

Zastosowanie statystycznych metod planowania doświadczeń ekstremalnych w spektralnej analizie emisyjnej

WSTĘP

Celem pracy eksperymentatora jest znalezienie optymalnych warunków prowadzenia procesu. Do chwili obecnej najczęściej osiąga się to poprzez działanie intuicyjne, oparte na własnym doświadczeniu. Wymaga to dużej ilości doświadczeń, często nie prowadzących do celu i niepotrzebnych, gdyż mechanizm badanego zjawiska jest na ogół znany tylko częściowo. Zastosowanie matematycznej teorii doświadczeń ekstremalnych [1] - [4] umożliwi otrzymanie modelu matematycznego takiego właśnie procesu.

Tworząc model matematyczny procesu opisuje się zwykle zależności badanej cechy w postaci funkcji:

$$y = f(x_1, x_2, \dots, x_k) \quad (1)$$

zwanej funkcją odpowiedzi /funkcją celu/, w której:

y - jest parametrem procesu podlegającym optymalizacji /funkcja odpowiedzi, celu/,

x_1, x_2, x_k - zmienne niezależne /tzw. czynniki/, którym przypisuje się odpowiednio zmieniające się wartości.

Ze względu na złożoność badanego procesu ustalenie analitycznej postaci funkcji celu na podstawie wyników badań podstawowych jest niemożliwe lub nieoptyczne. W takim przypadku dąży się do tego, aby na drodze empirycznej, na podstawie wyznaczonych doświadczalnie wartości odpowiedzi y w zaplanowanych punktach doświadczalnych, uzyskać wystarczająco dobre przybliżenie tej funkcji, czyli jej empiryczną formę matematyczną.

Na podstawie twierdzenia Taylora dla funkcji wielu zmiennych jako model funkcji odpowiedzi stosuje się najczęściej wielomian n -tego stopnia; gdy ograniczy się go do przybliżenia liniowego, otrzymuje się:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i \quad (2)$$

W przypadku bardziej złożonego przybliżenia, np. za pomocą wielomianu stopnia drugiego:

$$\hat{y} = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i < j} b_{ij} x_i x_j \quad (3)$$

W praktyce rzadko zachodzi konieczność stosowania modelu stopnia wyższego niż drugi. Współczynniki b_i wyznacza się metodą analizy regresji i noszą one nazwę współczynników regresji.

Sposób wyznaczania współczynników regresji najwygodniej przedstawić w postaci macierzowej.

W zapisie macierzowym ogólne równanie regresji ma postać:

$$Y = XB \quad (4)$$

gdzie: Y - wektor obserwacji

X - macierz planowania

B - macierz poszukiwanych wartości współczynników regresji.

Równanie $Y = XB$ rozwiązuje się za pomocą lewostronnego mnożenia przez macierz transponowaną X^T macierzy X

$$X^T Y = (X^T X) B \quad (5)$$

a następnie również poprzez lewostronne mnożenie przez $(X^T X)^{-1}$

$$(X^T X)^{-1} X^T Y = (X^T X)^{-1} (X^T X) B = JB = B \quad (6)$$

Końcowe równanie umożliwia bezpośrednie obliczanie przy użyciu komputera współczynników regresji dla dowolnie wybranego typu i rzędu planu eksperymentów [1], [3].

Macierz $X^T X$ nazywa się macierzą równań normalnych, a odwrotną w stosunku do niej macierz $(X^T X)^{-1}$ macierzą błędów /dyspersji, precyzji/. Problem wyznaczenia współczynników regresji sprawdza się więc do utworzenia macierzy równań normalnych, odwrócenie tej macierzy oraz wyznaczenia wartości wektora B przez kolejne pomnażanie macierzy.

1. PLANOWANIE DOŚWIADCZENIEM

Planowanie doświadczeń ma na celu umożliwienie wyznaczenia modelu matematycznego rozpatrywanego procesu.

Na podstawie apriorycznych wiadomości o procesie obiera się pewną strategię kierowania doświadczeniem.

Przebieg doświadczenia dzieli się zazwyczaj na oddzielne etapy.

W etapie pierwszym należy tak rozłożyć punkty w przestrzeni czynnikowej, aby uzyskać pojęcie o powierzchni odpowiedzi i ustalić kierunek przemieszczania się do obszaru optymalnego. Zwykle zaczyna się od przybliżenia liniowego.

w etapie drugim dokonuje się w obszarze optymalnym większej serii doświadczeń celem wyznaczenia dokładniejszego opisu matematycznego za pomocą wielomianu wyższego stopnia niż pierwszy.

Istnieje kilka eksperymentalnych metod optymalizacji [1], [2], [4]. Do najbardziej rozpowszechnionych należy metoda sympleksów i metoda gradientu.

