

PAWEŁ JANKOWSKI**ANDRZEJ ZIELIŃSKI****WIESŁAW MĄDRY**

Katedra Biometrii, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Analiza interakcji genotyp-środowisko dla pszenicy ozimej z wykorzystaniem metody graficznej biplot typu GGE Część II. Grupy środowisk i stabilność genotypów

Analysis of genotype-environment interaction for winter wheat using graphical GGE biplot method

Part II. Groups of environments and stability of genotypes

Biplot typu GGE jest graficzną metodą analizy tzw. efektów GGE, tzn. efektów głównych genotypów połączonych z efektami interakcji genotypowo-środowiskowych. W pracy metoda GGE biplot została zastosowana do interpretacji danych z 2004 roku, dotyczących plonów rodów pszenicy ozimej na terenie zachodniej Polski. Przedstawiono pewnego rodzaju analizę skupień badanych środowisk, polegającą na poszukiwaniu grup miejscowości (multiśrodowisk) o zbliżonym efekcie interakcji z wszystkimi analizowanymi genotypami. Metodę GGE biplot zastosowano następnie do oceny stabilności genotypów ze względu na wysokość plonu oraz do wskazania genotypów najbardziej zbliżonych do tzw. genotypu idealnego, definiowanego jako genotyp stabilny o największym średnim plonie poprzez środowiska.

Słowa kluczowe: biplot, efekty GGE, interakcja genotypowo-środowiskowa, pszenica ozima, stabilność genotypów

GGE biplot is a graphical method of analysis of the so-called GGE effects, i.e. main genotype effects combined with genotype-environment interaction effects. In the paper the GGE biplot method has been used in order to interpret the 2004 data for grain yield of the winter wheat variety trials in western Poland. First, a kind of the cluster analysis of tested environments has been presented. It led to an indication of the so-called multi-environments characterized by the similar genotype-environment interaction with all analyzed genotypes. Next, the method has been applied to the genotype stability analysis. Genotypes closest to the so-called ideal genotype defined as stable and with highest average yield in all environments have been found.

Key words: biplot, genotype-environment interaction, GGE effects, stability of genotypes, winter wheat

WSTĘP

Praca stanowi kontynuację artykułu Jankowski i wsp. (2006) prezentującego ideę oraz podstawy teoretyczne metody graficznej biplot typu GGE, służącej do analizy interakcji genotypowo-środowiskowej w hodowli roślin, a co za tym idzie, do oceny wartości użytkowej genotypów. Taka ocena należy do najważniejszych i wymagających zadań w zaawansowanych etapach procesu hodowli roślin (Caliński i in., 1987 a, b 1995, 1997; Basford i Cooper, 1998; Yan i in., 2000; Annicchiarico, 2002). Musi być ona oparta na wynikach z wielośrodowiskowej serii doświadczeń genotypowych, czyli doświadczeń wielokrotnych jednorocznych lub wieloletnich. W pracy wykorzystano możliwości metody biplot typu GGE do analizy danych z serii doświadczeń z genotypami, przeprowadzonych w określonych miejscowościach.

Głównym celem oceny wartości użytkowej genotypów w serii doświadczeń wielośrodowiskowych może być wybór odmian najlepiej dostosowanych do większości warunków przyrodniczych danego rejonu docelowego, czyli odmian o szerokiej adaptacji, tzn. najbardziej produktywnych przeciętnie w tym rejonie oraz stabilnych w sensie dynamicznym, czyli reagujących zmianą plonu o taką samą wielkość, o jaką zmienia się średnia wydajność środowisk tego rejonu (Becker i Leon, 1988; Braun i in., 1996; Caliński i in., 1997; Annicchiarico, 2002; Kang, 2002; Mądry i in., 2003; Mądry i Kang, 2005). Wyniki wspomnianej oceny genotypów mogą być także podstawą rejonizacji odmian o wąskiej adaptacji, czyli wysoko produktywnych tylko w pewnych tak zwanych megaśrodkach, tzn. w podrejonach wydzielonych z całego rejonu docelowego (Lin i Binns, 1994; Yan i in., 2000; Yan i Hunt, 2001; Annicchiarico, 2002).

