

Wojciech SKRZECZANOWSKI  
Andrzej ANTONIK

## ANALIZA WIDM SPEKTROSKOPII EMISYJNEJ ZA POMOCĄ WIELOWYMIAROWYCH METOD STATYSTYCZNYCH

**STRESZCZENIE** W pracy przedstawiono przykłady zastosowania wielowymiarowych metod statystycznych w badaniach widm spektroskopii emisyjnej. Krótko scharakteryzowano najistotniejsze właściwości dwóch metod, takich jak Analiza Czynnikiowa (FA – Factorial Analysis) i Analiza Głównych Składowych (ang. PCA – Principal Components Analysis). Zastosowano je do analizy widm emisyjnych uzyskanych ze źródeł ciągłych i impulsowych. Stwierdzono, że obróbka widm za pomocą metod FA/PCA pozwala wyznaczyć stopień podobieństwa widm, co z kolei umożliwia określenie różnic w charakterze źródeł, których emisja jest rejestrowana. W pracy zestawiono wyniki analizy FA/PCA widm LIBS (spektroskopii emisyjnej ze wzbudzeniem laserowym) w zastosowaniu badań stratygraficznych, identyfikacji różnych gatunków papieru, analizy powierzchniowej korozji wyrobów ze stopów stali o różnej strukturze powierzchni, wyładowań w gazie i żarówek. Zaprezentowano próby zastosowania metody FA do konstrukcji krzywych kalibracyjnych, umożliwiających badania ilościowe stopów metali, bez konieczności korzystania z danych atomowych właściwych składnikom stopów.

**Słowa kluczowe:** analiza czynnikowa, główne składowe, widmo emisyjne

---

**dr inż. Wojciech SKRZECZANOWSKI**  
e-mail: wskrzeczanowski@wat.edu.pl

**mgr inż. Andrzej ANTONIK**  
e-mail: aantoniak@wat.edu.pl

Instytut Optoelektroniki, Wojskowa Akademia Techniczna

PRACE INSTYTUTU ELEKTROTECHNIKI, zeszyt 255, 2012

## 1. WSTĘP

---

Analiza głównych składowych (z ang. PCA – *Principal Components Analysis*) i analiza czynnikowa (FA – *Factorial Analysis*) należą do grupy statystycznych metod czynnikowych [1, 3, 4] i są bardzo szeroko stosowane w wielowymiarowej analizie porównawczej. Pozwalają one na porównywanie obiektów lub zbiorów składających się z wielu zmiennych i znajdowanie w stosunkowo szybki sposób najważniejszych cech i właściwości tych zbiorów. Metody te sprowadzają się do przedstawienia za pomocą ortogonalnego przekształcenia wejściowego zbioru obiektów dowolnej natury (macierz wejściowa) jako kombinacji liniowej nowych, nieobserwowalnych charakterystyk, zwanych głównymi składowymi lub czynnikami. Wielowymiarowa analiza porównawcza prowadzi do:

- redukcji dużej liczby zmiennych, charakteryzujących obiekty do kilku podstawowych, które mogą być poddane szczegółowej analizie;
- grupowania obiektów charakteryzowanych przez bardzo zbliżone właściwości;
- wyodrębnienia najbardziej typowych zjawisk, procesów, co powoduje zmniejszenie nakładu czasu oraz kosztów przeprowadzanych badań;
- wyjaśnienia struktury powiązań pomiędzy charakterystykami obiektów.

Nieprawidłowy dobór zmiennych wejściowych może prowadzić do błędnych wyników w analizie porównawczej. Dlatego też zbiór zmiennych opisujących obiekty w wielowymiarowej analizie statystycznej powinien być ograniczony za pomocą kryteriów pozastatystycznych (merytorycznych i formalnych), i/lub też statystycznych. Wstępna selekcja zmiennych oparta na merytorycznych przesłankach prowadzi do wyodrębnienia potencjalnych zmiennych statystycznych, których ilość redukowana jest przy wykorzystaniu metod statystycznych. Prowadzi to do wyodrębnienia ostatecznego zespołu zmiennych, będących podstawą analizy porównawczej.

Przyjmuje się, że pierwsze próby opracowania metod wieloczynnikowych, a konkretnie analizy głównych składowych, pojawiły się w kryminalistyce w związku z potrzebą identyfikacji przestępców za pomocą zbioru cech antropometrycznych. W [1] podaje się, że autorem metody identyfikacji antropometrycznej był w 2. połowie XIX wieku Bertillon, który w tym celu posługiwał się dwunastoma różnymi pomiarami ciała. Jednak szybko okazało się, że niektóre z nich są bardzo mocno skorelowane ze sobą, jak np. długość przedramienia ( $F$ ), nogi ( $L$ ) i wzrost ( $S$ ). Edgeworth zaproponował zastąpienie

takich skorelowanych zmiennych, opisujących te parametry ciała, kombinacją nieskorelowanych zmiennych hipotetycznych  $f_1$ ,  $f_2$  i  $f_3$  w następujący sposób [1]:

$$f_1 = 0,16S + 0,51F + 0,39L$$

$$f_2 = -0,17S + 0,69F - 0,09L$$

$$f_3 = -0,15S - 0,25F + 0,52L$$

Każda z tych nowych zmiennych jest funkcją pomiarów, jednak nie mają one żadnego odniesienia do rzeczywistości (np.  $f_1$  nie jest fizyczną wielkością).

A. Balicki w [1] zaznacza, że teoretyczne podstawy analizy głównych składowych stworzył w 1902 roku Pearson. W początkowym okresie rozwoju metod wieloczynnikowych uważano, że zmienne wejściowe powinny mieć rozkład normalny, potem rozszerzono analizę na zmienne losowe (Hotelling, 1933, w [1]), zaś ostatnio, w epoce gwałtownego rozwoju komputerów i ich mocy obliczeniowej, obowiązuje zupełnie inny punkt widzenia. Aktualnie nie stawia się żadnych wymagań na rodzaje i rozkłady zmiennych. Metody czynnikowe traktuje się jako technikę opisu zbiorów danych, polegającą na optymalizacji pewnych kryteriów algebraicznych i geometrycznych.

Wielowymiarowe metody statystyczne znalazły stopniowo zastosowanie w naukach społecznych (socjologia, psychologia), przyrodniczych (biologia, medycyna), ekonomicznych, aż wreszcie w naukach fizycznych i technicznych, takich jak na przykład analiza obrazu, chemia, fizyka. Są one potężnym i skutecznym narzędziem analizy złożonych, wielowymiarowych obiektów, zjawisk i procesów we wszystkich dziedzinach nauki i działalności człowieka.

Do analizy prezentowanych w pracy wyników zastosowano pakiet oprogramowania STATISTICA.PL v.10.

## 2. ISTOTA METOD WIELOCZYNNIKOWYCH

---

### 2.1. Analiza głównych składowych (PCA)

---

Zakładając [1], że mamy  $p$  zmiennych,  $X_j$  ( $j = (1, 2, \dots, p)$ ), zaś obserwacje przeprowadzone na  $n$  jednorodnych obiektach ( $i = (1, 2, \dots, n)$ ) zgrupowano w postaci macierzy  $X$  pełnego rzędu, możemy zdefiniować dla tych zmiennych macierz korelacji  $R$  i macierz kowariancji  $S$ . Dodatkowo zakłada się, że macierze kowariancji i korelacji mają pewną liczbę różnych największych wartości własnych.

Idea analizy głównych składowych [1] polega na ortogonalnej transformacji układu badanych zmiennych wejściowych  $X_j$  w zbiór nowych, nieobserwowalnych zmiennych  $Y_i$ , które są liniowymi kombinacjami tych obserwowanych zmiennych. Można to zapisać w postaci macierzowej, jako

$$y = W'x$$

gdzie  $W = \{w_{ji}\}$  jest  $(p \times m)$  macierzą współczynników,  $x = [X_1, X_2, \dots, X_p]$  jest kolumnowym wektorem badanych zmiennych (wejściowych), natomiast  $y = [Y_1, Y_2, \dots, Y_m]$  jest kolumnowym wektorem zmiennych transformowanych, tzw. głównych składowych. Są one nieskorelowane między sobą (ortogonalne) i unormowane (suma kwadratów współczynników danej kombinacji  $w_{ji}$  jest równa 1), zaś suma wariancji wszystkich składowych  $Y_i$  jest równa ogólnej wariancji zmiennych  $X_j$ .

Celem analizy głównych składowych jest przede wszystkim analiza struktury zależności, z zapewnieniem prostoty opisu tej struktury. Zmienność  $p$  – wymiarową, której wyrazem są wzajemne zależności ujawnione w macierzy kowariancji lub macierzy korelacji, przedstawiana jest jako wynik pewnej liczby ortogonalnych składowych tak, aby mała liczba składowych odpowiadała możliwie największej części zmienności.

Miarą zasobów zmienności badanego zbioru (a więc i informacji zawartej w zbiorze wejściowym) jest suma wariancji zmiennych wejściowych. Idea metody PCA sprowadza się do redukcji wielu zmiennych wejściowych, które mogą być ze sobą skorelowane; nie zawsze w jasny sposób do znacznie mniejszej liczby nowych zmiennych – głównych składowych. Redukcja zmiennych odbywa się kosztem, niestety, zmniejszenia zasobu informacji wnoszonych przez dane wejściowe, co oznacza że pierwsze, najważniejsze nowe zmienne (główne składowe), nie opisują całej zmienności danych wejściowych, nie mniej jednak mogą dobrze opisywać zasoby zbioru wejściowego. Zawsze jednak suma wariancji wszystkich nowych zmiennych po przekształceniu PCA równa się sumie wariancji zbioru wejściowego, co oznacza odtworzenie wszystkich zasobów zmienności zbioru wejściowego.