Optymalizację metodą sympleksów rozpoczyna się od zbudowania sympleksu wyjściowego. Sympleks ten jest wyznaczony przez $k + 1$ punktów doświadczalnych P_0, P_1, \dots, P_k , które w k -wymiarowej przestrzeni czynnikowej stanowią wierzchołki wieloboku wypukłego. Najmniejsza liczba punktów doświadczalnych, potrzebna do przybliżenia liniowego powierzchni odpowiedzi, tworzy sympleks foremny/ stosuje się najczęściej sympleksy foremne/. Dla $k = 2$ / k - liczba czynników/ sympleksem foremnym jest trójkąt równoboczny, dla $k = 3$ - czworościan itd. W zaplanowanych wierzchołkach P_i sympleksu wyjściowego wykonuje się pomiary odpowiedzi y_i . Krok optymalizacyjny polega na zastąpieniu punktu P_w , któremu odpowiada najmniej korzystna wartość odpowiedzi y_w , nowym punktem P^* , będącym lustrzanym odbiciem punktu P_w względem krawędzi wspólnej dla obu sympleksów. Punkty P_i / $i \neq w$ / razem z punktem P^* tworzą nowy sympleks. Po wyznaczeniu odpowiedzi w punkcie P^* postępuje się z nowym sympleksem analogicznie jak z wyjściowym - tzn. tworzy się sympleks następny. Proces optymalizacji kończy się wówczas, gdy zostaną spełnione odpowiednio z góry dobrane kryteria optymalizacji [1], [2], [3].

Metodę gradientu opracowali Box i Wilson [6]. Proces optymalizacji dzieli się tu na 2 etapy: etap pierwszy jest związany z osiągnięciem tzw. obszaru prawie stacjonarnej funkcji odpowiedzi; drugi etap optymalizacji obejmuje badanie obszaru stacjonarnej oraz dokładniejszą lokalizację optimum. Na każdym etapie optymalizacji przeprowadza się badanie funkcji odpowiedzi opierając się na planie doświadczalnym, zwanym doświadczeniem czynnikowym. Cechą charakterystyczną doświadczenia czynnikowego jest równoczesna zmiana wszystkich parametrów wg określonego i ściśle przestrzegane algorytmu tzw. macierzy planowania. Macierz planowania dla dwóch czynników x_1 i x_2 przedstawiono przykładowo w tabl.1.

Plan czynnikowy typu 2^2

Tablica 1

Macierz planowania				Wektor obserwacji Y
\tilde{x}_0	\tilde{x}_1	\tilde{x}_2	$\tilde{x}_1 \times \tilde{x}_2$	
+ 1	- 1	- 1	+ 1	y_1
+ 1	+ 1	- 1	- 1	y_2
+ 1	- 1	+ 1	- 1	y_3
+ 1	+ 1	+ 1	+ 1	y_4

W początkowej fazie eksperymentowania czynnikom nadaje się tylko dwie wartości ekstremalne: poziom górny umownie oznaczony +1 i poziom dolny - 1. Wszystkie możliwe kombinacje dla 2 czynników ustalonych na 2 poziomach są wy-czerpane, jeżeli dokona się czterech pomiarów. Jest to tzw. pełny^{1/} plan czynni-kowy typu $2^s / s$ - liczba czynników/.

^{1/} W przypadku dużej liczby czynników stosuje się ułamkowe replikacje doświadcze-nia czynnikowego

W pierwszej kolumnie tabl. 1 podano wartość zmiennej fikcyjnej $x_0 = 1$, w drugiej i trzeciej kolumnie - wartości zmiennych x_1 i x_2 /te dwie kolumny obejmują właściwe planowanie/, w czwartej zapisano wartość iloczynu $x_1 x_2$. Ostatnia kolumna nie odnosi się bezpośrednio do macierzy planowania. Jest to wektor wartości wyników pomiarów. Pierwszy wiersz kolumny odpowiada pierwszemu pomiarowi, podczas którego obie zmienne niezależne x_1 i x_2 znajdują się na niższych poziomach. Podczas drugiego pomiaru zmienna x_1 znajduje się na wyższym poziomie, a zmienna x_2 - na niższym poziomie itd.