Zgodnie z powyższym główne cele analizy przeprowadzonej w pracy to: a) analiza skupień, czyli podział środowisk na multiśrodkach, w obrębie których poszczególne środowiska charakteryzują się podobnymi interakcjami ze wszystkimi badanymi genotypami; b) ocena stabilności genotypów pod względem wysokości plonu oraz wskazanie genotypów najbardziej zbliżonych do tak zwanego genotypu idealnego, czyli stabilnego o najwyższym średnim plonie spośród wszystkich badanych genotypów (Annicchiarico, 2002; Yan, Kang, 2003).

MATERIAŁ I METODY

W pracy analizie poddane zostały dane dotyczące plonu 56 rodów pszenicy ozimej w 7 miejscowościach (środowiskach). Środowiska różnią się typem gleby, przedplonem, nawożeniem oraz datą siewu i zbioru. Dane uzyskano w 2004 roku z doświadczeń przedrejestrów koordynowanych przez Oddział IHAR w Krakowie.

Oznaczenia miejscowości wraz z ich pełnymi nazwami i województwami, w których te miejscowości są położone, prezentuje tabela 1. Plony pszenicy ozimej w badanych miejscowościach przedstawione są w tabeli 2. Dane tworzą więc macierz o wymiarach 7×56 .

Tabela 1

Miejscowości (środowiska) Locations (environments)							
Oznaczenie Notation	DED	KOC	NAD	POB	SMH	STH	SZD
Miejscowość Location	Dębina	Kobierzyce	Nagradowice	Polanowice	Smolice	Strzelce	Szelejewo
Województwo Province	Pomor.	Dolnośl.	Wielkop.	Małop.	Wielkop.	Łódzkie	Wielkop.

Do analizy macierzy danych wykorzystana została metoda biplot typu GGE, będącą jednym z zastosowań metody graficznej biplot, zaproponowanej m.in. przez Gabriela (1971) i opartej na twierdzeniu o rozkładzie macierzy według wartości osobliwych (np. Rao, 1982, str. 60). Teoria ta ma wiele wspólnego z analizą składowych głównych (Zieliński i in. 2005). Każdy z wykresów biplot tworzony jest na podstawie pierwszych dwóch składowych głównych macierzy danych (w przypadku biplotu typu GGE są to tzw. środowiskowe składowe główne) i tłumaczy ten sam, co one, procent zmienności układu danych. Metoda biplot typu GGE, szczegółowo przedstawiona przez Yana i Kanga (2003; zob. też Yan i in., 2000; Yan, 2001; Yan i Hunt, 2001), bada tzw. efekty GGE (z ang. genotype + genotype-environment), będące sumą efektów głównych genotypów i efektów interakcji genotypowo-środowiskowej. Przez środowiska można w ogólności rozumieć miejscowości badane w jednym roku albo też miejscowości badane w ciągu kilku lat (każda kombinacja miejscowości i roku traktowana jest wówczas jako niezależne środowisko). W pracy przedmiotem analizy są wysokości plonów otrzymane w jednym roku. W takim przypadku uzasadnione jest wymienne stosowanie określeń środowisko i miejscowość (bądź lokalizacja), choć formalnie należałoby w miejsce GGE używać oznaczenia GGL (z ang. genotype + genotype-localization).

Podstawy teoretyczne, proces konstrukcji biplotu typu GGE oraz możliwości jego interpretacji zostały szczegółowo opisane w pierwszej części pracy (Jankowski i in., 2006). Biplot typu GGE składa się z punktów (wektorów) na płaszczyźnie, z których część reprezentuje genotypy, a część środowiska. Efekty typu GGE są funkcjami długości wektorów i kątów między nimi. Zatem analiza interakcji genotypowo-środowiskowej polega na odpowiedniej interpretacji własności geometrycznych i trygonometrycznych obiektów płaszczyzny.

Przy dużej liczbie danych do tworzenia wykresów typu GGE biplot najwygodniej jest zastosować jeden z wielu programów komputerowych. W pracy zostało wykorzystane specjalistyczne oprogramowanie GGEbiplot autorstwa Yana (Yan, 2001–2004), automatycznie generujące wykresy typu biplot dla kilku modeli GGE, opartych na standaryzowanej lub niestandaryzowanej macierzy danych. Posiada ono wiele narzędzi ułatwiających analizę wykresów i wyciąganie wniosków dotyczących zależności między genotypami a środowiskami. Dodatkowo program GGEbiplot umożliwia analizę modeli takich jak PCA czy AMMI. Przydatnym do tworzenia wykresów GGE biplot może być również bezpłatny mikroprogram, działający w arkuszu Excel, stworzony przez Lipkovicha i Smitha (Lipkovich, Smith, 2002).