Procedura transformacji prowadząca do wyznaczenia głównych składowych zbudowana jest w ten sposób, żeby pierwsza główna składowa wyjaśniała największy zakres zmienności danych wejściowych, następna nieco mniejszy zakres, itd. Przyjmuje się, że aby poprawnie opisać zbiór danych wejściowych, dwie pierwsze główne składowe powinny opisywać powyżej 70% zmienności (wariancji) zbioru wejściowego [1, 3, 4]. W analizie PCA nowe zmienne, to jest główne składowe, nie zawsze muszą posiadać interpretację fizyczną – tworzą one niejako wirtualną przestrzeń, w której nasze zmienne

i dane wejściowe dają się prościej opisać i w której można łatwiej wykryć ich strukturę oraz wzajemne współzależności (jeśli istnieją).

## 2.2. Analiza czynnikowa (FA)

Analiza czynnikowa jest metodą badania struktury wewnętrznych zależności obserwacji wielowymiarowych [3]. W metodzie tej każda zmienna obserwowalna (wejściowa) przedstawiana jest jako kombinacja liniowa pewnej liczby nieobserwowalnych zmiennych, zwanych czynnikami, wspólnych dla całego zbioru wejściowego, oraz jednego nieobserwowalnego czynnika swoistego dla tej zmiennej. Jednocześnie czynniki wspólne i swoiste nie są ze sobą skorelowane (są ortogonalne). Dzięki takim założeniom, wariancję każdej zmiennej wejściowej (będącej, jak wspomniano w podrozdziale 2.1, miarą jej zasobu informacyjnego o badanym procesie lub zjawisku), można podzielić na wariancję wyjaśnianą przez czynniki wspólne i wariancję wyjaśnianą przez czynnik swoisty (właściwy tylko dla tej zmiennej). Celem analizy czynnikowej jest znalezienie takiego zbioru czynników wspólnych oraz określenie ich związków ze zmiennymi obserwowalnymi, który pozwala na wyjaśnienie struktury powiązań między zmiennymi obserwowalnymi. Prekursorem analizy czynnikowej jest Spearman (1904, [1, 3]).

Podobnie jak w analizie głównych składowych, punktem wyjścia w analizie czynnikowej jest budowa macierzy danych wejściowych [3]. Zakładając, że wszystkie zmienne wejściowe mają rozkład normalny, każdą z nich, po uprzednim wystandaryzowaniu, polegającym na redukcji np. wpływu skali lub jednostek pomiarowych, przedstawia się jako kombinację liniową nie tylko zmiennych nieobserwowalnych zwanych czynnikami wspólnymi (głównymi), przenoszącymi informacje wspólne dla zmiennych wejściowych, ale także pojedynczej zmiennej zwanej czynnikiem swoistym (specyficznym), przenoszącej informacje zawarte wyłącznie w tej zmiennej wejściowej i niepowielane przez inne zmienne wejściowe.

Można to przedstawić następująco:

$$\mathbf{Z} = \mathbf{WF} + \mathbf{BU} \quad (1)$$

gdzie  $\mathbf{W} = [w_{jl}]$  jest macierzą ładunków czynnikowych czynników wspólnych ( $n \times s$ ), przy czym  $w_{jl}$  jest ładunkiem czynnikowym znajdującym się przy  $j$ -tej zmiennej i  $l$ -tym ładunku wspólnym;  $\mathbf{F} = [f_{jl}]$  jest macierzą czynników wspólnych ( $s \times n$ ), przy czym  $f_{jl}$  jest wartością  $l$ -tego czynnika wspólnego w  $i$ -tym obiekcie,  $\mathbf{B} = [\text{diag}(b_j)]$  jest macierzą diagonalną ładunków czynnikowych czynników

swoistych ( $n \times m$ ), przy czym  $b_j$  jest ładunkiem czynnikowym  $j$ -tego czynnika swoistego, a  $U = [u_{ji}]$  – macierzą czynników swoistych ( $n \times m$ ), przy czym  $u_{ij}$  jest wartością  $j$ -tego czynnika swoistego w  $i$ -tym obiekcie.

Wariancję zmiennej wejściowej można przedstawić w postaci sumy wariancji wyjaśnianej przez czynniki wspólne i wariancji wnoszonej przez czynnik swoisty następująco:

$$S^2(z_j) = h_j^2 + d_j^2 = \sum_{l=1}^s w_{jl}^2 + b_j^2 = 1, \quad j = 1, 2, \dots, m \quad (2)$$

gdzie:  $h_j^2$  – zasoby zmienności wspólnej  $j$ -tej zmiennej;  $d_j^2$  – zasoby zmienności swoistej  $j$ -tej zmiennej.

W analizie czynnikowej dąży się do eliminacji wpływu czynników swoistych na rzecz czynników wspólnych, co jest równoważne z minimalizacją wpływu wszystkich innych zmiennych na zmienne wejściowe poza czynnikami wspólnymi. Odbywa się to przez zastąpienie macierzy korelacji  $R$  współczynników korelacji znajdujących się na głównej przekątnej (współczynników korelacji o wartościach równych 1) zasobami zmienności wspólnej (to jest wartościami zwykle mniejszymi od 1). Uzyskujemy wówczas zredukowaną macierz korelacji o postaci

$$\tilde{R} = \begin{cases} \tilde{r}_{jj'} = r_{jj'}, & \text{dla } j \neq j' \\ \tilde{r}_{jj'} = h^2_{j'}, & \text{dla } j = j' \end{cases} \quad j, j' = 1, 2, \dots, m \quad (3)$$

Ostatecznie układ równań (1) sprowadza się do postaci

$$Z = WF \quad (4)$$

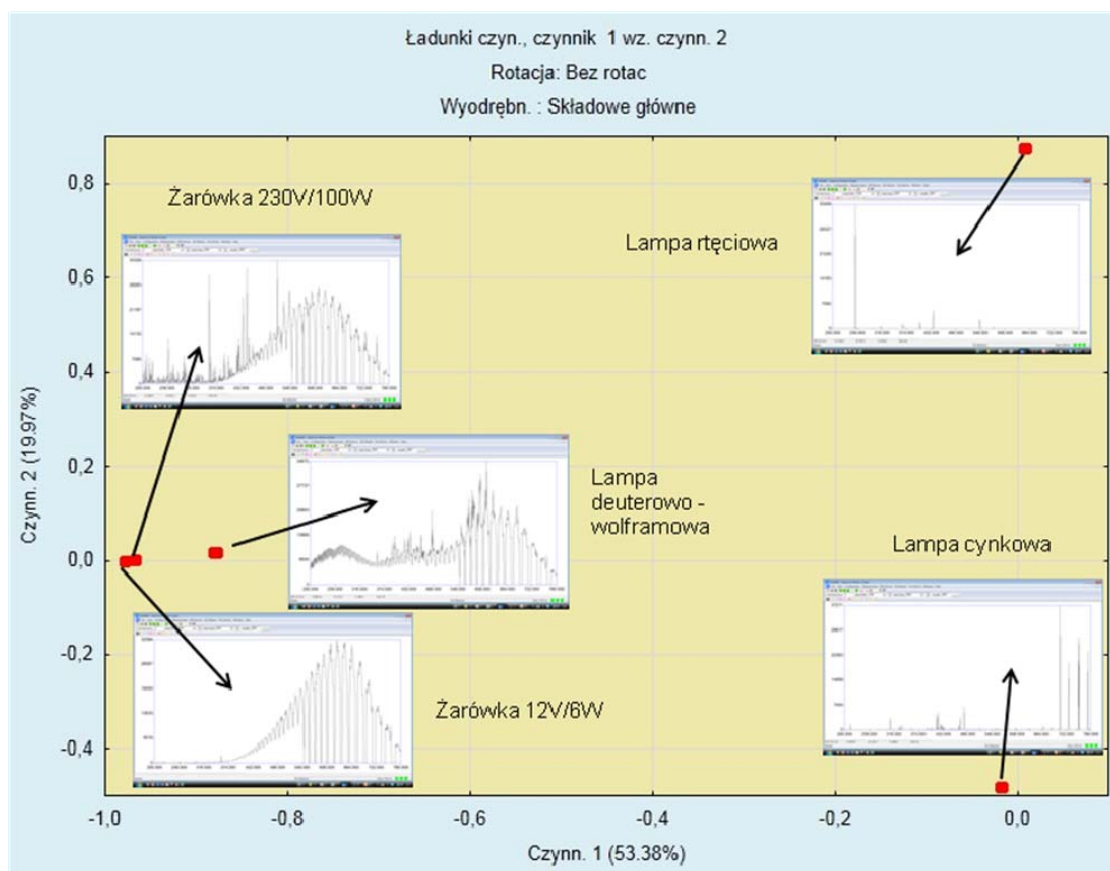
Ta zredukowana macierz korelacji (3) stanowi podstawę metod wyznaczenia ładunków czynnikowych, z których najczęściej stosowane, to metoda osi głównych (tak jak w analizie głównych składowych – jedyna różnica to zastąpienie pełnej macierzy korelacji macierzą korelacji zredukowaną), metoda centroidalna, metoda największej wiarygodności i metoda najmniejszych reszt. Bardziej szczegółowy opis procedur wyznaczania ładunków czynnikowych dla głównych składowych i czynników wspólnych można znaleźć w [1, 3, 4].