Stosując takie planowanie można obliczyć współczynniki regresji równania 1 stopnia:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 \quad (7)$$

lub niepełnego równania II stopnia:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1^2 x_2^2 \quad (8)$$

oraz wyznaczyć gradient funkcji odpowiedzi określający kierunek dalszych pomiarów. Znalezione w wyniku doświadczenia czynnیکowego punkt optymalny staje się środkiem kolejnego doświadczenia czynnیکowego. Cykl ten powtarza się aż do osiągnięcia obszaru prawie stacjonarnego. Obszaru takiego nie można już opisać przybliżeniem liniowym. W tej części powierzchni odpowiedzi dominującą rolę odgrywają współczynniki regresji, charakteryzujące efekty współdziałania. Zwykle obszar stacjonarny udaje się dobrze opisać wielomianem stopnia drugiego. W tym celu należy rozporządzać systemem planowania, w którym każda zmienna przyjmuje przynajmniej trzy różne wartości. Plany tego typu są na ogół bardzo skomplikowane i wymagają dużej liczby doświadczeń. Jednym z planów wielopoziomowych jest plan Boxa-Behnkena [5], [7]. Macierz planowania, np. dla $s = 3$, wygląda następująco /tabl. 2/.

Tablica 2

Plan Boxa-Behnkena

Nr doświadczenia	Zmienne kodowane								
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	$\tilde{x}_1 \tilde{x}_2$	$\tilde{x}_1 \tilde{x}_3$	$\tilde{x}_2 \tilde{x}_3$	\tilde{x}_1^2	\tilde{x}_2^2	\tilde{x}_3^2
1	-1	-1	0	+1	0	0	+1	+1	0
2	+1	-1	0	-1	0	0	+1	+1	0
3	-1	+1	0	-1	0	0	+1	+1	0
4	+1	+1	0	+1	0	0	+1	+1	0
5	-1	0	-1	0	+1	0	+1	0	+1
6	+1	0	-1	0	-1	0	+1	0	+1
7	-1	0	+1	0	-1	0	+1	0	+1
8	+1	0	+1	0	+1	0	+1	0	+1
9	0	-1	-1	0	0	+1	0	+1	+1
10	0	-1	+1	0	0	-1	0	+1	+1
11	0	+1	-1	0	0	-1	0	+1	+1
12	0	+1	+1	0	0	+1	0	+1	+1

Nr doświadczenia	Zmienne kodowane								
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	$\tilde{x}_1\tilde{x}_2$	$\tilde{x}_1\tilde{x}_3$	$\tilde{x}_2\tilde{x}_3$	\tilde{x}_1^2	\tilde{x}_3^2	\tilde{x}_3^2
13	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0

Planowanie Boxa-Behnkena wymaga mniejszej liczby pomiarów niż całkowite doświadczenie czynnikowe typu 3^5 . Plan ten służy do wyznaczania funkcji regresji w postaci:

$$\hat{y} = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n b_{ii} x_i^2 + \sum_{i < j} b_{ij} x_i x_j \quad (9)$$

Poszczególne współczynniki regresji są obliczane wg wzorów:

$$b_0 = \frac{\sum y_0}{3} \quad (10)$$

$$b_i = \frac{1}{8} \sum_{n=1}^{15} x_{in} y_n \quad (11)$$

$$b_{ii} = \frac{1}{8} \sum_{n=1}^{15} x_{in}^2 y_n - 0,0208 \left(\sum_{n=1}^{15} x_{1n}^2 y_n + \sum_{n=1}^{15} x_{2n}^2 y_n + \sum_{n=1}^{15} x_{3n}^2 y_n \right) - \frac{\sum y_0}{6} \quad (12)$$

$$b_{ij} = \frac{1}{4} \sum_{n=1}^{15} x_{in} x_{jn} y_n \quad (i \neq j; i = 1, 2, 3) \quad (13)$$

Stosując plan Boxa-Behnkena otrzymuje się dużą koncentrację punktów w obszarze optymalnym. Umożliwia to prawidłowe wybranie najbardziej korzystnych parametrów procesu.

2. ANALIZA STATYSTYCZNA RÓWNIANIA REGRESJI

Wszystkie modele matematyczne uzyskiwane w trakcie optymalizacji poddaje się badaniom statystycznym [1], [3]. Analiza ta obejmuje między innymi: obliczenie wariancji eksperymentu, obliczenie wariancji zgodności modelu oraz przyjęcie lub odrzucenie hipotezy o adekwatności modelu.

Wariancję eksperymentu oblicza się na podstawie kilku powtórzeń; najczęściej w środku planu, tzn. dla poziomu podstawowego każdego z parametrów, na podstawie wzoru

$$s_y^2 = \frac{\sum_{n=1}^q (y_n - \bar{y})^2}{q - 1} \quad (14)$$

gdzie: y_n - kolejne wyniki pomiarów w środku planu
 \bar{y} - średnia arytmetyczna wyników
 q - liczba doświadczeń w środku planu.