Tabela 2

Plony pszenicy ozimej 56 rodów w 7 miejscowościach
Yield of 56 genotypes of winter wheat in 7 locations

Miejscowość Location	DED	KOC	NAD	POB	SMH	STH	SZD
Ród Genotype							
1	2	3	4	5	6	7	8
CHD_342/98	10,72	12,50	10,17	10,13	11,63	9,64	8,75
DAD_2522/00	10,31	11,26	10,29	9,99	10,81	9,81	8,16
CHD_507/00	10,33	10,77	10,39	10,48	10,66	9,09	7,91
DAD_408/99	10,05	11,57	10,55	11,39	10,56	9,72	7,05
LAD_425/99	10,03	11,45	10,20	9,50	10,50	10,75	8,21
DAD_128/00	10,17	10,71	9,82	9,20	10,74	9,72	7,85
LAD_758/00	9,85	11,86	9,84	10,67	11,09	10,06	8,64
LAD_317/99	9,23	10,60	9,62	10,82	10,80	8,50	7,29
AND_886/98	9,18	10,17	9,68	9,25	10,02	9,25	7,99
AND_418/99	10,99	12,08	10,73	10,42	11,01	10,11	9,27
AND_981/97	9,72	12,08	10,54	10,45	9,943	10,12	8,09
AND_1237/99	10,10	11,37	10,66	10,15	11,47	9,70	8,28
NAD_016/98	9,53	11,35	10,63	10,59	10,11	10,58	8,24
NAD_017/98	9,73	10,78	10,98	9,80	10,74	10,44	7,73
NAD_038/98	9,12	11,37	10,02	10,28	8,96	8,60	7,57
NAD_088/98	9,74	11,92	10,63	10,86	10,36	9,34	8,73
HEC_659/99	10,31	11,98	10,39	10,47	11,12	9,55	8,36
HEC_780/00	10,47	11,56	10,53	11,14	10,89	9,63	8,21
KOC_1437/01_RS.9	10,05	11,80	10,37	11,33	11,00	9,20	8,36
KOC_1491/01_926-	10,37	10,83	10,00	9,08	10,67	9,61	8,20
KOC_2047/01_949-	10,56	11,37	10,95	11,08	11,13	10,09	8,77
KOC_2267/01_RS.9	9,57	11,25	9,69	10,81	10,76	9,23	8,06
KOC_2374/01_991-	10,00	11,68	10,19	10,30	10,54	10,08	8,07
KOC_2339/01_984-	9,89	11,77	9,43	11,40	11,17	10,23	7,75
HRSM_26	10,33	11,39	10,38	11,81	11,23	10,09	7,39
HRSM_91	10,18	11,60	10,38	11,77	11,04	8,92	8,36
HRSM_111	10,00	10,52	10,17	8,84	11,12	9,52	7,49
HRSM_122	9,11	11,17	10,09	10,50	10,37	9,29	7,32
HRSM_204	10,38	10,63	10,25	9,69	10,33	10,23	7,09
HRSM_213	9,39	10,61	9,63	9,97	10,68	9,75	8,15
SMH_6847	10,33	10,98	10,48	10,71	11,56	10,09	8,00
SMH_6868	9,19	9,97	9,12	9,95	9,42	9,06	7,06
STH_5441	9,83	10,23	10,03	8,81	10,99	9,65	8,32
STH_1599	9,47	10,46	8,80	9,74	9,63	8,95	7,41
STH_1427	10,31	11,84	10,62	11,44	10,96	10,82	9,10
STH_1596	9,52	11,64	8,93	10,20	10,84	10,63	8,82
STH_5338	10,42	11,82	10,45	10,68	10,75	10,52	9,01
STH_5504	8,83	9,76	10,22	8,94	9,81	9,17	8,88
STH_43	10,23	10,24	9,85	9,53	10,45	9,34	8,80
STH_175	9,29	11,65	9,59	9,29	11,21	11,09	8,67
POB_126/00	10,52	11,77	10,71	9,83	10,62	10,48	8,19
POB_470/01	9,93	11,02	10,43	9,75	10,69	9,98	8,98
POB_495/01	10,17	11,57	10,89	10,45	10,91	9,76	7,52
POB_452/01	10,13	12,18	10,27	9,75	10,88	10,12	8,66
POB_654/95	10,10	12,06	10,40	8,99	10,79	10,76	9,00
MOB_3926/99	10,76	12,37	9,41	10,77	10,80	10,87	9,57
MOB_1737/00	9,16	11,63	9,47	10,23	10,07	8,79	7,86
MOB_2355/97	9,24	11,43	10,00	10,75	9,84	9,38	9,20
MOB_3191/99	10,19	12,24	10,25	10,27	10,80	9,95	9,85

