \quad (3)$$

gdzie:

X — oznacza wektor zmiennych predykcyjnych oraz ich interakcji,

h_i — funkcje bazowe,

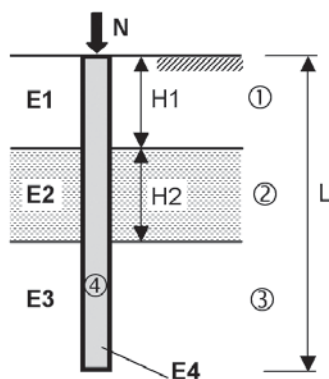
b_i — wyznaczone współczynniki.

Algorytm obliczeń ma przebieg dwuetapowy. Wpierw przeszukiwany jest zbiór danych wartości wejściowych. Do modelu dodawane są kolejne funkcje bazowe, wybierane ze zbioru

wszystkich dopuszczalnych funkcji, tak by zmaksymalizować ogólny poziom dopasowania. Następnie zostaje przeprowadzone upraszczanie modelu, poprzez redukcję liczby funkcji bazowych i poprzez eliminację najmniej znaczących funkcji. W obu etapach o stopniu złożoności modelu, a następnie stopniu redukcji decyduje użytkownik, określając maksymalną liczbę funkcji bazowych, rząd interakcji pomiędzy zmiennymi niezależnymi, wartość kary za dodanie do modelu kolejnej funkcji bazowej, próg zabezpieczający przed nadmiernym dopasowaniem modelu. Miara dopasowania modelu określana jest wartością współczynnika *R*-kwadrat (patrz [1], [3]).

3. Analiza współpracy pała w ośrodkiem trójwarstwowym

Celem prezentowanego przykładowego rozwiązania była analiza zależności wielkości przemieszczeń głowicy pała od wybranych czynników. Analizie poddano pał o długości $L = 12$ m i stałej średnicy $D = 0,6$ m, poddany obciążeniu osiowemu równemu $N = 1$ kN, zagłębiony w trójwarstwowym podłożu gruntowym. Założono przy tym, że układ warstw będzie odwzorowywał typowe dla zastosowań pali warunki geotechniczne: ze słabą warstwą nr 2, spoczywającą na mocnej warstwie nr 3.



Rys. 1. Pał w trójwarstwowym podłożu gruntowym

Dane do analizy *MARS* zaczerpnięto z prac [5, 6], w postaci pliku 300 różnych wyników przemieszczeń otrzymanych z obliczeń MES dla modelu liniowo sprężystego, obejmujących zestawy danych z kombinacjami 4 różnych wielkości $H1$ i 5 różnych wielkości $H2$ oraz odpowiednio 3 różnych wielkości $E1$ i 5 wielkości $E2$. Pozostałe warunki brzegowe zadania przyjęto jako stałe. Rozpatrywane różne grubości warstw i wielkości modułu sprężystości Younga dla warstwy górnej i środkowej podłoża pokazano w tabeli 1.

Analizie poddano 300 rezultatów symulacji komputerowej współpracy pała z trójwarstwowym ośrodkiem gruntowym według rysunku 1 i tabeli 1, poszukując zależności między wielkością pionowych przemieszczeń głowicy pała, a kombinacją wielkości wybranych czterech czynników bezwymiarowych [6]:

TABELA 1
Parametry podłoża gruntowego i analizowanego pala

Nr warstwy	Grubość warstwy H [m]	Moduł Younga E [MPa]	Współczynnik Poissona ν
1	2,4; 3,6; 4,8; 6,0	60, 80, 100	0,25
2	1,2; 1,8; 2,4; 3,0; 3,6	10, 20, 30, 40, 50	0,10
3	—	100	0,25
pal	—	30000	0,15

$$X_1 = \frac{H1}{L}, \quad X_2 = \frac{H2}{L}, \quad X_3 = \frac{E1}{E3}, \quad X_4 = \frac{E2}{E3}, \quad (4)$$

Do ustalenia zależności wielkości przemieszczeń głowicy pala od zdefiniowanych czynników wykorzystano model kwadratowej regresji powierzchni odpowiedzi oraz MARSplines. Obliczenia wykonano za pomocą pakietu statystycznego Statistica 9.0 [7]. Przy generowaniu modelu za pomocą MARSplines, posłużono się również pakietem Salford Systems Predictive Mining Suite [8].

Opracowując model kwadratowej regresji powierzchni odpowiedzi przy użyciu pakietu Statistica stwierdzono, iż przy ustalonym poziomie istotności $p = 0,05$, statystycznie nieistotne okazały się współczynniki przy X_1^2 oraz przy interakcji $X_1 X_2$. Następnie, stosując regresję krokową wsteczną, eliminowano z modelu zmienne mające najmniej istotny wpływ na wartość przewidywanych przemieszczeń. Ostatecznie równanie regresji przyjęło postać:

$$przem = b_0 + b_1 X_2 + b_2 X_2 X_4 + b_3 X_1 + b_4 X_1 X_3 \quad (5)$$

dla podanych w tabeli 2 wartości parametrów zadania. Dla analizowanego modelu miara jego dopasowania wynosi R -kwadrat = 0,98118. Dla porównania, równanie regresji wielorakiej wyznaczonej jako kombinacja liniowa parametrów X_1, X_2, X_3, X_4 uzyskało „jakość” z miarą R -kwadrat = 0,91357.