Wariancję zgodności modelu oblicza się wg wzoru:

$$s_{y-\hat{y}}^2 = \frac{\sum_{n=1}^N (y_n - \hat{y})^2}{N - k} \quad (15)$$

gdzie: y_n - wielkości znalezione doświadczalnie
 \hat{y} - wielkość obliczona z równania regresji
 N - liczba doświadczeń
 k - liczba współczynników regresji

Jeżeli uzyskany model matematyczny dla zmiennej y opisuje dobrze powierzchnię odpowiedzi, to statystyka

$$F = \frac{s_{y-\hat{y}}^2}{s_y^2} \quad (16)$$

ma rozkład F-Snedecora. Jeżeli jej wartość liczbowa jest mniejsza niż podano w tablicach [8] wartość funkcji F-Snedecora dla $N-k$ i $q-1$ stopni swobody, nie ma podstaw do odrzucenia hipotezy, że przyjęty model jest dobrym przybliżeniem badanej zależności. W przeciwnym przypadku należy przyjąć nowy model i obliczenia zacząć od nowa.

W części doświadczalnej przedstawiono zastosowanie podanej w skrócie matematycznej teorii planowania doświadczeń ekstremalnych do wyboru optymalnych warunków oznaczenia zanieczyszczeń metalicznych w grafitowym koncentracie, uzyskanym po oddzieleniu pierwiastka podstawowego podczas spektrochemicznej analizy niektórych materiałów wysokiej czystości.

3. CZĘŚĆ DOŚWIADCZALNA

W analizie spektralnej materiałów wysokiej czystości głównym celem jest uzyskanie wysokiej intensywności linii widmowych oznaczonych zanieczyszczeń śladowych. Jak wykazano we wcześniejszych pracach autora [9] intensywność linii jest wynikiem wielu procesów zachodzących w płazmie. Na procesy te wywierają wpływ między innymi takie czynniki, jak skład matrycy, natężenie prądu, skład atmosfery gazowej, w której odbywa się wzbudzenie, rodzaj i ilość nośnika, kształt elektrod, szybkość przepływu gazu itp.

Przedmiotem niniejszej pracy było zbadanie wpływu na natężenie linii widmowych następujących czynników: zawartości tlenu w argonie x_1 , %/, czasu ekspozycji x_2 , s/ i szybkości przepływu gazu x_3 , l/min/. Uzyskano w ten sposób informację o warunkach przechodzenia próbki z elektrody do plazmy, w zależności od czynnika będącego w dużym nadmiarze w stosunku do masy próbki, a mianowicie od otaczającej atmosfery. Do realizacji tego zadania wykorzystano matematyczne metody planowania doświadczeń ekstremalnych. Kryterium optymalizacji był logarytm stosunku intensywności linii do intensywności tła ($y = \log I_{Me}/I_T$).

W wyniku przeprowadzonych doświadczeń utworzono model matematyczny zależności y wybranych linii Zn, Cd, Bi, Mn, Ti, Pd, Fe i Ca od rozpatrywanych czynników w szerokim zakresie zmienności badanych parametrów. Proces optymalizacji przedstawiono na przykładzie linii wapnia Ca 317,9 nm.

3.1. Technika analityczna

W pracy zastosowano spektrograf siatkowy PGS-2. Jako źródła wzbudzenia użyto łuku prądu stałego o natężeniu prądu 10 A. Badaniom poddawano wzorzec spektrograficzny wykonany na bazie proszku grafitowego, służący do spektrochemicznej analizy następujących substancji wysokiej czystości: Ga, Ag, Sb, SbCl₃, SbCl₅, PCl₃, POCl₃, AsCl₃. Ponieważ w procesie wzbogacania bznoczane w tych materiałach pierwiastki śladowe przechodzą do roztworu w postaci soli chlorkowych, stwierdzono doświadczalnie, że jest celowe stosowanie jako nośnika chlorku sodu.

3.2. Przebieg optymalizacji

Założono, że zależność y od x_1 , x_2 , x_3 można dokładnie opisać za pomocą liniowej funkcji regresji postaci:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$$

W celu ułożenia planu doświadczeń trzeba było ustalić zakresy zmienności poszczególnych czynników. Przyjęto za interesujący zakres stężenia tlenu w argonie od 5 do 25%. Powyżej 25% "konsumpcja" /utlenianie/ elektrod jest zbyt duża,

poniżej 5%, jak wynika z pracy [9], zawartość tlenu jest zbyt mała by mogła wpływać w sposób widoczny na proces wypalania się pierwiastków. Na podstawie krzywych parowania pierwiastków [9] czas ekspozycji od 20 do 60 s wydawał się wystarczający na całkowite odparowanie próbki. Większość materiału wyparuje w ciągu pierwszych 20 s palenia się łożka.

Zastosowany palnik wymagał jeszcze zbadania szybkości przepływu gazu. Zbyt mały przepływ powoduje nadmierne rozgrzanie i zapylenie elektrod; duży przepływ powoduje za szybkie przenoszenie badanego materiału poza obszar świecącej plazmy. Na podstawie wyników pracy [9] przyjęto, że szybkość przepływu gazu należy zbadać w przedziale od 1 do 3 l/min.