### 3. WYNIKI POMIARÓW

Zaprezentowane w tym rozdziale wyniki pomiarów dotyczą statystycznej analizy widm spektroskopii emisyjnej metodami opisanymi skrótowo w rozdziale 2. Widma te uzyskano ze źródeł stacjonarnych (żarówki, lampy spektralne) oraz impulsowych, wywoływanych oddziaływaniem impulsowego lasera na powierzchnię badanej próbki (widma LIBS). Opis metody LIBS i typowy układ eksperymentalny stosowany wielokrotnie przez autorów i przedstawiany na poprzednich konferencjach POOMT można znaleźć w [2, 5, 6].

#### 3.1. Badania źródeł stacjonarnych

Na rysunku 1 przedstawiono wyniki analizy czynnikowej dla kilku źródeł stacjonarnych (dwie żarówki i trzy lampy spektralne).



Rys. 1. Położenie zmiennych wejściowych (czerwone punkty) i odpowiadające im widma różnych źródeł

Dwa pierwsze czynniki wspólne obejmują łącznie ponad 73% całkowitej wariancji zmiennych wejściowych (widm poszczególnych źródeł), więc reprezentację zbioru wejściowego można uznać za poprawną. Wzajemne położenie punktów określa zarazem stopień podobieństwa widm. Jak widać, po lewej stronie wykresu zgromadzone zostały widma o charakterze ciągłym, zaś po prawej – widma liniowe. Można więc zinterpretować czynnik pierwszy jako określający stopień „liniowości” widma (im większa wartość czynnika, tym widmo jest bardziej liniowe).

Z kolei ze wzrostem czynnika drugiego rośnie „krótkofalowość” widma (widmo lampy rtęciowej z silną linią 253 nm w prawym górnym rogu wykresu, a widmo lampy cynkowej z przewagą linii powyżej 700 nm w prawym dolnym). Wszystkie widma zarejestrowano za pomocą spektrometru typu echelle, a ponieważ nie zastosowano korekcji spektralnej, to w widmach ciągłych (po lewej stronie wykresu) widoczna jest cykliczna struktura, związana wyłącznie ze zmienną czułością spektralną w poszczególnych rzędach dyfrakcji, typową dla siatki typu echelle.

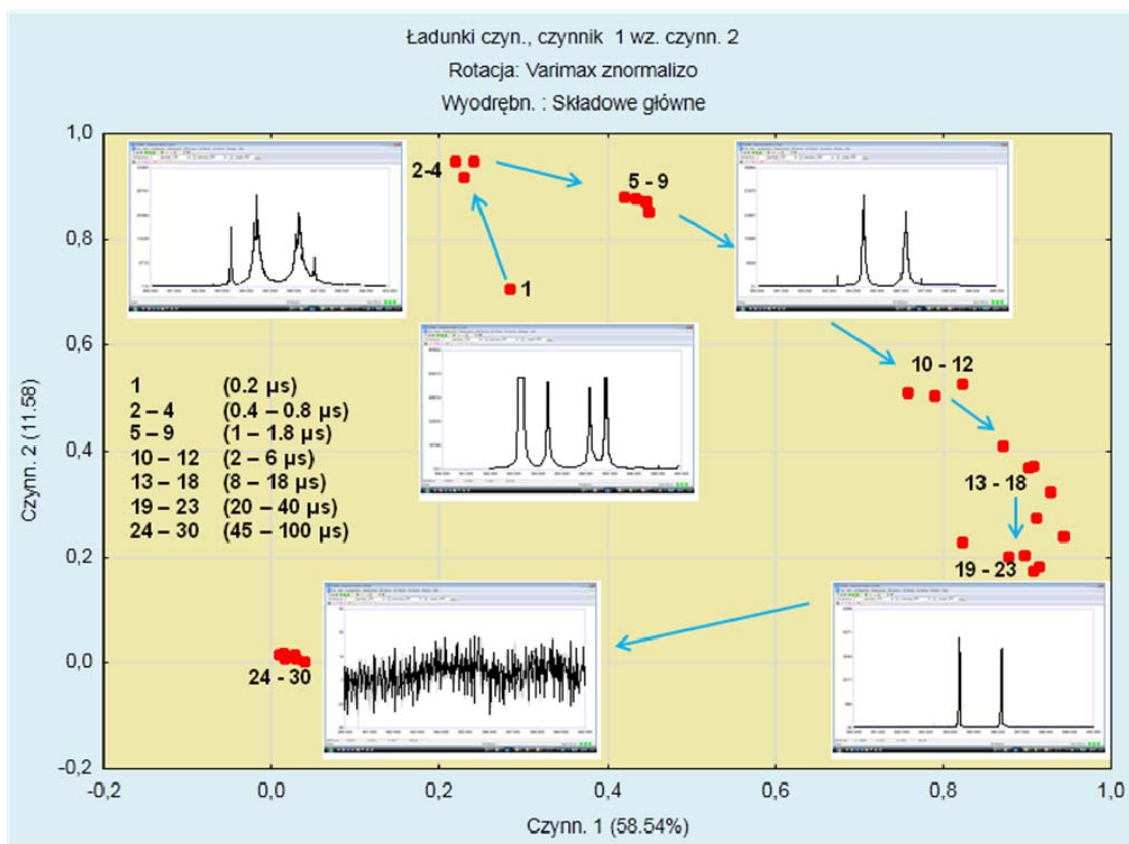
### 3.2. Badania źródeł impulsowych (LIBS)

Na rysunku 2 przedstawiono w reprezentacji wieloczynnikowej 30 widm LIBS z zakresu 200-780 nm, rejestrowanych na płycie duralowej ze zmiennym opóźnieniem w stosunku do impulsu laserowego wytwarzającego plazmę.

Do opóźnienia 2  $\mu$ s widma rejestrowano co 200 ns, między 2  $\mu$ s a 20  $\mu$ s co 2  $\mu$ s, między 20  $\mu$ s a 50  $\mu$ s co 5  $\mu$ s, a między 50  $\mu$ s a 100  $\mu$ s co 10  $\mu$ s. Dla charakterystycznych grup widm dołączono typowe widma plazmy z obszaru 390-400 nm. Dwa pierwsze czynniki wspólne reprezentują 70% zmienności wszystkich widm, to jest 30 zmiennych wejściowych. Zmienne te zgromadzone są w kilku grupach, co oznacza, że w pewnych przedziałach czasu widma są do siebie podobne, a tym samym parametry plazmy muszą być do siebie zbliżone.

Wstawki w postaci widm o szerokości 10 nm ilustrują charakter zmian zachodzących w plazmie wraz z upływem czasu. Dla małych czasów opóźnień względem impulsu laserowego rejestrowane są widma gorącej plazmy z bardzo poszerzonymi liniami widmowymi (widmo na wstawce nr 1). Z upływem czasu poszerzenie linii stopniowo spada, maleje także ich natężenie. Po upływie ponad 40  $\mu$ s od impulsu laserowego wytwarzającego plazmę promieniowanie plazmy zanika.





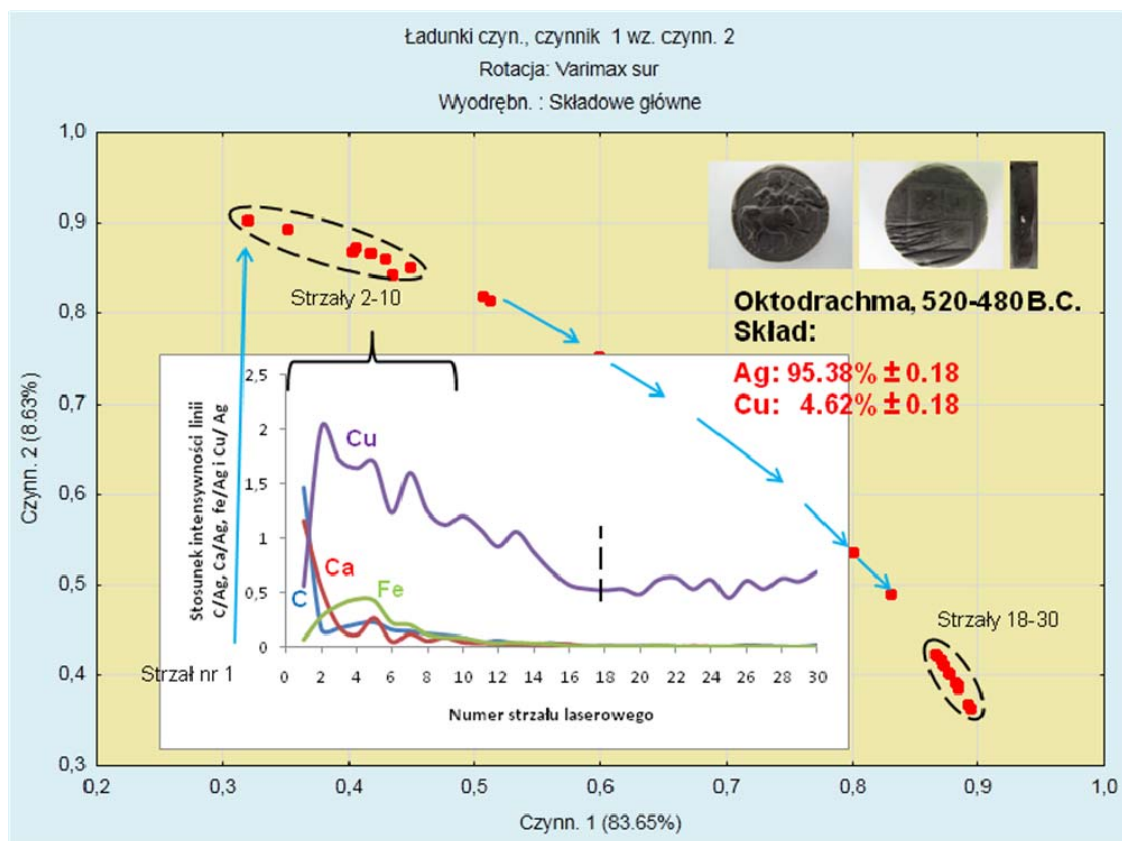
Rys. 2. Analiza czynnikowa widm LIBS wytwarzanych na płytce Al-Mg, rejestrowanych z różnym opóźnieniem względem laserowego impulsu wytwarzającego plazmę

### 3.3. Stratygrafia LIBS i analiza czynnikowa starożytnej monety

Analizę czynnikową można również skutecznie zastosować w badaniach stratygraficznych, czego dowodzą wyniki badań LIBS starogreckiej oktodrachmy sprzed 2500 lat, zaprezentowane na rysunku 3.

