

WIESŁAW MĄDRY
ADRIANA DEREJKO

Katedra Doświadczalnictwa i Bioinformatyki
Szkoła Główna Gospodarstwa Wiejskiego, Warszawa

Metody statystyczne analizy danych w kompletnej klasyfikacji Odmiana × Agrotechnika × Miejscowość × Rok (G×M×L×Y) z PDOiR

Statistical methods for data analysis in the complete classification Cultivar × Crop Management × Location × Year (G×M×L×Y) from PVTs

Doświadczenia porejestrowe są prowadzone od 1998 roku w ramach Porejestrowego Doświadczalnictwa Odmianowego i Rolniczego (PDOiR). W tym systemie wykonuje się serie doświadczeń odmianowych i odmianowo-agrotechnicznych. Doświadczenia PDOiR stanowią ostatni etap wdrażania postępu biologicznego do praktyki rolniczej. Pod względem merytorycznym i metodycznym system PDOiR jest koordynowany przez Centralny Ośrodek Badania Odmian Roślin Uprawnych. Realizacja serii doświadczeń w PDOiR odbywa się na terenie całego kraju w środowiskach (stacjach doświadczalnych) dobrze reprezentujących przestrzenną zmienność agro-ekosystemów w najważniejszych rejonach uprawy danego gatunku roślin w Polsce. W niniejszej pracy przedstawione zostały podstawy teoretyczne proponowanych, klasycznych, adaptowanych i adekwatnie rozwiniętych metod statystycznych, tj. łączna analiza wariancji, szczegółowe porównania wielokrotne, analiza AMMI oraz analiza skupień. Ponadto zaprezentowano użyteczność tych metod do wykorzystania ich w analizie danych w kompletnej klasyfikacji Odmiana × Agrotechnika × Miejscowość × Rok, pochodzących z PDOiR.

Słowa kluczowe: model mieszany ANOVA, łączna analiza wariancji, analiza AMMI, analiza skupień, Porejestrowe Doświadczalnictwo Odmianowe i Rolnicze (PDOiR)

Postregistration Experiments are conducted since 1998 as part of the Post-registration Variety Testing System (PVTs). In this system series of varietal and varietal-agronomic experiments are performed. Experiments in PVTs represent the last stage in the implementation of biological progress to agricultural practice. PVTs is coordinated by the Research Centre for Cultivar Testing in terms of design and methodology. The implementation of a series of experiments in PVTs is held throughout the country in environments (Cultivar Testing Stations) representing well the spatial variability of agro-ecosystems in the major growing areas of the particular plant species in Poland. In this paper the

theoretical basis is proposed, as well as classical, adapted and adequately developed statistical methods, i.e. the combined analysis of variance, multiple comparison, AMMI analysis and cluster analysis are presented. Moreover, the usefulness of these methods for analyzing the data in the complete classification of Cultivar \times Crop Management \times Location Year, coming from PVTS is presented.

Key words: mixed linear ANOVA model, combined analysis of variance, AMMI analysis, cluster analysis, Post-registration Variety Testing System (PVTS)

ŁĄCZNA ANALIZA WARIANCJI

Model mieszany łącznej analizy wariancji

Średnie z powtórzeń (tzw. średnie poprawione) badanej cechy rolniczej roślin, zwykle plonu, obliczone na podstawie danych z serii doświadczeń dwuczynnikowych w Porejstrowym Doświadczalnictwie Odmianowym i Rolniczym tworzą 4-kierunkową kompletną klasyfikację o postaci Odmiana \times Agrotechnika \times Miejscowość \times Rok ($G \times M \times L \times Y$). Łączna analiza wariancji (ang. combined ANOVA) tych średnich poprawionych może być wykonana na podstawie odpowiedniego modelu liniowego. Taki model jest stosowany w dwustopniowej analizie danych z serii doświadczeń odmianowych lub odmianowo-agrotechnicznych (Smith i in., 2005; Piepho i Möhring 2006; Möhring i Piepho, 2009; Welham i in., 2010). Analizę danych z serii doświadczeń PDOiR w podejściu dwustopniowym przeprowadzimy w dwóch etapach.