Tablica 3

Poziomy badanych czynników

Wyszczególnienie	Zmienne niezależne rzeczywiste		
	zawartość tlenu w argonie %	czas ekspozycji s	szybkość przepływu gazu l/min
Poziom dolny	5	20	1
Poziom górny	25	60	3
Poziom podstawowy /środek planu/	15	40	2
Krok eksperymentu	10	20	1
	Zmienne niezależne zakodowane		
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3
Poziom dolny	- 1	- 1	- 1
Poziom górny	+ 1	+ 1	+ 1
Poziom podstawowy	0	0	0

W pierwszym etapie do sprawdzenia założonego modelu matematycznego przyjęto przedstawione w tabl. 3 poziomy badanych czynników uzasadnione poprzednio. Dla uproszczenia obliczeń stosowano tzw. zmienne kodowane $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$. Przejście z wielkości naturalnych na zakodowane jest określone następującymi wzorami:

$$\tilde{x}_1 = \frac{x_1 - A_1}{\Delta_1}; \quad \tilde{x}_2 = \frac{x_2 - A_2}{\Delta_2}; \quad \tilde{x}_3 = \frac{x_3 - A_3}{\Delta_3} \quad (17)$$

w którym A_1, A_2, A_3 oznaczają wartości zmiennych naturalnych x_1, x_2, x_3 w środku planu a wielkości $\Delta_1, \Delta_2, \Delta_3$ - przedziały zmian odpowiednio $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$.

Zgodnie z planem wykonano 2^3 doświadczenia i dodatkowo dla określenia wariancji eksperymentu 4 doświadczenia w środku planu. Macierz planowania przedstawiono w tabl. 4. Wyznaczono wszystkie cztery współczynniki regresji dla następujących pierwiastków: Zn, Cd, Bi, Mn, Ti, Pd, Fe i Ca /tabl. 5/.

Plan typu 2^3

Nr doświadczenia n	Macierz planowania			
	\tilde{x}_0	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3
1	+1	+1	+1	+1
2	+1	-1	+1	+1
3	+1	+1	-1	+1
4	+1	-1	-1	+1
5	+1	+1	+1	-1
6	+1	-1	+1	-1
7	+1	+1	-1	-1
8	+1	-1	-1	-1
9	+1	0	0	0
10	+1	0	0	0
11	+1	0	0	0
12	+1	0	0	0

Model matematyczny opisujący np. zależność y od x_1, x_2, x_3 dla linii wapienia Ca 317,9 nm ma postać:

$$y = 1,27 - 0,0275 \tilde{x}_1 + 0,22 \tilde{x}_2 - 0,0025 \tilde{x}_3$$

Tablica 5

Wartości współczynników regresji

Pierwiastek	b_0	b_1	b_2	b_3
Zn	0,5425	0,0656	0,101	0,0294
Cd	0,9621	- 0,0219	0,00562	- 0,0894
Bi	1,6383	0,0844	0,0406	- 0,112
Mn	1,5171	0,146	0,162	- 0,107
Ti	1,1870	0,191	0,206	- 0,0075
Fe	1,3154	0,165	0,189	- 0,130
Pd	0,8858	0,307	0,161	- 0,0237
Ca	1,27	- 0,0275	0,22	- 0,0025

Wariancja s_y^2 charakteryzująca błąd doświadczenia /zgodnie z wzorem 14/ wynosi 0,00196. Dane liczbowe do obliczenia wariancji zgodności modelu /wg wzoru 15/ przedstawiono w tabl. 6.

Jak wynika z tabl. 6 dla $n=3 \dots 7$ istnieje dobra zgodność wyników uzyskanych drogą doświadczalną z odpowiednimi oszacowaniami z równania regresji. Na końcach przedziału, tzn. dla $n=1, n=2, i n=8$, obserwuje się natomiast duże różnice wartości obliczonych i otrzymanych w wyniku doświadczenia. Korzystając z obliczonych wariancji sprawdzono za pomocą testu F-Snedecora, czy założony model rzeczywiście dobrze opisuje zależność y od x_1, x_2, x_3 dla linii Ca 317,9 nm.