Na rysunku 3 przedstawiono w układzie dwóch głównych czynników wspólnych, które łącznie opisują ponad 92% ogólnych zasobów informacyjnych (wariancji) wprowadzonych do zbioru wejściowego przez 30 widm, położenia widm uzyskanych z różnych głębokości od powierzchni monety. Rozkład stratygraficzny składu pierwiastkowego monety uzyskano „strzelając” laserem w to samo miejsce na obwodzie (grzbiet) oktodrachmy. Oczywiście, wraz z głębokością, widma zmieniały się, co skutkowało różnym położeniem w układzie współrzędnych czynnik 1 – czynnik 2. Widma te odzwierciedlają rozkład stratygraficzny wybranych pierwiastków pokazany we wstawce na rysunku. Badanie to było przeprowadzone w ramach próby określenia autentyczności

monety, którą potwierdzono, wyznaczając za pomocą krzywych kalibracyjnych skład chemiczny głównego materiału, z którego wykonano monetę (wysokoprocentowy stop srebra i miedzi).



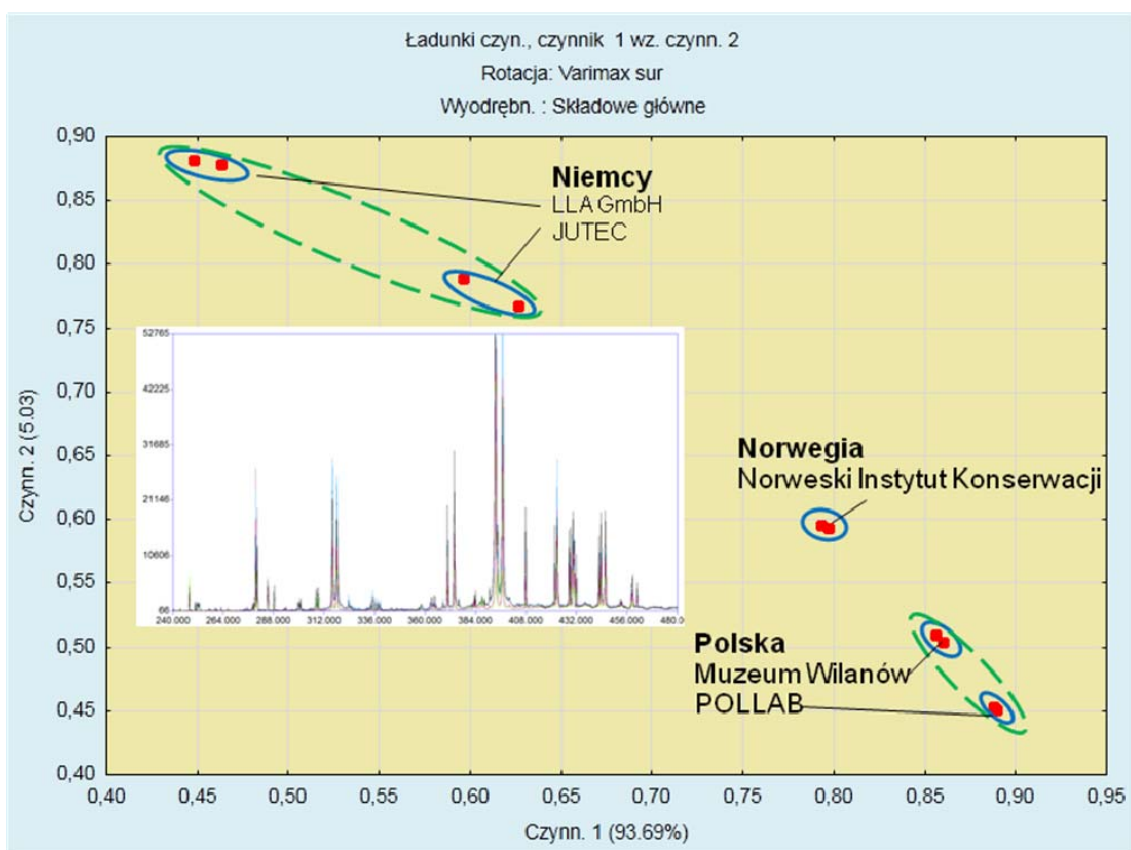
Rys. 3. Wyniki badań LIBS antycznej monety greckiej i analiza czynnikowa widm LIBS

### 3.4. Badanie papieru kredowego

Kolejnym przykładem przydatności zastosowania analizy czynnikowej do określania podobieństwa i zróżnicowania różnych materiałów są wyniki badań identyczności papieru kredowego, przedstawione na rysunku 4.

Badane próbki papieru to przykłady wydawnictw akcydensowych, pochodzące z trzech krajów. Analiza czynnikowa obejmuje przy dwóch czynnikach głównych prawie 99 procent zmienności, czyli praktycznie całość zasobów informacyjnych wnoszonych przez 10 zmiennych wejściowych (po dwa widma dla każdej z pięciu próbek). Widać wyraźnie, że zmienne wejściowe reprezentowane na wykresie przez widma leżą w różnych, daleko od siebie położonych miejscach (są to widma uzyskane w trzecim i czwartym strzale lasera w daną

próbkę papieru. Wstawka obrazująca pięć nałożonych na siebie widm z trzeciego strzału w każdą próbkę ilustruje, jak bardzo podobne są te widma i jak pracochołonna musiałaby być tradycyjna analiza widm, mająca na celu wykrycie zróżnicowania. Widoczna jest tu przewaga efektywności analizy czynnikowej w wyznaczeniu podobieństwa próbek. Ponadto, oprócz ewidentnych różnic w składzie papieru pochodzącego z różnych krajów, metoda czynnikowa pozwoliła dodatkowo wychwycić różnice w próbkach papieru z danego kraju.