W rozpatrywanych badaniach w pierwszym etapie wykonamy oddzielnie dwuczynnikowe analizy wariancji dla surowych danych, pochodzących z doświadczeń pojedynczych, planowanych w niekompletnym układzie split-block w każdej miejscowości i roku oceny odmian. Analiza wariancji danych z doświadczeń pojedynczych PDOiR, oparta na modelu stałym (odmiana i intensywność agrotechniki są traktowane jako czynniki stałe — Welham i in., 2010) dla układu split-block (Elandt, 1964) jest przedstawiona w tabeli 1. Algorytmy średnich kwadratów dla wszystkich źródeł zmienności są opracowane i oprogramowane w pakiecie SERGEN 4.0 (Caliński i in., 1998). Celem tego postępowania jest wyznaczenie średnich poprawionych (estymatorów BLUE) badanej cechy dla kombinacji odmian i intensywności agrotechniki oraz średnich kwadratów dla trzech błędów doświadczalnych w każdym środowisku (Piepho i in., 2008). Średnie poprawione z wielokrotnej serii doświadczeń PDOiR, wykonanej w kilku latach, porządkuje się w postaci klasyfikacji $G \times M \times L \times Y$ przedstawiono w tabeli 1, gdzie: b_{ij} jest liczbą bloków niekompletnych ze względu na odmiany w pojedynczym doświadczeniu wykonanym w i -tym roku oraz j -tej miejscowości, K_{ij} jest liczbą badanych odmian w pojedynczym doświadczeniu w i -tym roku oraz j -tej miejscowości, Λ jest liczbą badanych poziomów intensywności agrotechniki, $v_{e_1(ij)}$ jest liczbą stopni swobody dla błędu 1 w analizie wariancji danych z doświadczenia w i -tym roku oraz j -tej miejscowości, $s_{e_1(ij)}^2$ jest średnim kwadratem błędu 1 w niekompletnym układzie split-block dla i -tego roku oraz j -tej miejscowości, $v_{e_2(ij)}$ jest liczbą stopni swobody dla błędu 2 w analizie wariancji danych z doświadczenia w i -tym roku oraz j -tej miejscowości, $s_{e_2(ij)}^2$ jest średnim kwadratem błędu 2 w niekompletnym układzie split-block dla i -tego roku oraz j -tej miejscowości, $v_{e_3(ij)}$ jest

liczbą stopni swobody dla błędu 3 w analizie wariancji danych z doświadczenia w i -tym roku oraz j -tej miejscowości, $s_{e_3(ij)}^2$ jest średnim kwadratem błędu 3 w niekompletnym układzie *split-block* dla i -tego roku oraz j -tej miejscowości.

Tabela 1

Tabela analizy wariancji danych z pojedynczego, w i -tym ($i=1,2, \dots, I$) roku oraz j -tej ($j=1,2, \dots, J$) miejscowości, doświadczenia dwuczynnikowego w Porejestrowym Doświadczalnictwie Odmianowym i Rolniczym, założonego w niekompletnym układzie *split-block*
Analysis of variance, data from a single two-factorial experiment in an incomplete *split-block* setup, in an i^{th} year ($i = 1, 2, \dots, I$) and j^{th} location ($j = 1, 2, \dots, J$), performed within Post-registration Variety Testing System

| Źródła zmienności Sources of variation | Stopnie swobody Degrees of freedom | Średnie kwadraty Mean squares | $F_{\text{emp.}}$ |
|--|---------------------------------------|----------------------------------|-------------------------|
| Bloki Blocks | $b_{ij}-1$ | $M1_{ij}$ | |
| Odmiana Cultivar | $K_{ij}-1$ | $M2_{ij}$ | $M2_{ij}/s_{e_1(ij)}^2$ |
| Błąd 1 Error 1 | $V_{e_1(ij)}$ | $M3_{ij}=s_{e_1(ij)}^2$ | |
| Agrotechnika Agrotechnical procedures | $\Lambda-1$ | $M4_{ij}$ | $M4_{ij}/s_{e_2(ij)}^2$ |
| Błąd 2 Error 2 | $V_{e_2(ij)}$ | $M5_{ij}=s_{e_2(ij)}^2$ | |
| Odmiana \times Agrotechnika Cultivar \times Agrotechnical procedures | $(K_{ij}-1)(\Lambda-1)$ | $M6_{ij}$ | $M6_{ij}/s_{e_3(ij)}^2$ |
| Błąd 3 Error 3 | $V_{e_3(ij)}$ | $M7_{ij}=s_{e_3(ij)}^2$ | |