Tabela pomocnicza do obliczania wariancji zgodności modelu funkcji liniowej

Nr doświadczenia n	y_n	\hat{y}	$y_n - \hat{y}$	$(y_n - \hat{y})^2$
1	1,24	1,46	0,22	0,0484
2	1,57	1,51	0,06	0,0036
3	1,01	1,02	0,01	0,0001
4	1,02	1,07	0,05	0,0025
5	1,45	1,46	0,01	0,0001
6	1,47	1,52	0,05	0,0025
7	1,04	1,02	0,02	0,0004
8	0,90	1,08	0,18	0,0324
Σ	-	-	-	0,0900

W rozpatrywanym przykładzie /zgodnie z wzorem 16/ otrzymano:

$$F = \frac{0,0225}{0,00196} = 11,4$$

Odczytano z tablic rozkładu F-Snedecora wartość krytyczną $F_{\alpha, \phi_1, \phi_2}$ dla $\phi_1 = N-k = 4$ i $\phi_2 = n-i = 3$ stopni swobody wynosi 9,12 dla poziomu istotności $\alpha = 0,05$ /tzn. prawdopodobieństwo, że wartość zmiennej F przekroczy wartość 9,12 wynosi 5%./

Ponieważ obliczona wartość F jest większa od wartości tablicowej, nie można przyjąć hipotezy o adekwatności modelu, tzn. model liniowy nie jest dostatecznym przybliżeniem dla opisu badanej zależności, co potwierdzają dane zawarte w tabl. 6.

Ponieważ model liniowy jest nieodpowiedni, do uwzględnienia wzajemnych wpływów poszczególnych czynników przyjęto niepełny model funkcji kwadratowej o postaci:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \quad (19)$$

Obliczono współczynniki regresji przy współdziałaniach:

$$b_{12} = -0,06 \quad b_{13} = -0,0575 \quad b_{23} = -0,025$$

Model opisujący zmienną y dla linii Ca 317,9 nm ma teraz postać:

$$y = 1,27 - 0,0275 \bar{x}_1 + 0,22 \bar{x}_2 - 0,0025 \bar{x}_3 - 0,06 \bar{x}_1\bar{x}_2 - 0,0575 \bar{x}_1\bar{x}_3 - 0,025 \bar{x}_2\bar{x}_3$$

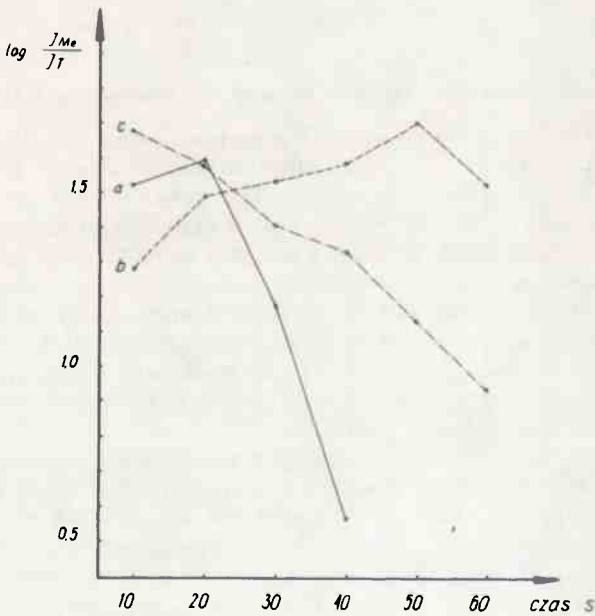
Wyniki pomiarów potrzebne do obliczenia zgodności modelu podano w tabl. 7.

Tabela pomocnicza do obliczania wariancji zgodności modelu funkcji kwadratowej

Nr doświadczenia n	y_n	\hat{y}	$y_n - \hat{y}$	$(y_n - \hat{y})^2$
1	1,24	1,32	0,08	0,0064
2	1,57	1,61	0,04	0,0016
3	1,01	1,05	0,04	0,0016
4	1,02	1,10	0,08	0,0064
5	1,45	1,49	0,04	0,0016
6	1,47	1,55	0,08	0,0064
7	1,04	1,12	0,08	0,0064
8	0,90	0,94	0,04	0,0016
Σ	-	-	-	0,0320

Wariancja zgodności modelu $s_{y-\hat{y}}$ wynosi 0,008. Obliczona zmienna F-Snedecora dla nowego modelu jest równa 4,08. Jak widać obliczona wartość F jest mniejsza od wartości tablicowej $F_{0.05.4.3}$. Wobec tego przyjęto, że model dostatecznie przybliża badaną zależność.

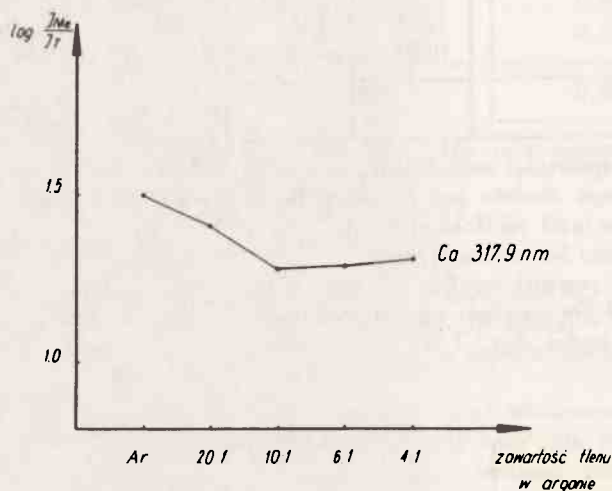
Z równania regresji wynika, że największy wpływ na wielkość y dla linii wapnia Ca 317,9 nm wywiera czas ekspozycji x_2 . Jest to zgodne z krzywą parowania tego pierwiastka /rys. 1/.