c.d. Tabela 2

1	2	3	4	5	6	7	8
MOB_3963/99	10,41	11,63	9,73	9,50	10,41	10,40	8,11
MOB_3977/99	9,76	11,43	10,27	11,79	10,80	10,31	9,12
SZD_601	10,88	11,52	9,89	10,47	11,32	9,28	8,20
SZD_597	9,49	11,29	9,57	11,30	10,65	10,07	8,823
Tonacja	9,78	11,08	10,24	9,89	10,52	10,54	8,30
Kobra	10,03	11,22	9,87	9,70	10,83	9,71	7,83
Korweta	9,02	11,24	10,06	10,95	10,36	9,11	8,67

WYNIKI I DYSKUSJA

Analiza środowiskowych składowych głównych dla ocen efektów GGE otrzymanych na podstawie danych z tabeli 2 wskazuje, że powinno się uwzględnić trzy pierwsze składowe, tłumaczące razem 75% zmienności układu danych. Pominięcie pozostałych składowych głównych przyniesie utratę 25% informacji (tab. 3).

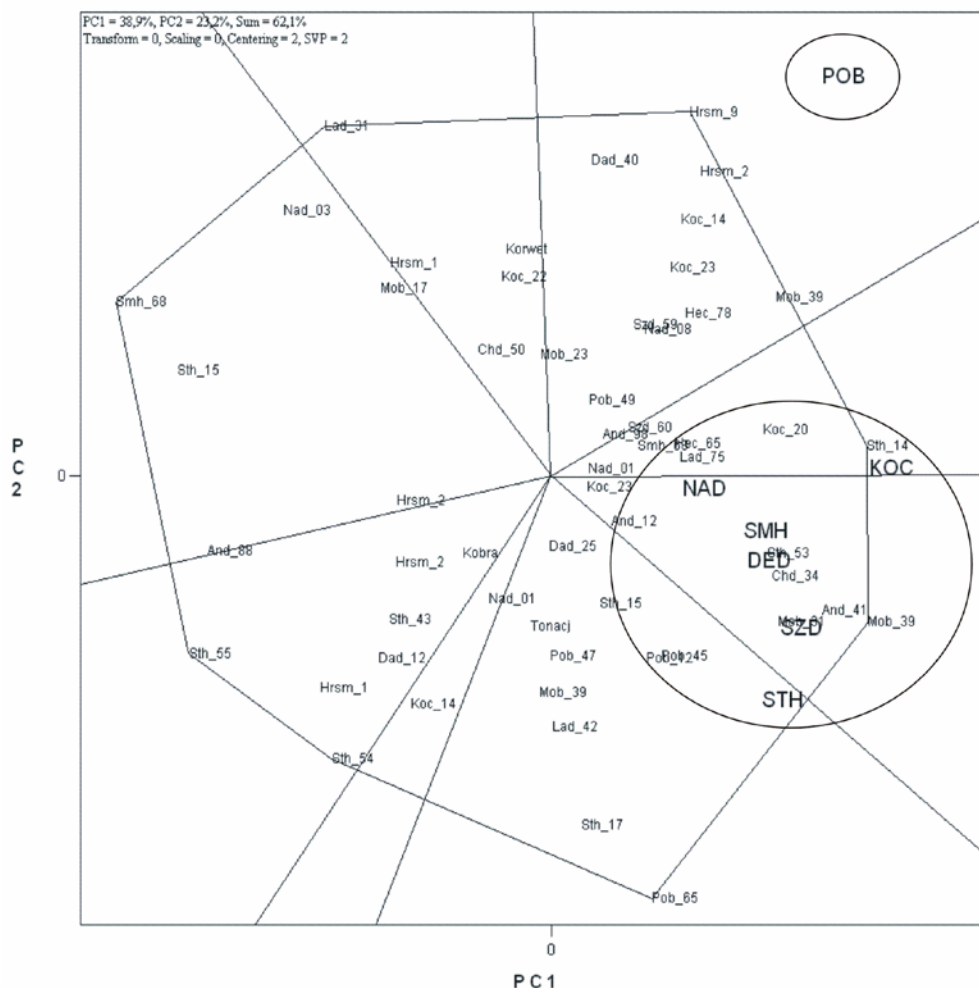
Tabela 3

Procent zmienności danych tłumaczony przez kolejne składowe główne macierzy danych
Percentage of the variability of the data explained by the successive principal components

Składowa główna — Principal component	1	2	3
Procent zmienności — Percentage of variability	39%	23%	13%

Metoda GGE biplot może być bardzo przydatna w końcowych etapach hodowli zbóż, a także innych gatunków roślin. Pierwszym z jej głównych zastosowań jest podział środowisk na grupy reprezentujące, możliwie wiernie, obszary potencjalnej uprawy wyhodowanych odmian, czyli tzw. rejonów docelowych. Innymi słowy wykresy typu GGE biplot umożliwiają wyznaczenie tzw. multiśrodowisk, w obrębie których środowiska cechują się podobnymi interakcjami ze wszystkimi badanymi genotypami.