W drugim etapie zastosujemy łączną analizę wariancji średnich poprawionych w klasyfikacji $G \times M \times L \times Y$, opartą na odpowiednim modelu analizy wariancji. W rozpatrywanym przypadku taki model powinien zawierać efekty główne i interakcyjne czynników badanych w wielokrotnych seriach doświadczeń PDOiR w wielu latach (uwzględnionych w omawianej klasyfikacji średnich poprawionych) oraz trzy błędy doświadczalne, odpowiednie dla układu *split-block*. W rozważanej metodyce analizy statystycznej średnich poprawionych, trzy czynniki, tj. odmiana, intensywność agrotechniki i miejscowość traktujemy jako stałe, zaś rok uznajemy jako czynnik losowy (McIntosh, 1983; Annicchiarico i in., 2006, 2010 a). Wobec tego, model mieszany łącznej analizy wariancji (standardowy liniowy model mieszany lub mieszany model ANOVA, ang. ANOVA type mixed model — Smith i in., 2005; Hu i Spilke, 2011) dla rozważanych średnich poprawionych w klasyfikacji $G \times M \times L \times Y$ z serii doświadczeń PDOiR, planowanych w układzie *split-block*, ma następującą postać (McIntosh, 1983; Brancourt-Hulmel i in., 2003; Annicchiarico i in., 2010 a):

$$X_{ijkl} = m + Y_i + L_j + YL_{ij} + G_k + GY_{ik} + GL_{jk} + GLY_{ijk} + e_{1ijk} + M_l + MY_{il} + ML_{il} + MLY_{ijl} + e_{2ijl} + GM_{kl} + GMY_{ikl} + GML_{ljk} + GMLY_{ijkl} + e_{3ijkl} \quad (1)$$

gdzie:

X_{ijkl} — jest średnią poprawioną obserwacji plonu z n powtórzeń dla 4-czynnikowej kombinacji i -tego ($i=1,2, \dots, I$) roku, j -tej ($j=1,2, \dots, J$) miejscowości, k -tej ($k=1,2, \dots, K$) odmiany oraz l -tej ($l=1,2, \dots, L$) intensywności agrotechniki,

m — jest średnią ogólną, Y_i jest losowym efektem głównym i -tego roku, L_j jest stałym efektem głównym j -tej miejscowości, G_k jest stałym efektem głównym k -tej odmiany, M_l jest stałym efektem głównym l -tej intensywności agrotechniki, YL_{ij} jest losowym efektem interakcji podwójnej i -tego roku z j -tą miejscowością, GY_{ik} jest losowym efektem interakcji podwójnej k -tej odmiany z i -tym rokiem, GL_{jk} jest stałym efektem interakcji podwójnej k -tej odmiany z j -tą miejscowością, MY_{il} jest losowym efektem interakcji podwójnej l -tej intensywności agrotechniki z i -tym rokiem, ML_{jl} jest stałym efektem interakcji podwójnej l -tej intensywności agrotechniki z j -tą miejscowością, GM_{kl} jest stałym efektem interakcji podwójnej k -tej odmiany z l -tą intensywnością agrotechniki, GLY_{ijk} jest losowym efektem interakcji potrójnej k -tej odmiany z j -tą miejscowością oraz i -tym rokiem, GMY_{ikl} jest losowym efektem interakcji potrójnej k -tej odmiany z l -tą intensywnością agrotechniki oraz i -tym rokiem, GML_{jkl} jest stałym efektem interakcji potrójnej k -tej odmiany z l -tą intensywnością agrotechniki oraz j -tą miejscowością, MLY_{ijl} jest losowym efektem interakcji potrójnej l -tej intensywności agrotechniki z j -tą miejscowością oraz i -tym rokiem, $GMLY_{ijkl}$ jest losowym efektem interakcji poczwórnej k -tej odmiany z l -tą intensywnością agrotechniki, j -tą miejscowością oraz i -tym rokiem, e_{1ijk} , e_{2ijl} oraz e_{3ijkl} są średnimi (poprzez powtórzenia) błędami w układzie split-block, odpowiednio średnim błędem pierwszym, drugim i trzecim, związanymi z obserwacjami X_{ijkl} .

W mieszanym modelu ANOVA (1) przyjmujemy standardowe założenia, że efekty losowe mają rozkłady normalne i są wzajemnie niezależne z jednakową wariancją, σ^2 , w każdej grupie tych efektów (McIntosh, 1983; Smith i in., 2005; Hu i Spilke, 2011). Zatem, przyjmujemy, że

$$Y_i \sim NID(0, \sigma_Y^2), YL_{ij} \sim NID(0, \sigma_{YL}^2), GY_{ik} \sim NID(0, \sigma_{GY}^2), MY_{il} \sim NID(0, \sigma_{MY}^2), \\ GLY_{ijk} \sim NID(0, \sigma_{GLY}^2), GMY_{jkl} \sim NID(0, \sigma_{GMY}^2), MLY_{ijl} \sim NID(0, \sigma_{MLY}^2) \text{ oraz} \\ GMLY_{ijkl} \sim NID(0, \sigma_{GMLY}^2), e_{1ijk} \sim NID(0, \sigma_{e_1}^2), e_{2ijl} \sim NID(0, \sigma_{e_2}^2), \\ e_{3ijkl} \sim NID(0, \sigma_{e_3}^2) \text{ — (de la Vega i in., 2007 b; Peltonen-Sainio i in., 2009; DeLacy i} \\ \text{in., 2010; Welham i in., 2010).$$