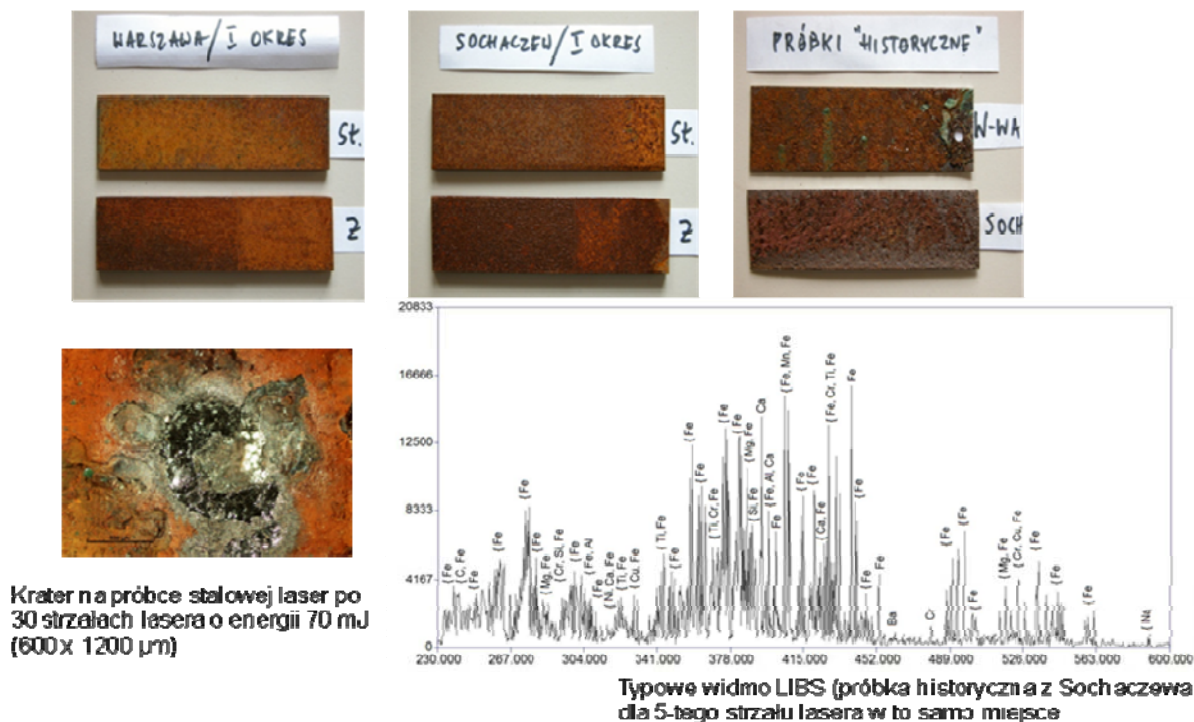


Rys. 4. Analiza czynnikowa papieru kredowego

### 3.5. Analiza czynnikowa procesów korozji stopów żelaza

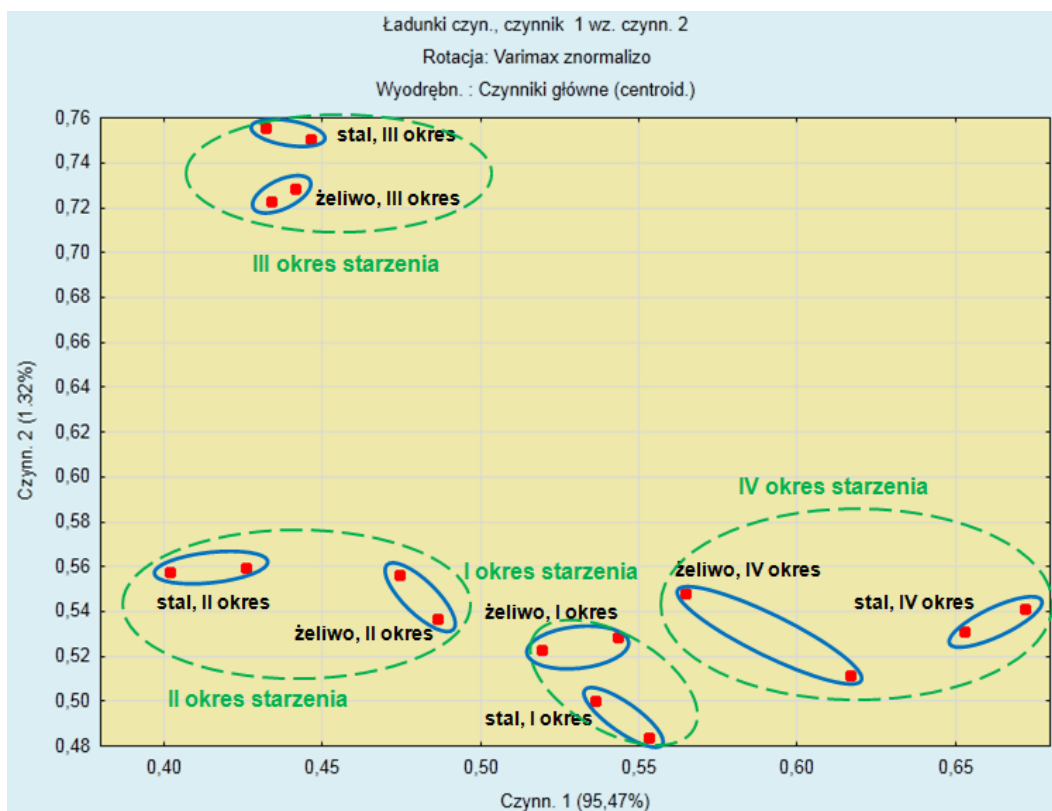
Badania korozji wykonano w 2011 roku w ramach polsko-norweskiego projektu „Ocalić od zardzewienia”. Zastosowano typową aparaturę do badań metodą LIBS, opisaną na poprzedniej konferencji POOMT, [6]. Badania dotyczyły procesów korozji próbek stali i żeliwa o różnej strukturze powierzchni (szczotkowana i szlifowana), eksponowanych na działanie środowiska w dwóch

lokalizacjach – w Warszawie i Sochaczewie. Próbkę badano w cyklu kwartalnym, oceniając stopień korozji po trzech, sześciu, dziewięciu i jedenastu miesiącach ekspozycji.

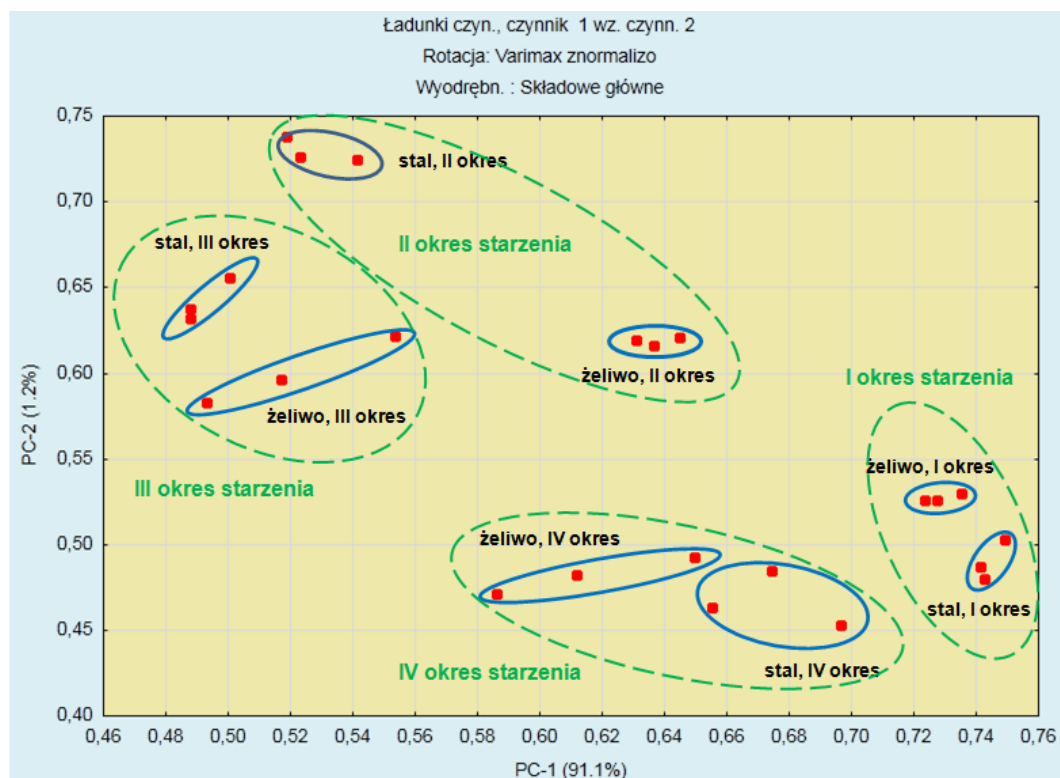


Rys. 5. Próbkę do badań korozji stopów żelaza

Typowe próbki stali i żeliwa po trzech miesiącach ekspozycji Warszawie i Sochaczewie, oraz próbki historyczne, typowe widmo LIBS i skutek badania metodą LIBS po trzydziestu strzałach lasera pokazano na rysunku 5. Na rysunkach 6-9 przedstawiono wyniki statystycznej analizy wieloczynnikowej procesów korozji zachodzących na powierzchni próbek ze stali i żeliwa w Warszawie i Sochaczewie w ciągu kolejnych jedenastu miesięcy (od listopada 2010 r. do września 2011 r.). Zaznaczane na wykresach okresy starzenia oznaczają kwartalne ekspozycje na działanie środowiska, np. I i III okres starzenia oznaczają, że próbka w pierwszym przypadku korodowała jeden kwartał (trzy miesiące), zaś w drugim – próbka była eksponowana przez dziewięć miesięcy (korozja przez III kwartały ekspozycji). We wszystkich przypadkach badania widm próbek z żelaza (patrz rys. 6-9) pierwsze dwa czynniki główne obejmują od 92,30% do 96,80% ogólnej wariancji zmiennych wejściowych, więc reprezentacja od 8 do 24 widm przez te dwa czynniki główne jest bardzo dobra.



Rys. 6. Analiza czynnikowa procesów korozji na powierzchni szlifowanej dla próbek stalowych i z żeliwa wystawionych na działanie środowiska w Warszawie