Rysunek 1 jest wykresem typu biplot utworzonym na podstawie dwóch pierwszych składowych głównych i przedstawia podział wszystkich środowisk na dwa główne multiśrodowiska. Z powodu znacznej liczby badanych genotypów przypisywanie każdej z miejscowości rodu o największym w niej efekcie GGE odbywa się w sposób nieco bardziej złożony, niż było to przedstawione w części I pracy (Jankowski i in., 2006). Mianowicie odnajdywane są genotypy najbardziej oddalone od początku układu współrzędnych w ten sposób, by wielokąt, którego wierzchołkami będą te genotypy, zawierał w swoim wnętrzu wszystkie genotypy pozostałe. Linie prostopadłe oddzielają miejscowości, w których największy plon osiąga ród będący wierzchołkiem wielokąta zawartym w danej części wykresu. I tak z rysunku 1 wynika, że ród Hrsm_9 (Hrsm_91) daje najwyższy plon w miejscowości POB, zaś rody Sth_14 (STH_1427) oraz Mob_39 (Mob_3926/99) we wszystkich pozostałych miejscowościach. Z drugiej strony rody znajdujące się w pięciu pozostałych wierzchołkach są najgorszymi dla części lub dla wszystkich środowisk.



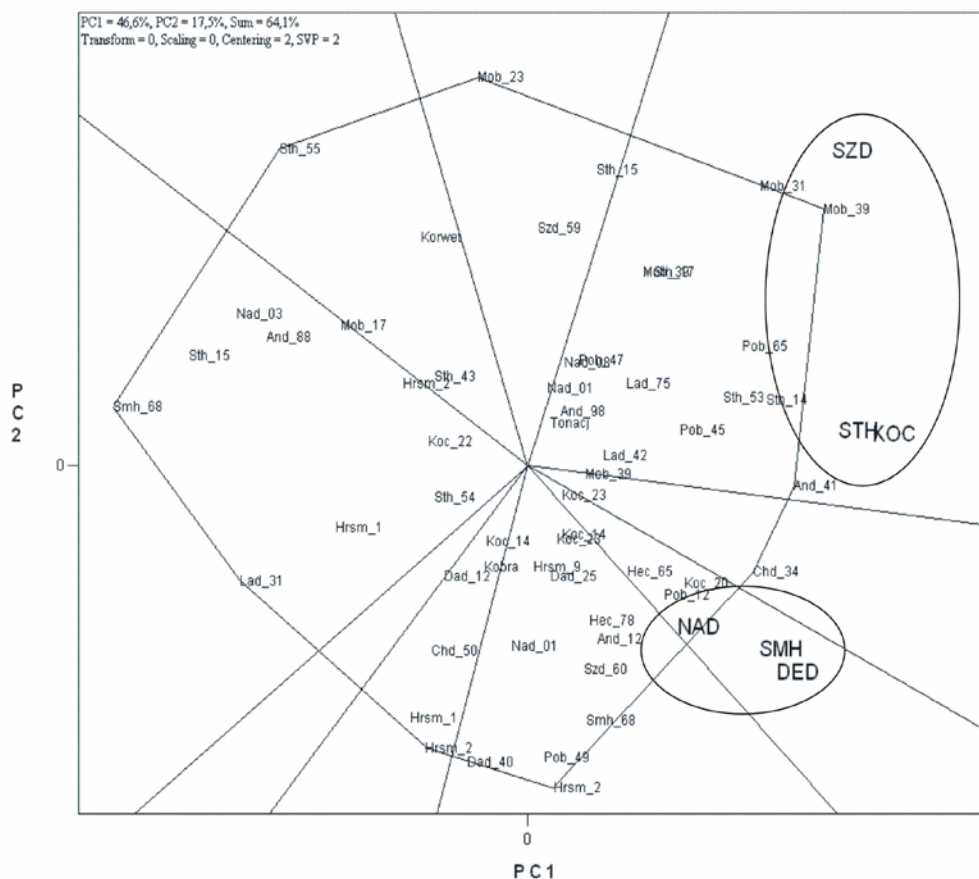
Rys. 1. Podział środowisk na dwa główne multiśrodowiska na podstawie wykresu biplot typu GGE opartego na pierwszej i drugiej składowej głównej