Rys. 1. Krzywe parowania wapnia: (a) w powietrzu, (b) w argonie, (c) w mieszaninie Ar + O₂ /4:1/

Wapń został zaliczony przez Schrolla do grupy pierwiastków trudno lotnych w powietrzu, w redukującej atmosferze elektrody węglowej, na równi z Mg, Ti, V, Cr, Fe, Co, Ni, Pd, Be, Al i Au.

Krzywa parowania wapnia w atmosferze argonu jest bardzo rozciągnięta w czasie. W mieszaninie argonu z tlenem parowanie jest również powolne, ale wyraźnie intensywność linii maleje w czasie. Czas ekspozycji 20 s jest tu stanowczo za krótki, bo po upływie 60 s jeszcze dużo wapnia zostaje w kraterze elektrody. Mała i ujemna wartość współczynnika regresji $x_3/b_3 = -0,0275/$ przy zmiennej x_1 , również znajduje swoje potwierdzenie w minimalnej zmianie intensywności linii tego pierwiastka w zależności od zawartości tlenu w argonie /rys. 2/. Jak widać na rys. 2 intensywność linii Ca 317,9 nm bardzo powoli maleje w miarę wzrostu zawartości tlenu. Bardzo niska wartość współczynnika $b_3 = -0,0025$ wskazuje na nieistotne znaczenie tego czynnika w rozpatrywanym zakresie zmienności.



Rys. 2. Wpływ zawartości tlenu w argonie na intensywność linii Ca 317,9 nm

Można wnioskować, że wzrost szybkości przepływu gazu, jeśli już wpływa na intensywność linii, to raczej ujemnie. Natomiast zaobserwowano spore współdziałanie czynników $x_1 \times x_2$ i $x_1 \times x_3$, co znajduje pełne potwierdzenie na rys. 1 i 2. Z równania regresji wynika, że badany proces będzie miał korzystniejszy przebieg przy czasie ekspozycji przekraczającym przyjęty poziom górny, przy równoczesnym obniżeniu szybkości przepływu gazu.

Rezultatem etapu pierwszego było ustalenie kierunku optymalizacji. W przypadku linii wapnia Ca 317,9 nm najlepsze wyniki uzyskano dla punktu 2 doświadczenia czynnikowego, a mianowicie $y=1,57$. Odpowiada to następującym wartościom czynników: $x_1 = 5\%$, $x_2 = 60$ s, $x_3 = 3$ l/min. Zgodnie z metodą gradientu jest to punkt wyjściowy do dalszych poszukiwań.

W celu uzyskania bliższych informacji o powierzchni odpowiedzi w otoczeniu tego punktu, ponownie zastosowano plan czynnikowy, ale tym razem typu 2^2 . Czynniki x_3 jako bardzo mało istotny zostawiono na poziomie -1 , a zmieniano tylko x_1 i x_2 . Na podstawie wyników etapu pierwszego przyjęto następujące zakresy zmian: $5 \leq x_1 \leq 7\%$, $90 \leq x_2 \leq 120$ s. Nie uzyskano lepszych rezultatów. Przedłużenie czasu ekspozycji znacznie zwiększyło I_0 ; czyli zmniejszeniu uległ stosunek intensywności linii do I_0 . Bardzo zbliżone wyniki wskazały na to, że osiągnięto obszar prawie stacjonarny tzw. powierzchni odpowiedzi.

Badania obszaru stacjonarnego, jak wykazano w części teoretycznej, wymaga planowania co najmniej trójpoziomowego.

Zastosowano plan doświadczeń wg Boxa-Behnkena o macierzy planowania przedstawionej w tabl. 2. Na podstawie wyników etapów pierwszego i drugiego, w etapie trzecim przyjęto następujące zakresy zmian: $5 < x_1 < 9\%$, $60 < x_2 < 80$ s, $1.0 < x_3 < 1,2$ l/min.