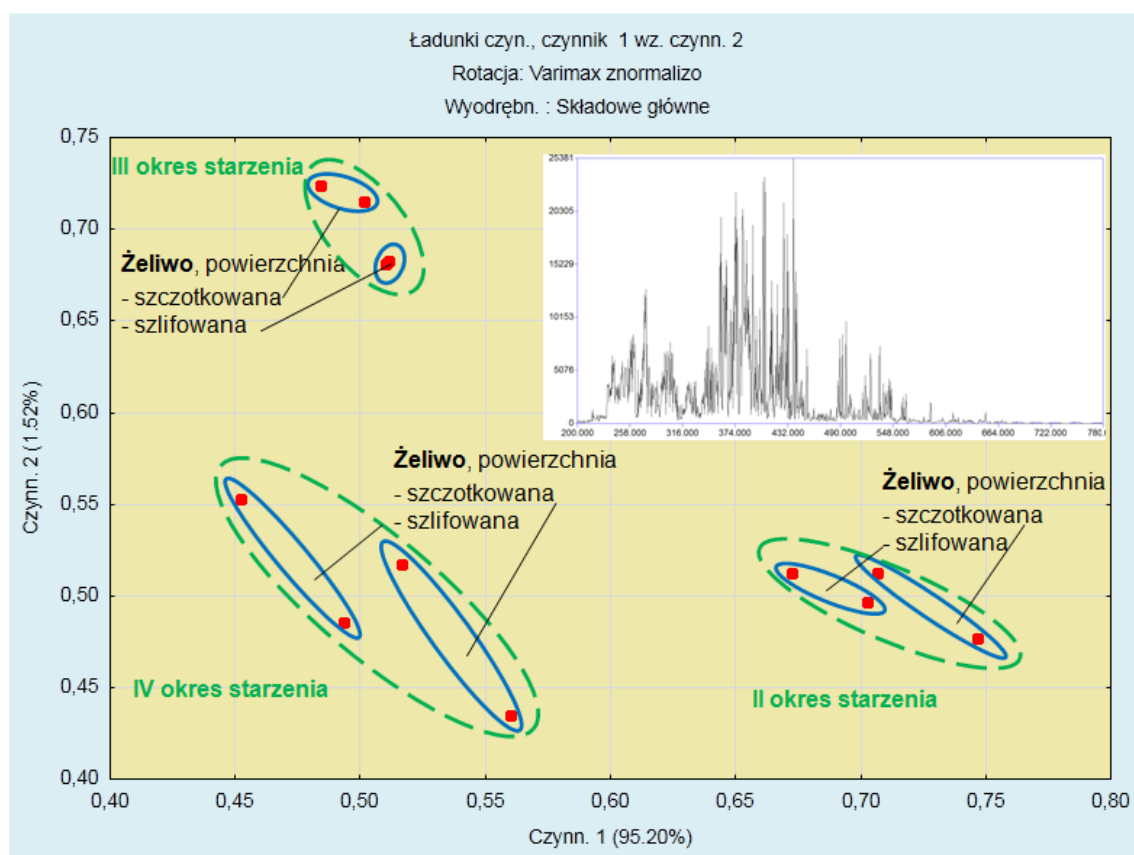


Rys. 7. Analiza czynnikowa procesów korozji na powierzchni szlifowanej dla próbek stalowych i z żeliwa wystawionych na działanie środowiska w Sochaczewie

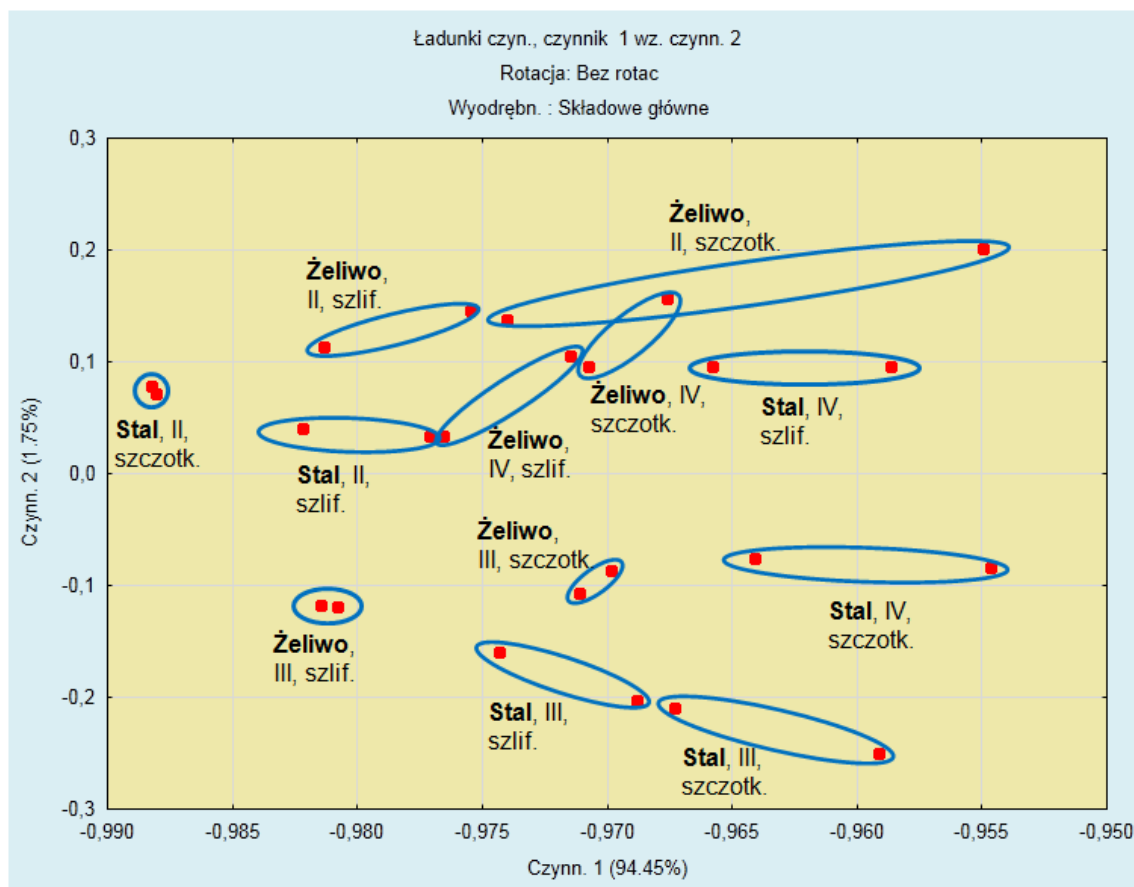


I tak, z rysunków 6 i 7 wynika, że można dobrze odróżnić widma zarejestrowane na próbkach nie tylko ze stali, jak i żeliwa, ale także widma rejestrowane po poszczególnych okresach ekspozycji (w kolejnych okresach starzenia próbek).

Badane były powierzchnie o różnej strukturze, przy czym obszary szlifowane i szczotkowane znajdowały się na każdej próbce. Z rysunków 8 i 9 wynika, że analiza czynnikowa pozwala wyodrębnić, poprzez wykazanie różnic, obszary szczotkowane i szlifowane oraz wykazuje, że procesy korozji przebiegają na nich w różny sposób. Należy podkreślić, że za pomocą statystycznych metod obróbki widm można dowieść, że korozja przebiega w różny sposób na obszarach stykających się ze sobą lub odległych o kilka milimetrów.



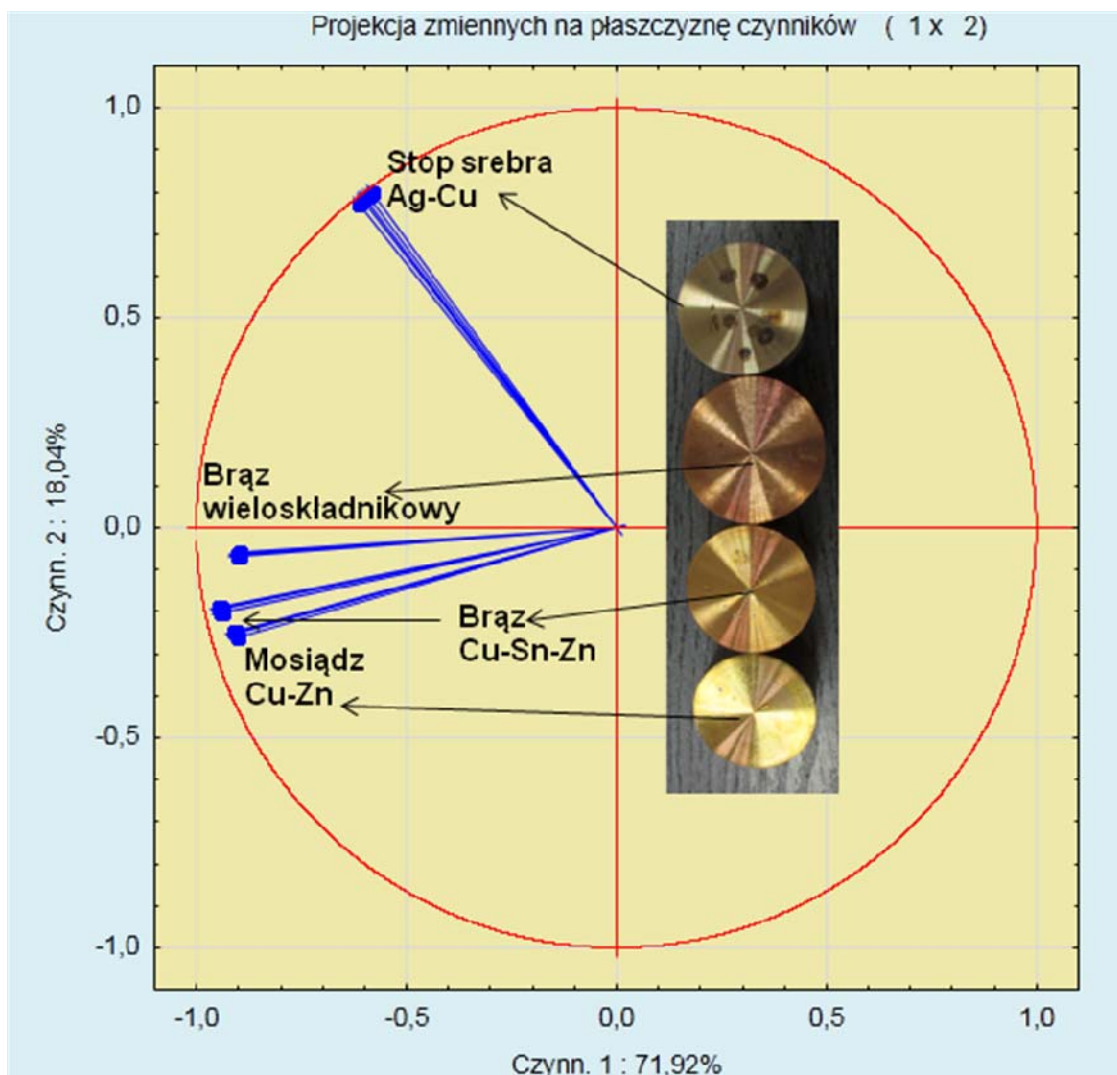
Rys. 8. Analiza czynnikowa procesów korozji na powierzchni szlifowanej i szczotkowanej dla próbek z żeliwa wystawionych na działanie środowiska w Warszawie