Fig. 1. Two multi-environments as the result of grouping based on the GGE biplot drawn by means of the first and second principal components

Przyjrzyjmy się dokładniej wyodrębnionej grupie środowisk. Na rysunku 2 sporządzonym po wyłączeniu z macierzy danych kolumny związanej ze środowiskiem POB ujawnia się głębsza struktura grupy, niewidoczna wcześniej z powodu znacznych różnic między pierwotnie otrzymanymi grupami. Okazuje się, że grupę sześciu miejscowości możemy podzielić na jeszcze dwie podgrupy, z których jedną tworzą środowiska KOC, SZD i STH, a drugą DED, NAD i SMH.

Otrzymane wyniki potwierdza dodatkowy biplot typu GGE sporządzony na podstawie drugiej i trzeciej składowej głównej i prezentowany na rysunku 3. Zasadność jego analizy

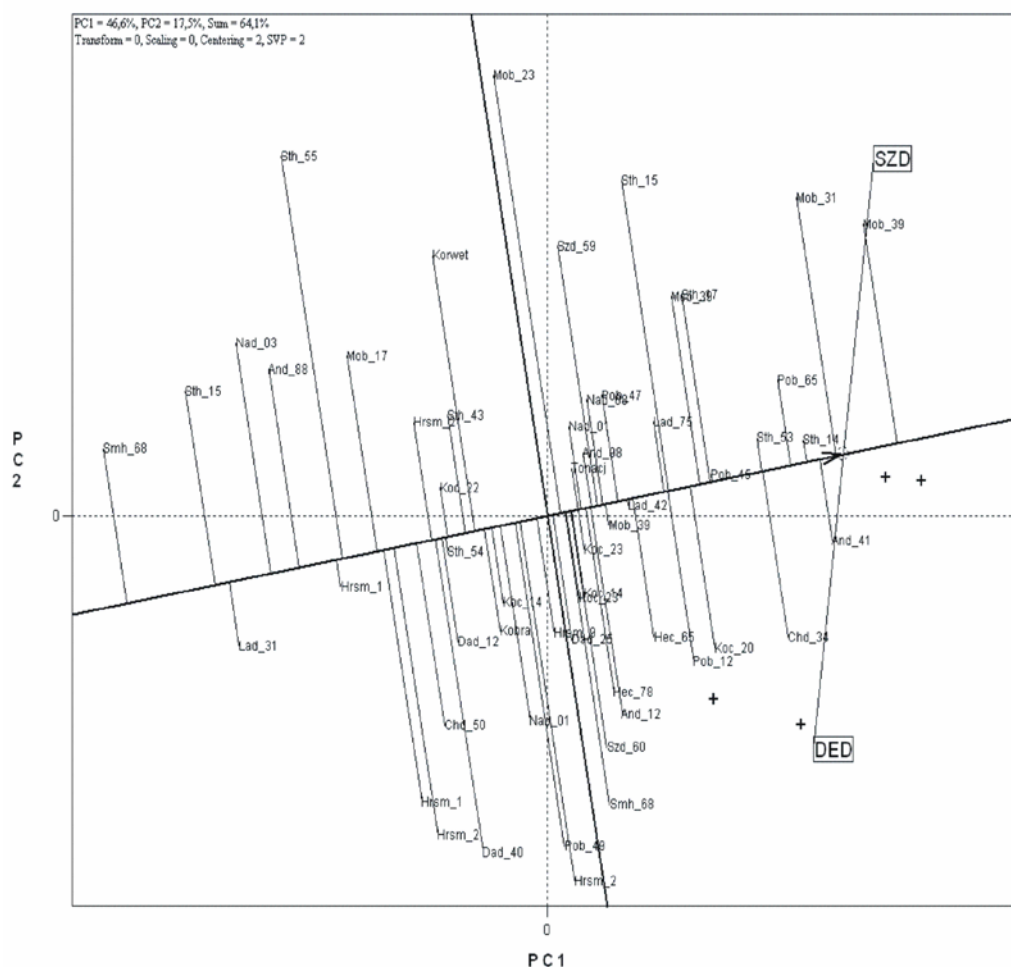
tłumaczy fakt, że pierwsza składowa główna nie ujawnia zróżnicowania miejscowości, wszystkie, bowiem koncentrują się wokół wartości pierwszej składowej głównej równej 3 (patrz rys. 1). Jak widać, dodatkowy biplot ukazuje ten sam podział na grupy środowisk jak przedstawiony powyżej.