Mieszany model łącznej analizy wariancji (1) dla danych kompletnych w klasyfikacji $G \times M \times L \times Y$ jest podstawą estymatorów typu BLUE dla efektów stałych, funkcji liniowych tych estymatorów (średnich obiektowych oraz różnic tych średnich) i ich błędów standardowych, a także łącznej czteroczynnikowej analizy wariancji.

Tabela łącznej analizy wariancji

Na podstawie modelu mieszanego ANOVA (1), można wysunąć wiele hipotez statystycznych, traktujących o zerowych wartościach poszczególnych rodzajów, merytorycznie ważnych efektów czynnikowych. Każda z tych hipotez jest przypuszczeniem badacza o braku odpowiedniego efektu: jednego, dwóch, trzech lub czterech

czynników. Merytorycznie, najważniejsze są hipotezy dotyczące zerowych wartości efektów głównych, interakcji podwójnej i potrójnej dla odmian, intensywności agrotechniki i miejscowości.

Tabela 2

Łączna analiza wariancji oparta na modelu mieszanym dla średnich poprawionych w klasyfikacji $G \times M \times L \times Y$, obliczonych z danych z serii doświadczeń dwuczynnikowych PDOiR, planowanych w układzie split-block z n powtórzeniami

Overall analysis of variance based on a mixed model for corrected means in classification $G \times M \times L \times Y$, that were calculated using data from a series of two-factorial experiments of PVTS planned in a split-block setup with n replications

| Źródła zmienności Sources of variation | Stopnie swobody Degrees of freedom | Średnie kwadraty [†] Mean squares [†] | F _{emp} |
|--|---------------------------------------|--|------------------|
| Rok — year | I-1 | M1 | M1/M8 |
| Miejscowość — location | J-1 | M2 | M2/M3 |
| Rok × Miejscowość — year × location | (I-1)(J-1) | M3 | M3/M8 |
| Odmiana — cultivar | K-1 | M4 | M4/M5 |
| Odmiana × Rok — cultivar × year | (I-1)(K-1) | M5 | M5/M8 |
| Odmiana × Miejscowość — cultivar × location | (J-1)(K-1) | M6 | M6/M7 |
| Odmiana × Miejscowość × Rok | (I-1)(J-1)(K-1) | M7 | M7/M8 |
| Cultivar × location × year | (I-1)(J-1)(K-1) | M7 | M7/M8 |
| Łączny średni błąd 1 Overall mean error 1 | $\sum_{ij} v_{e1(ij)}$ | M8 = S_{e1}^2 | |
| Agrotechnika — Agrotechnical procedures | $\Lambda-1$ | M9 | M9/M10 |
| Agrotechnika × Rok | (I-1)($\Lambda-1$) | M10 | M10/M13 |
| Agrotechnical procedures × year | (I-1)($\Lambda-1$) | M10 | M10/M13 |
| Agrotechnika × Miejscowość | (J-1)($\Lambda-1$) | M11 | M11/M12 |
| Agrotechnical procedures × location | (J-1)($\Lambda-1$) | M11 | M11/M12 |
| Agrotechnika × Miejscowość × Rok | (I-1)(J-1)($\Lambda-1$) | M12 | M12/M13 |
| Agrotechnical procedures × location × year | (I-1)(J-1)($\Lambda-1$) | M12 | M12/M13 |
| Łączny średni błąd 2 Overall mean error 2 | $\sum_{ij} v_{e2(ij)}$ | M13 = S_{e2}^2 | |
| Rok — year | (K-1)($\Lambda-1$) | M14 | M14/M15 |
| Miejscowość — location | (I-1)(K-1)($\Lambda-1$) | M15 | M15/M18 |
| Rok × Miejscowość — year × location | (J-1)(K-1)($\Lambda-1$) | M16 | M16/M17 |
| Odmiana — cultivar | (I-1)(J-1)(K-1)($\Lambda-1$) | M17 | M17/M18 |
| Łączny średni błąd 3 Overall mean error 3 | $\sum_{ij} v_{e3(ij)}$ | M18 = S_{e3}^2 | |