Najwyższą wartość y dla linii Ca 317,9 nm otrzymano przy następujących wartościach parametrów: $x_1 = 9\%$, $x_2 = 80$ s, $x_3 = 1,1$ l/min. Po obliczeniu współczynników regresji model matematyczny wyrażający zależność y od x_1 , x_2 , x_3 dla linii Ca 317,9 nm ma postać:

$$y = 1,47 + 0,011 \tilde{x}_1 + 0,031 \tilde{x}_2 - 0,0075 \tilde{x}_3 + 0,0755 \tilde{x}_1^2 + 0,030 \tilde{x}_2^2 + 0,0163 \tilde{x}_3^2 + 0,020 \tilde{x}_1 \tilde{x}_2 - 0,00175 \tilde{x}_1 \tilde{x}_3 - 0,0025 \tilde{x}_2 \tilde{x}_3$$

Wariancja eksperymentu dla $\bar{y} = 1,47$ i $q-1 = 2$ stopni swobody wynosi 0,00165; wariancja zgodności modelu jest równo 0,0057. Wartość testu F wynosi 3,45, natomiast $F_{0,05,2,2} = 19$. Wartości liczbowe F wskazują, że przyjęty model dobrze opisuje badaną funkcję. W tabelicy 8 przedstawiono wartości y uzyskane w wyniku doświadczenia i obliczone z równania regresji.

Tablica 8

Wartości y uzyskane w wyniku doświadczenia i obliczone z równania regresji \hat{y}

Nr doświadczenia n	y_i	\hat{y}_i	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
1	1,56	1,55	0,01	0,0001
2	1,57	1,53	0,04	0,0016
3	1,57	1,57	0	0
4	1,66	1,63	0,03	0,0009
5	1,57	1,54	0,03	0,0009
6	1,60	1,59	0,01	0,0001
7	1,56	1,56	0	0
8	1,52	1,54	0,02	0,0004
9	1,43	1,49	0,06	0,0036
10	1,45	1,48	0,03	0,0009
11	1,51	1,56	0,05	0,0025
12	1,52	1,54	0,02	0,0004
Σ	-	-	-	0,0114

4. OMÓWIENIE WYNIKÓW

Zastosowanie matematycznej teorii planowania doświadczeń ekstremalnych umożliwiło otrzymanie modelu matematycznego zależności natężenia linii widmowych od składu atmosfery gazowej otaczającej łuk, czasu ekspozycji i szybkości przepływu gazu.

Pracę podzielono na trzy etapy. W etapie pierwszym, w wyniku pełnego doświadczenia czynnikowego typu 2^k , ustalono kierunek optymalizacji. Wykonano pomiary współczynników regresji dla kilku pierwiastków łatwo lotnych /Zn, Cd, Bi/, średnio lotnych /Mn/ i trudno lotnych /Ti, Pd, Fe, Ca/ /klasyfikacja wg Schrolla/.

Na podstawie uzyskanych wyników stwierdzono, że wraz ze wzrostem temperatury wrzenia pierwiastka rośnie wpływ no intensywność linii zawartości tlenu w argonie i czasu ekspozycji. Dane tablicowe znajdują pełne potwierdzenie w wartościach intensywności linii uzyskanych w innym eksperymencie autora, przeprowadzonym tradycyjną metodą tzw. prób i błędów [9].

Proces optymalizacji poprowadzono w kierunku uzyskania maksymalnej wartości stosunku intensywności linii do intensywności tła. Metodę optymalizacji zilustrowano na przykładzie linii wapnia Co 317,9 nm.

5. WNIOSKI

Zastosowanie statystycznych metod planowania doświadczeń umożliwia przedstawienie w postaci modelu matematycznego zależności między natężeniem linii widmowych a czynnikami wpływającymi na to natężenie. Współczynniki regresji stanowią oszacowania indywidualnych efektów i wzajemnych oddziaływań poszczególnych parametrów na intensywność linii. Matematyczna teoria planowania doświadczeń zapewnia uzyskanie równomiernej informacji o całym obszarze n-wymiarowej przestrzeni czynnikowej przy minimalnej liczbie doświadczeń.

LITERATURA

1. Mańczak K.: Technika planowania eksperymentu. Warszawa, WNT 1976
2. Nalimow W.W., Czernowa N.A.: Statystyczne metody planowania doświadczeń ekstremalnych. Warszawa, WNT 1967
3. Zieliński R.: Wybrane zagadnienia optymalizacji statystycznej. Warszawa, PWN 1974
4. Parczewski A., Rokosz A.: Przem. Chem. 53, 200, 1974
5. Box G.E.P., Behnken D.W.: Technometrics 2, 455, 1960
6. Box G.E.P., Wilson K.B.: J. Roy. Stat. Soc. B 13, 1, 1951
7. Arpadjan S., Doerffel K., Holland-Letz, Much H., Pannach M.: Z. Anal. Chem. 270, 257, 1974
8. Zieliński R.: Tablice statystyczne. Warszawa, 1972
9. Sakołowska W.: Sprawozdanie ONPMP nr 798, 1976