TABELA 2
Wartości współczynników regresji dla równania (1)

Zmienne	Znormalizowane współczynniki regresji	Współczynniki równania regresji
wyraz wolny		1,49E-06
X_2	0,8860	1,80E-06
$X_2 X_4$	-0,7292	-2,85E-06
X_1	1,0822	1,39E-06
$X_1 X_3$	-1,1169	-1,49E-06

Tworząc model przy użyciu biblioteki MARSplines pakietu Statistica, wygenerowano funkcję sklejaną. Do konstrukcji funkcji bazowych wykorzystano 4 czynniki główne X_1, X_2, X_3, X_4 oraz ich interakcje. Otrzymano równanie regresji w postaci:

$$\begin{aligned}
 przem = & b_0 + b_1 \cdot \max(0; X_3 - 0,6) + b_2 \cdot \max(0; X_4 - 0,1) + \\
 & + b_3 \cdot \max(0; X_2 - 0,1) + b_4 \cdot \max(0; X_1 - 0,1) \cdot \max(0; X_4 - 0,1) + \\
 & + b_5 \cdot \max(0; X_1 - 0,2) + b_6 \cdot \max(0; X_1 - 0,2) \cdot \max(0; X_3 - 0,6) + \\
 & + b_7 \cdot \max(0; X_2 - 0,1) \cdot \max(0; X_3 - 0,6) + \\
 & + b_8 \cdot \max(0; X_1 - 0,2) \cdot \max(0; X_4 - 0,1) + \\
 & + b_9 \cdot \max(0; X_3 - 0,6) \cdot \max(0; X_4 - 0,1)
 \end{aligned} \tag{6}$$

gdzie współczynniki b_i ($i = 0, \dots, 9$) zapisane jako wektor wynoszą odpowiednio:

$$b = [17,3; 2,81; -3,05; 15,6; -26,6; 5,24; 14,5; -4,53; -1,72; 1,09] \cdot 10^{-7} \tag{4}$$

Wykorzystano w nim 14 funkcji bazowych, opartych o wyselekcjonowane 10 zmiennych spośród 4 czynników głównych i ich interakcji. Ocena dopasowania modelu utworzonego w pakiecie MARSplines — Statistica wyniosła R -kwadrat = 0,984. Lepszą jakość dopasowania modelu otrzymano generując równanie regresji przy użyciu pakietu Salford Systems Predictive Mining Suite. Zależność miary R -kwadrat od liczby wykorzystanych w modelu funkcji bazowych przedstawia tabela 3. Jako optymalny wskazany został model oparty o 14 funkcji bazowych.

TABELA 3
Zależność R -kwadrat od liczby funkcji bazowych użytych w modelu MARS zbudowanym w pakiecie Salford Systems Predictive Mining Suite

Liczba użytych funkcji bazowych	Liczba czynników głównych	R -kwadrat
15	4	0,99848
14	4	0,99851
13	4	0,99844
12	4	0,99813
11	4	0,99778
10	4	0,99699
9	4	0,99573
8	4	0,98998
7	4	0,98103
6	4	0,97143
5	4	0,96727
4	3	0,91792
3	3	0,86440
2	3	0,82198
1	2	0,46565

Zbudowany model MARS pozwala również na procentową ocenę istotności poszczególnych zmiennych, w zakresie ich wpływu na prognozowaną wielkość. Jak widać z zestawienia w tabeli 4, w przedstawionym modelu wartość przemieszczeń głowicy pała w największym stopniu zależy od czynników X_3 i X_4 .

TABELA 4
**Istotność poszczególnych czynników w modelu MARS
 zbudowanym w pakiecie Salford Systems Predictive Mining Suite**

zmiennie	istotność zmiennej w modelu
X_3	100,00 %
X_4	97,20 %
X_2	82,25 %
X_1	45,18 %

4. Podsumowanie

Zaprezentowany przykład zastosowania metody MARS, jako jednej z predykcyjnych technik eksploracji danych, dotyczy zadania geotechnicznego, w którym mamy do czynienia ze stosunkowo niewielką liczbą danych. Pokazuje jednak możliwości poszukiwania uogólnień i relacji między zmiennymi w zadaniach silnie wielowymiarowych, dla których nie ma dotychczas wystarczająco dokładnych rozwiązań teoretycznych. Jest to podejście atrakcyjne również ze względu na swój uniwersalizm, czyli brak szczególnych ograniczeń i wymagań co do rodzaju badanych zależności i rozkładów zmiennych. Połączenie efektywności metody elementów skończonych, jako narzędzia do realizacji wieloczynnikowych eksperymentów numerycznych oraz współczesnych technik statystycznej eksploracji danych wielowymiarowych, stwarza nowe możliwości analizy zależności między zmiennymi w złożonych zadaniach geotechniki.

LITERATURA

- [1] *Dobosz M.*: Wspomagana komputerowo statystyczna analiza wyników badań. Wydawnictwo EXIT, Warszawa, 2004
- [2] *Friedman J.H.*: Multivariate Adaptive Regression Splines (with Discussion). *Annals of Statistics*, 19, 1991, pp. 1–141
- [3] *Giudici P.*: Applied Data Mining. Statistical Methods for Business and Industry. John Wiley & Sons Ltd, Chichester, 2003
- [4] *Hastie T., Tibshirani R., Friedman J.H.*: The Elements of Statistical Learning. Data mining, inference, and prediction. Springer Verlag, New York, 2001
- [5] *Kania M.M.*: Pewne aspekty odwzorowania MES pracy pała w podłożu warstwowanym. Materiały VI Krajowej Konferencji Mechaniki Gruntów i Fundamentowania, tom II. Poznań, 1984, s. 193–198
- [6] *Kania M.M.*: Współpraca fundamentu palowego z podłożem warstwowanym, Rozprawa doktorska. Politechnika Poznańska, Wydział Budownictwa Lądowego, Poznań, 1983 (maszynopis)
- [7] Statistica. Program Manual. StatSoft Inc., 2009
- [8] Salford Systems Predictive Mining Suite Manual. Salford Systems, San Diego, 2000