[†] Średnie kwadraty dla wszystkich efektów czynnikowych oblicza się według klasycznych wzorów dla 4-kierunkowej analizy wariancji danych (tutaj średnich poprawionych) w klasyfikacji kompletnej

[†] Mean squares for all factorial effects were calculated according to the classical method for 4-factorial analysis of variance in data (here: corrected means) in complete classification

Hipotezy o zerowych wartościach efektów głównych badanych czynników wskazują, że nie wpływały one przeciętnie (poprzez poziomy wszystkich pozostałych, badanych czynników) na daną cechę rolniczą rozważanego gatunku roślin. Natomiast, hipotezy o zerowych wartościach efektów interakcji podwójnej i potrójnej, odpowiednio dwóch lub trzech czynników wskazują, że każdy z nich może wpływać w zróżnicowany sposób na badaną cechę, zależnie od poziomów drugiego i trzeciego czynnika. Omawiane hipotezy mogą być testowane za pomocą testu F w łącznej analizie wariancji na podstawie modelu

mieszanego (1) dla średnich w kompletnej klasyfikacji $G \times M \times L \times Y$ z serii doświadczeń PDOiR. Łączna analiza wariancji oparta na modelu (1) została przedstawiona w tabeli 2. Jest ona odpowiednio zmodyfikowaną postacią podobnej analizy wariancji dla wielokrotnej i wieloletniej serii dwuczynnikowych doświadczeń planowanych w układzie split-plot, podaną przez McIntosh (1983). Ta analiza wariancji została odpowiednio dostosowana do układu split-block tych doświadczeń, zgodnie z teorią statystyczną modelowania analizy danych w takim układzie doświadczalnym (Elandt, 1964; Mejza, 1999; Mintenko i in., 2002).

Testowanie istotności efektów głównych odmian (G), intensywności agrotechniki (M), miejscowości (L) i lat (Y), a także interakcji podwójnych, potrójnych i poczwórnej między tymi czynnikami powinno być odpowiednio dostosowane do układu doświadczalnego, w jakim założono doświadczenia oraz do charakteru statystycznego czynników, tzn. traktowania ich jako stałe lub losowe (Elandt, 1964; Annicchiarico 2002 b; Mintenko i in., 2002; Annicchiarico i in., 2010 a). Testowanie hipotez o braku rozważanych efektów czynnikowych za pomocą testu F, zostało przedstawione w tabeli 2. Istotność efektów głównych dla czynników stałych, tj. odmian, agrotechniki i miejscowości oraz ich efektów interakcji jest testowana za pomocą testu F, w którym statystyki F_{emp} są określone jako ilorazy średnich kwadratów dla każdego z tych efektów oraz średnich kwadratów dla interakcji testowanych efektów z latami, jako czynnikiem losowym (McIntosh, 1983; Elandt, 1964; Annicchiarico, 2002 b; Annicchiarico i in., 2010 a). Do testowania istotności efektów głównych lat i efektów interakcji czynników stałych z latami wykorzystuje się funkcje testowe F_{emp} , które są ilorazami średnich kwadratów dla tych efektów i średnich kwadratów dla odpowiadających im błędów — 1, 2 lub 3 (McIntosh, 1983). W omawianych seriach doświadczeń odmianowo-agrotechnicznych uznanie odmiany, agrotechniki i miejscowości za czynniki stałe, ma swoją konsekwencję metodyczno-statystyczną. Polega ona na odnoszeniu wniosków o plonowaniu badanego gatunku roślin tylko do tych odmian, intensywności agrotechniki i miejscowości, które uczestniczyły w serii doświadczeń PDOiR.

Szczegółowa analiza porównawcza średnich

W celu przeprowadzenia wielostronnej analizy statystycznej 4-kieunkowej klasyfikacji danych w postaci Odmiana \times Agrotechnika \times Miejscowość \times Rok, łączna analiza wariancji powinna być uzupełniona o szczegółowe porównanie różnych średnich obiektowych, stanowiących wartości liniowych funkcji estymatorów BLUE dla efektów stałych w modelu (1), tj. średnich dla poziomów niektórych badanych czynników stałych oraz średnich dla kombinacji dwuczynnikowych). Takie postępowanie można przeprowadzić za pomocą odpowiednich metod wielokrotnych porównań średnich. Umożliwia ono szczegółową ocenę:

1. przeciętnego wpływu każdego z badanych czynników (głównie odmian i intensywności agrotechniki) na badaną cechę roślin, za pomocą porównania średnich dla poziomów każdego z tych czynników,
2. zróżnicowania średnich odmianowych (obliczonych poprzez miejscowości i lata) przy każdej intensywności agrotechniki, za pomocą porównania odpowiednich średnich dla odmian przy każdej agrotechnice,