Rys. 2. Dalszy podział jednego z pierwotnych multiśrodowisk na podstawie wykresu biplot typu GGE opartego na pierwszej i drugiej składowej głównej

Fig. 2. Further grouping applied to one of the original multi-environments based on the GGE biplot drawn by means of the first and second principal components

W przypadku obu wydzielonych trójelementowych grup środowisk można przypuszczać, że w ich obrębie nie występuje interakcja genotypowo-środowiskowa. Gdyby taka hipoteza została potwierdzona danymi z kolejnych lat, można by wybrać po jednym środowisku reprezentującym daną podgrupę. Jak pokazuje rysunek 4, środowiskami najlepiej spełniającymi tę rolę byłyby SZD oraz DED o najdłuższych w obrębie swoich multiśrodowisk wektorach w przestrzeni pierwszej i drugiej składowej głównej. Stąd



Rys. 6. Porównanie efektu GGE poszczególnych genotypów w środowiskach SZD i DED
Fig. 6. Comparison of the genotypes according to their GGE effects in the SZD and DED environments

Dodatkowe dwie linie na wykresie, równoległe do osi średniego środowiska, wydzielają genotypy o dużej stabilności (genotypy zawarte pomiędzy tymi liniami). Podobnie dodatkowa linia prostopadła w sposób przykładowy pokazuje możliwy wybór rodów zbliżonych do genotypu idealnego. Kryteria wyboru genotypów „wystarczająco” stabilnych lub zbliżonych do genotypu idealnego, jako w znacznym stopniu arbitralne, wymagają doświadczenia hodowcy i jego wiedzy na temat badanych rodów i środowisk.

Fig. 7. Stability analysis of the genotypes. The arrow visible in the middle of the circle shows the location of the ideal genotype. The arrow is situated on the average environment coordinate (AEC) abscissa representing stable genotypes. The line perpendicular to the AEC abscissa shows the directions of instability increase indicated by the arrows

PODSUMOWANIE

72

miejscowości na terenie zachodniej Polski, potwierdza swą skuteczność przy ocenie przydatności tych rodów do hodowli. Duża łatwość interpretacyjna wykresów biplot typu GGE umożliwiła podział miejscowości na grupy jednorodne pod względem interakcji genotypowo-środowiskowej (wydzielone zostały trzy takie grupy). Przy pomocy tej metody można też wskazać genotypy o dużej stabilności i charakteryzujące się wysokim plonem.

Pewną wadą omawianej metody jest to, że wykresy oddają rzeczywiste zależności w sposób jedynie przybliżony (jest to wada każdej metody opartej na analizie dwóch lub trzech pierwszych składowych głównych). Dotyczy to szczególnie przypadków, gdy porównywane obiekty różnią się w niewielkim stopniu. Uniknięcie dowolności interpretacyjnych w takich sytuacjach wymaga dodatkowego doświadczenia hodowcy oraz jego wiedzy na temat badanych rodów i środowisk.