Rys. 9. Analiza czynnikowa procesów korozji na powierzchni szlifowanej i szrotkowanej dla próbek i ze stali i z żeliwa wystawionych na działanie środowiska w Warszawie

### 3.6. Analiza głównych składowych i analiza czynnikowa stopów

Do badań stopów zastosowano zarówno analizę głównych składowych, jak i analizę czynnikową. Badania dotyczyły stopów miedzi – stopu srebra z miedzią, mosiądzu dwuskładnikowego (Cu-Zn), brązu trójskładnikowego (Cu-Sn-Zn) i brązu wieloskładnikowego. Na rysunkach 10, 11 i 12 przedstawiono kolejno graficzne wyniki analizy PCA, wyniki analizy czynnikowej oraz wykresu wartości własnych dla dwudziestu zmiennych wejściowych (widm dla czterech stopów – po pięć dla każdego). Wynik analizy PCA przedstawiono w postaci wykresu położenia zmiennych wejściowych, czyli widm w kole jednostkowym. Jeśli zmienne są położone blisko siebie, oznacza to, że są silnie skorelowane. Jak widać, widma położone są w czterech grupach, właściwych dla poszczególnych stopów. Bliskość położenia oznacza podobieństwo składu. „Odstawanie” widm dla stopu Cu-Ag oznacza, że w pozostałych stopach srebra nie ma.

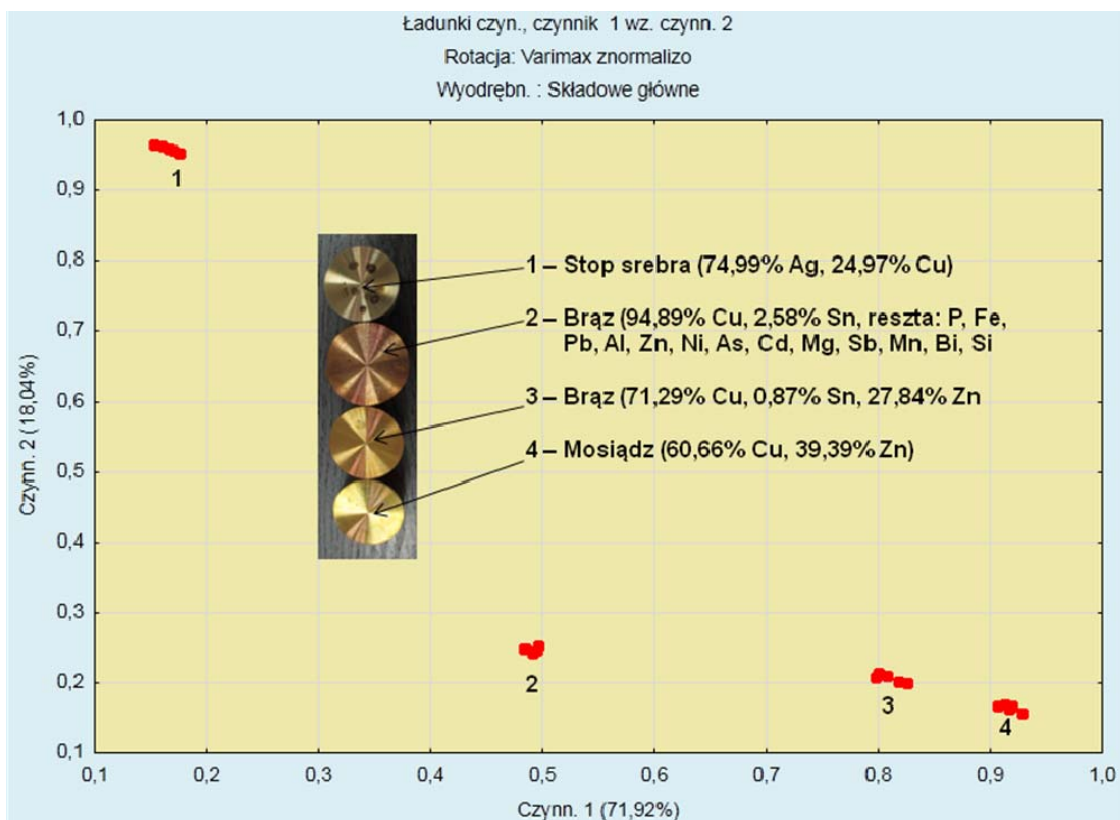


Rys. 10. Analiza głównych składowych dla stopów miedzi

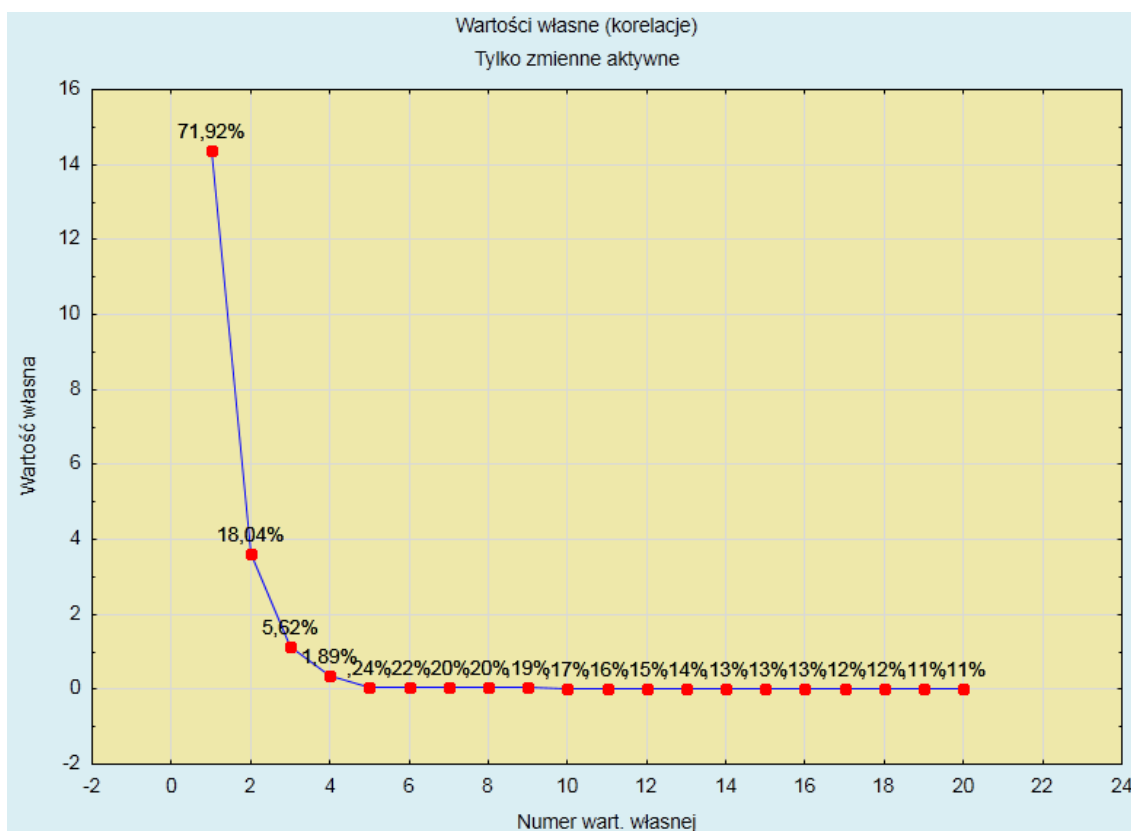
Generalnie, dla przypadków bardzo silnie skorelowanych zmiennych wejściowych nie zawsze warto przedstawiać graficzną analizę PCA – wówczas na wykresie zmiennych w kole jednostkowym są one położone bardzo blisko siebie i trudno uzyskać jasną informację o właściwościach układu wejściowego. Tak się dzieje wtedy, gdy np. pierwsza główna składowa przenosi powyżej 95% informacji.

Zarówno dla analizy PCA, jak FA, dwie pierwsze składowe główne (w programie STATISTICA oznaczane jako czynniki), jak też główne czynniki wspólne przenoszą prawie 90% zasobu informacji, czyli zmienności ogólnej (wariancji) widm wejściowych. Pozostałych osiemnaście – tylko 10% zmienności (rys. 12).