3. reakcji odmian pod względem badanej cechy na intensywność agrotechniki, średnio poprzez miejscowości i lata, za pomocą porównania odpowiednich średnich dla badanych intensywności agrotechniki w odniesieniu do każdej z odmian,
4. reakcji odmian na intensywność agrotechniki, średnio poprzez lata, oddzielnie w różnych grupach środowisk rolniczych, za pomocą porównania odpowiednich średnich dla agrotechniki w odniesieniu do każdej z odmian w grupach miejscowości,
5. przeciętnej reakcji danej cechy roślin rozpatrywanego gatunku (średniej reakcji poprzez odmiany i lata) na intensywność agrotechniki w każdym badanym środowisku rolniczym, za pomocą porównania odpowiednich średnich dla intensywności agrotechniki w każdej z badanych miejscowości.

Istnieje wiele różnych metod wielokrotnych porównań średnich, opartych na różnych procedurach jednoczesnego testowania istotności różnic między średnimi. Często stosowane są metody Tukeya oraz Newman-Keulsa (N-K), stanowiące rozwinięcie zwykłego testu t-Studenta do porównań średnich dla dwóch obiektów (Mądry, 2003; Quinn i Keough, 2003; Steinberg, 2011). Dla obu tych metod najmniejsza istotna różnica (NIR) jest obliczana według tej samej ogólnej zasady, jako iloczyn oceny błędu standardowego różnicy średnich (s_r) i wartości krytycznej, $t_{\alpha, \nu}$, odpowiedniego rozkładu teoretycznego dla stopni swobody, ν , błędu doświadczalnego oraz przy poziomie istotności α , zapewniającym wymagany jednoczesny poziom istotności porównań wielokrotnych średnich, nie mniejszy, niż 0,05. Gdy liczba porównywanych średnich obiektowych wynosi 2, wówczas stosuje się zwykły test t-Studenta (Elandt, 1964; Steinberg, 2011). Jego wykorzystanie w klasyfikacji danych w postaci $G \times M \times L \times Y$ może być uzasadnione przy porównaniu odpowiednich średnich tylko dla dwóch badanych intensywności agrotechniki. Taki przypadek jest typowy dla doświadczeń PDOiR ze zbożami.

Najmniejsza istotna różnica (NIR) do porównania różnych rodzajów średnich dla dwóch badanych intensywności agrotechniki, oparta na teście t-Studenta, ma ogólną postać:

$$NIR_{\alpha} = t_{\alpha, \nu} s_r \quad (2)$$

gdzie:

$t_{\alpha, \nu}$ jest wartością krytyczną rozkładu t-Studenta z ν stopniami swobody dla odpowiedniej interakcji z latami przy poziomie istotności α ,

$s_r = \sqrt{\frac{2M \bullet}{N}}$ jest standardowym błędem różnicy dwóch średnich obliczonych z N obserwacji (średnich poprawionych w klasyfikacji $G \times M \times L \times Y$), $M \bullet$ jest średnim kwadratem odpowiedniej interakcji porównywanych efektów z rokiem, jako czynnikiem losowym, określonym w łącznej analizie wariancji (tab. 2).

Wobec podanych zasad ogólnych, w naszym przypadku NIR_{α} , odpowiednia do porównania średnich dla dwóch intensywności agrotechniki u każdej z badanych odmian, obliczonych poprzez I lat oraz J miejscowości, \bar{x}_{kl} , ma postać:

$$NIR_{\alpha, (\bar{x}_{kl=1} - \bar{x}_{kl=2})} = t_{\alpha, v=(I-1)(K-1)(\Lambda-1)} \sqrt{\frac{2M15}{IJ}} \quad (3)$$

gdzie: $t_{\alpha, v=(I-1)(K-1)(\Lambda-1)}$ jest wartością krytyczną rozkładu t-Studenta przy poziomie istotności α , dla liczby stopni swobody, v , średniego kwadratu ($M15$), dotyczącego interakcji odmiana \times intensywność agrotechniki \times rok w łącznej analizie wariancji.

Natomiast NIR_{α} , odpowiednia do porównania średnich dla dwóch intensywności agrotechniki u każdej z badanych odmian, obliczonych poprzez I lat oraz grupę $Z < J$ miejscowości, \bar{x}'_{kl} , ma postać:

$$NIR_{\alpha, (\bar{x}'_{kl=1} - \bar{x}'_{kl=2})} = t_{\alpha, v=(I-1)(K-1)(\Lambda-1)} \sqrt{\frac{2M15}{IZ}} \quad (4)$$

(użyte symbole, oprócz Z , mają znaczenie takie, jak w formule (3)).