LITERATURA

- Annicchiarico P. 2002. Genotype $\times$ environment interactions — challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant Production and Protection Paper-174. Food and Agriculture Organization of the United Nations, Rome.
- Basford H. E., Cooper M. 1998. Genotype $\times$ environment interactions and some considerations of their implications for wheat breeding in Australia. *Austr. J. Agric. Res.* 49: 153 — 174.
- Becker H. C., Leon J. 1988. Stability analysis in plant breeding. *Plant Breeding* 101, 1 — 23.
- Braun H. J., Rajaram S., van Ginkel M. 1996. CIMMYT's approach to breeding for wide adaptation. *Euphytica* 92, 175 — 183.
- Caliński T., Czajka S., Kaczmarek Z. 1987 a. A model for the analysis of a series of experiments repeated at several places over a period of years. I. Theory. *Biul. Oceny Odmian* 17-18: 7 — 33.
- Caliński T., Czajka S., Kaczmarek Z. 1987 b. A model for the analysis of a series of experiments repeated at several places over a period of years. II. Example. *Biul. Oceny Odmian* 17-18: 35 — 71.
- Caliński T., Czajka S., Kaczmarek Z. 1997. A multivariate approach to analysing genotype-environment interactions. In: *Advances in Biometrical Genetics*. Krajewski P., Kaczmarek Z. (Ed), Poznań: 3 — 14.
- Caliński T., Czajka S., Kaczmarek Z., Krajewski P., Siatkowski I. 1995. SERGEN — a computer program for the analysis of series of variety trials. *Biuletyn Oceny Odmian* 26-27: 39 — 41.
- Gabriel K. R. 1971. The biplot-graphic display of matrices with application to principal component analysis. *Biometrika* 58: 453 — 467.
- Jankowski P., Zieliński A., Mądry W. 2006. Analiza interakcji genotyp-środowisko dla pszenicy ozimej z wykorzystaniem metody graficznej biplot typu GGE. Część I. Podstawy teoretyczne. *Biul. IHAR* 240/241: .
- Kang M. S. 2002. Genotype-environment interaction: Progress and prospects. In: *Quantitative genetics, genomics and plant breeding*. M. S. Kang (Ed), CAB International Wallingford, UK: 221 — 243.
- Lin C. S., Binns M. R. 1994. Concepts and methods for analyzing regional trial data for cultivar and location selection. *Plant Breeding Reviews* 12: 271 — 297.
- Lipkovich H., Smith P. 2002. Biplot and singular value decomposition Macros for Excel®. *J. of Statist. Software* 7 (5): 1 — 15.
- Mądry W., Kang M. S. 2005. Scheffé-Calinski and Shukla Models: Their Interpretation and Usefulness in Stability and Adaptation Analyses. *Journal of Crop Improvement*, Vol. 14, 1/2: (w druku).
- Mądry W., Drzazga T., Laudański Z., Rajfura A., Krajewski P., Sieczko L. 2003. Statystyczna analiza adaptacji genotypów pszenicy ozimej na podstawie mieszanych modeli genotypowo-środowiskowych. *Fragmenta Agronomica* 80: 10 — 31.
- Rao C. R. 1982. *Modele liniowe statystyki matematycznej*. PWN, Warszawa: 60.

- Yan W. 2001. GGE biplot — a Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agron. J.* 93: 1111 — 1118.
- Yan W. 2001-2004. Strona World Wide Web: <http://www.ggebiplot.com/biplot.htm>.
- Yan W., Hunt L.A., Sheng Q., Szlavnick Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40: 597 — 605.
- Yan W., Hunt L.A. 2001. Interpretation of genotype-environment interaction for winter wheat yield in Ontario. *Crop Sci.* 41, 19 — 25.
- Yan W., Kang M. S. 2003. GGE biplot analysis: a graphical tool for breeders, geneticists and agronomists. CRC Press, Boca Raton, FL.
- Zieliński A., Jankowski P., Mądry W. 2005. Biplot typu GGE jako narzędzie do graficznej analizy danych z serii doświadczeń z genotypami. *XXXV Coll. Biom.*, 207 — 224.