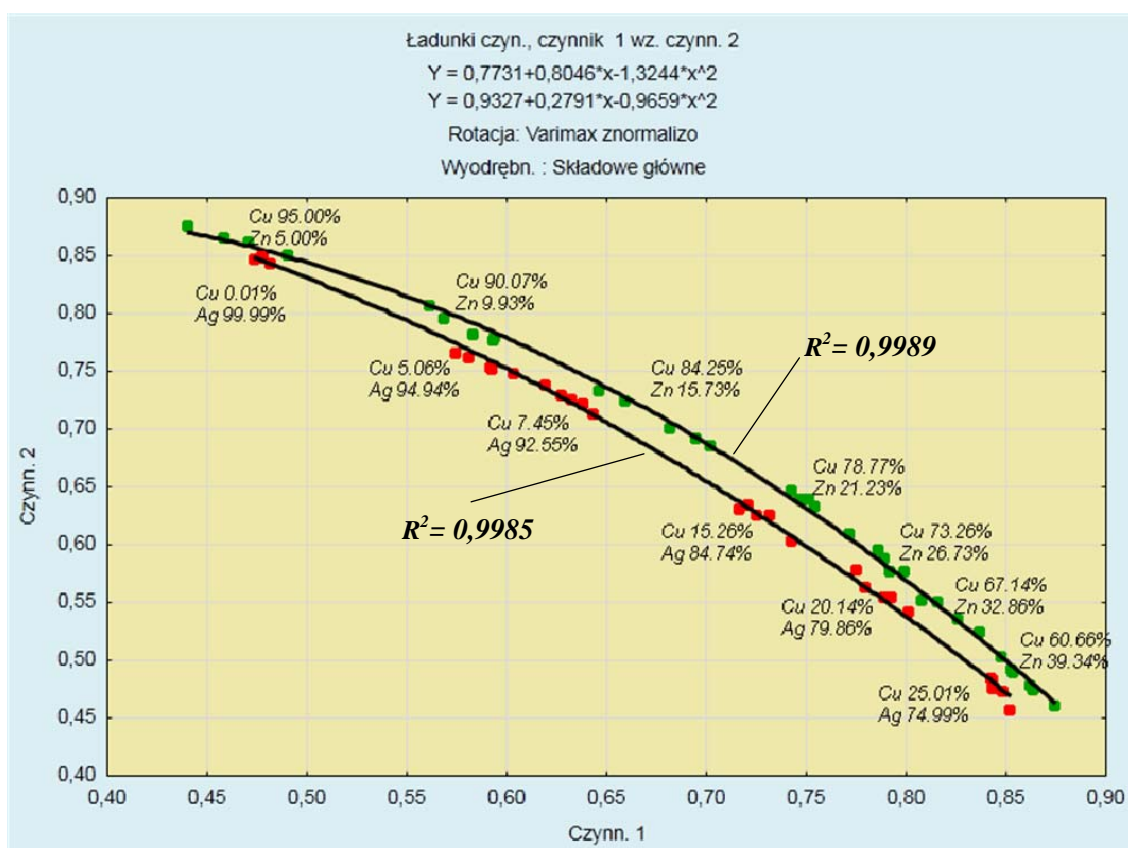


Rys. 11. Analiza czynnikowa dla stopów miedzi



Rys. 12. Wykres osypiska dla 20 widm otrzymanych dla 4 stopów miedzi

Powyższą analizę statystyczną można również zastosować do konstrukcji krzywych kalibracyjnych, w wyniku czego można wykonywać badania ilościowe. Przykład takiego podejścia zaprezentowano na rysunku 13, na którym przedstawiono dwie krzywe kalibracyjne dla dwóch dwuskładnikowych stopów miedzi (ze srebrem i cynkiem).



Rys. 13. Krzywe kalibracyjne dla stopów miedzi otrzymane metodą wieloczynnikową

Wartości R-kwadrat określające stopień dopasowania krzywych trendu do danych wejściowych są bardzo bliskie 1, a więc bardzo wiarygodne. Świadczy to zarazem o tym, że analiza czynnikowa doskonale nadaje się również do konstrukcji krzywych kalibracyjnych, za pomocą których można określać skład chemiczny stopów o nieznanym proporcjach składników, bez konieczności prowadzenia żmudnej pracy polegającej na identyfikacji, właściwym wyborze i wyznaczeniu natężenia linii widmowych w zależności od stężenia składnika stopu, a także znajomości ich stałych atomowych. W omawianym przypadku wystarczy uzyskać tylko widmo.

Jedynym warunkiem dla metody krzywych kalibracyjnych, oprócz identyczności warunków pomiaru przy kalibracji i pomiarze nieznannej próbki, jest identyczny, lub bardzo zbliżony skład wzorców i nieznanego materiału.

## 4. PODSUMOWANIE

---

W pracy skrótowo opisano podstawowe własności dwóch wielowymiarowych metod statystycznych – analizy głównych składowych (PCA) i analizy czynnikowej (FA). Korzystając z komercyjnego pakietu oprogramowania STATISTICA.PL v.10, zastosowano obie metody do analizy widm promieniowania uzyskiwanych w pomiarach metodą spektroskopii emisyjnej. Jako przykłady aplikacji metody PCA i FA przedstawione zostały analizy widm kilku źródeł stacjonarnych oraz widm impulsowych otrzymywanych dla różnych zastosowań metody LIBS (*Laser-Induced Breakdown Spectroscopy*). Ze względu na właściwości obu metod, sprowadzających się, w dużym uproszczeniu, do wyjaśniania struktury, właściwości i ukrytych często podobieństw elementów zbiorów wejściowych, przedstawione w pracy przykłady ewidentnie wykazują bardzo dużą przydatność metod do analizy widm, skutkującej wykrywaniem minimalnych z pozoru różnic w materiałach, z których uzyskiwane są widma.

Jednym z dobrych przykładów są tu wyniki analizy procesów korozji, opisane w rozdziale 3.5. W oparciu o analizę czynnikową potwierdzono fakt, że proces korozji na powierzchni próbki ze stopu żelaza przebiega zależnie od jej struktury powierzchni, analiza pozwala również zróżnicować stan powierzchni (stopień korozji) w zależności od długości trwania procesu korozji. W przypadku pozornie jednakowych materiałów, takich jak papier kredowy, analiza FA pozwoliła wykazać różnice wynikające z miejsca pochodzenia (i zapewne technologii produkcji). Analiza stopów dowiodła z kolei, że metody wieloczynnikowe można z powodzeniem zastosować w badaniach ilościowych.

Podsumowując, zaprezentowane wyniki zastosowania analizy czynnikowej i głównych składowych do badania widm spektroskopii emisyjnej świadczą o bardzo dużej przydatności metod wielowymiarowej analizy statystycznej w określaniu różnic, podobieństwa, a także struktury dowolnego rodzaju materiałów. Metody te można również wykorzystać w badaniach półilościowych, np. do porównania podobieństwa/identyczności nieznanego materiału z wzorcem lub innego rodzaju obiektem odniesienia. To z kolei stwarza możliwość szybkiej analizy autentyczności czy też falsyfikacji badanego materiału pod względem zgodności z charakterystykami znanego obiektu. Metody wielowy-

miarowej analizy statystycznej mogą być stosowane w badaniach materiałowych z wykorzystaniem dowolnego rodzaju widm (absorpcyjnych, Ramana, rentgenowskich, fluorescencyjnych, w podczerwieni, etc.).

## LITERATURA

1. Balicki A.: Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk, 2009.
2. Editors: Miziolek A., Pasleschi V., Schechter I.: Laser-Induced Breakdown Spectroscopy, Cambridge University Press, Cambridge, 2006.
3. Frątczak E., Gołata E., Klimanet T., Ptak-Chmielewska A., Pęczkowski M.: Wielowymiarowa analiza statystyczna. Teoria – przykłady zastosowań z systemem SAS, Szkoła Główna Handlowa w Warszawie, Warszawa, 2009.
4. Panek T.: Statystyczne metody wielowymiarowej analizy porównawczej, Szkoła Główna Handlowa w Warszawie, Warszawa, 2009.
5. Skrzeczanowski W.: Badania dzieł sztuki prowadzone w IOE WAT za pomocą spektroskopii emisyjnej wzbudzanej laserem, Prace Instytutu Elektrotechniki, 234, 217, 2008.
6. Skrzeczanowski W.: Stratygrafia warstw wierzchnich metalowych dzieł sztuki metodą spektroskopii emisyjnej ze wzbudzeniem laserowym, Prace Instytutu Elektrotechniki, 244, 35, 2010.

*Rękopis dostarczono dnia 29.03.2012 r.*

## ANALYSIS OF EMISSION SPECTRA USING MULTIVARIATE STATISTIC METHODS

Wojciech SKRZECZANOWSKI, Andrzej ANTONIK

**ABSTRACT** *The most important characteristics of the Principal Components Analysis and Factorial Analysis are briefly described. They were applied to analysis of emission spectra obtained from stationary and pulse sources. It was stated that processing of spectra using PCA/FA methods reveals similarity level of the spectra, which in turn allows for determination of differences between analysed sources which emission is measured. Results of PCA/FA analysis of LIBS spectra applied to stratigraphy measurements, identification of paper sort, corrosion process of surface layers of iron alloys, discharges in gas and light bulbs are shown. Attempts of the use of the FA method for construction of calibration curves is presented which can be applied in quantitative measurements without need of use of atomic data.*

**Keywords:** *principal components, factorial analysis, emission spectra.*

**Dr inż. Wojciech SKRZECZANOWSKI** jest specjalistą z dziedziny spektroskopii optycznej, metrologii laserowej i systemów zarządzania jakością. Autor i współautor ponad 100 artykułów i referatów prezentowanych w czasopismach i na konferencjach naukowych. Aktualnie w Instytucie Optoelektroniki Wojskowej Akademii Technicznej zajmuje się badaniami z zakresu spektroskopii emisyjnej ze wzbudzeniem laserowym.



**Mgr inż. Andrzej ANTONIK** jest specjalistą z dziedziny elektroniki laserowej i systemów zarządzania jakością. Autor i współautor ponad 50 artykułów, patentów i referatów prezentowanych w czasopismach i na konferencjach naukowych. Aktualnie zajmuje się systemami zarządzania w działalności badawczej.