Natomiast, NIR_{α} , odpowiednia do porównania średnich dla dwóch intensywności agrotechniki w każdej z badanych miejscowości, obliczonych poprzez I lat oraz K odmian, \bar{x}_{jl} , ma postać:

$$NIR_{\alpha, (\bar{x}_{jl=1} - \bar{x}_{jl=2})} = t_{\alpha, v=(I-1)(J-1)(\Lambda-1)} \sqrt{\frac{2M12}{IK}} \quad (5)$$

gdzie: $t_{\alpha, v=(I-1)(J-1)(\Lambda-1)}$ jest wartością krytyczną rozkładu t-Studenta przy poziomie istotności α , dla liczby stopni swobody, v , średniego kwadratu ($M12$), dotyczącego interakcji intensywność agrotechniki \times miejscowość \times rok w łącznej analizie wariancji.

W celu dokonania wielokrotnych porównań średnich dla liczby obiektów większej, niż 2, można zastosować procedurę Newman-Keulsa (Wójcik i Laudański, 1989; Mądry 2003; Steinberg, 2011). W rozpatrywanych badaniach jesteśmy zainteresowani wielokrotnym porównaniem średnich odmianowych, \bar{x}_k , obliczonych poprzez I lat, J miejscowości oraz Λ (w ogólnym przypadku) intensywności agrotechniki. Najmniejsze istotne różnice w procedurze Newman-Keulsa do porównania wymienionych średnich, mają postać:

$$NIR_{\alpha, (\bar{x}_k - \bar{x}_k')} = q_{\alpha, s, v=(I-1)(K-1)} \sqrt{\frac{2M5}{IJA}} \quad (6)$$

gdzie: $q_{\alpha, s, v=(I-1)(K-1)}$ jest wartością krytyczną studentyzowanego rozstępu przy poziomie istotności α , dla liczby stopni swobody, v , średniego kwadratu ($M5$), dotyczącego interakcji odmiany \times lata w łącznej analizie wariancji oraz s porównywanych odmian, gdzie $s < K$.

Także, jesteśmy zainteresowani wielokrotnym porównaniem średnich odmianowych, \bar{x}_{kl} , oddzielnie dla każdej intensywności agrotechniki, obliczonych poprzez I lat oraz J miejscowości. Wobec tego, w celu porównania wymienionych średnich odmianowych oddzielnie dla dwóch intensywności agrotechniki, najmniejsze istotne różnice w procedurze Newman-Keulsa mają postać:

$$NIR_{\alpha, (\bar{x}_{ki} - \bar{x}_{k1})}^{N-K} = q_{\alpha, s, v=(I-1)(K-1)} \sqrt{\frac{2M5}{IJ}} \quad (7)$$

gdzie:

$q_{\alpha, s, v=(I-1)(K-1)}$ oraz $M5$ są wartościami odpowiednich parametrów podanych we wzorze (6).

BADANIE REAKCJI ODMIAN NA ŚRODOWISKO ZA POMOCĄ ANALIZY AMMI

Model AMMI dla efektów interakcji Odmiana × Miejscowość

Ważnym celem wnioskowania w doświadczeniach odmianowych i odmianowo-agrotechnicznych jest ocena adaptacyjnej reakcji odmian pod względem cech rolniczych na warunki środowiskowe. W tym badaniu polecanym podejściem jest zastosowanie analizy AMMI oraz wykorzystanie i interpretacja jej wyników, wraz z wynikami ANOVA, w celu efektywnej wizualizacji wymienionej reakcji odmian w postaci wykresu plonu nominalnego odmian (Annicchiarico, 2002 b; Gauch i in., 2008; Annicchiarico i in., 2010 a; De Vita i in., 2010; Mądry i in., 2012).

Analiza AMMI stałych efektów interakcji GL_{jk} w liniowym modelu mieszanym ANOVA (1) jest oparta na matematycznym ujęciu tych efektów w postaci multiplikatywnej. Ujęcie to polega na podziale efektów interakcji GL_{jk} na składniki, będące iloczynami odpowiednich parametrów multiplikatywnych dla k -tej odmiany oraz j -tej miejscowości (Annicchiarico, 2002 b; Crossa i in., 2002; Gauch i in., 2008), zgodnie ze stałym modelem AMMI dla efektów interakcji Odmiana × Miejscowość, który ma postać (Annicchiarico, 2009; Annicchiarico i in., 2010 a; De Vita i in., 2010):

$$GL_{jk} = \sum_{t=1}^T \lambda_t u_{tk} v_{tj} \quad (8)$$

gdzie:

GL_{jk} jest stałym efektem interakcji dla k -tej ($k=1,2, \dots, K$) odmiany z j -tą ($j=1,2, \dots, J$) miejscowością, u_{tk} jest odmianowym parametrem multiplikatywnym t -tej ($t=1, \dots, T=\min\{(J-1), (K-1)\}$) interakcyjnej składowej głównej dla k -tej odmiany (ang. genotypic score of the t -th interaction principal component, GIPC t) lub odmianowym parametrem t -tego składnika multiplikatywnego efektów interakcji GL_{jk} , v_{tj} jest środowiskowym parametrem multiplikatywnym t -tej interakcyjnej składowej głównej dla j -tej miejscowości (ang. environmental score of the t -th interaction principal component, EIPC t) lub środowiskowym parametrem t -tego składnika multiplikatywnego efektów interakcji GL_{jk} ,

λ_t — jest parametrem skalującym dla t -tego składnika multiplikatywnego.

Można przyjąć, że parametry skalujące, λ_t , są określone w jednostkach miary badanej zmiennej zależnej (X), natomiast parametry odmianowe u_{tk} i środowiskowe v_{tj} są niemianowane. Zatem, każdy t -ty składnik multiplikatywny efektów interakcji GL_{jk} jest wyrażony w jednostkach miary zmiennej zależnej. Nie ma przeszkód, aby parametry

skalujące λ_t były włączone jednocześnie do parametrów odmianowych, u_{tk} i środowiskowych, v_{tj} . Taki zabieg polega na utworzeniu symetrycznie skalowanych parametrów o następującej postaci: $u'_{tk} = u_{tk}\sqrt{\lambda_t}$ oraz $v'_{tj} = v_{tj}\sqrt{\lambda_t}$ (Annicchiarico 2002 b; Ebdon i Gauch, 2002; Gauch i in., 2008; Paderewski, 2008).

Zatem, model (8) przyjmuje postać:

$$GL_{jk} = \sum_{t=1}^T u'_{tk} v'_{tj} \quad (9)$$

Model mieszany łącznej analizy wariancji (1), z wykorzystaniem multiplikatywnej postaci stałych efektów interakcji GL_{jk} (wzór 8), ma następującą postać:

$$X_{ijkl} = m + Y_i + L_j + YL_{ij} + G_k + GY_{ik} + \sum_{t=1}^T u'_{tk} v'_{tj} + GLY_{ijk} + e_{1ijk} + M_l + MY_{il} + ML_{il} + MLY_{ijl} + e_{2ijl} + GM_{kl} + GMY_{ikl} + GML_{ijk} + GMLY_{ijkl} + e_{3ijkl} \quad (10)$$

Równanie (10) stanowi model addytywno-multiplikatywny (ang. linear-bilinear model) — (Crossa i Cornelius, 2002; Crossa i in., 2002, 2010), który jest podstawą dwóch metod statystycznych, tj. analizy wariancji oraz analizy AMMI. Addytywno-multiplikatywny model (10) może być wykorzystany do konstrukcji modelu i wykresu plonu nominalnego odmian, jako narzędzi do analizy adaptacyjnej reakcji odmian na środowisko, przeciętnie poprzez poziomy intensywności agrotechniki oraz lata, na podstawie danych w klasyfikacji kompletnej o postaci Odmiana \times Agrotechnika \times Miejscowość \times Rok (Ma i in., 2004; Anicchiarico i in., 2010 a; Mohammadi i in., 2011). Jest to jedna z możliwości wnioskowania o ocenie reakcji odmian na środowiska, uzasadniona wtedy, gdy interakcje odmian i miejscowości z agrotechniką, tj. $G \times M$, $L \times M$ oraz $G \times M \times L$, decydujące o zróżnicowanych rodzajach reakcji odmian na środowisko, są nieistotne. Nieistotne interakcje $G \times M$, $L \times M$ oraz $G \times M \times L$ dla plonu ziarna pszenicy zwyczajnej i twardej uzyskali w swoich badaniach Ma i in. (2004), Annicchiarico i in. (2010 a) oraz Mohammadi i in. (2011). Jednakże, częstym i potwierdzonym empirycznie zjawiskiem są także istotne interakcje $G \times M$, $L \times M$ oraz $G \times M \times L$ (Brancourt-Hulmel, 2003; Stiller i in., 2004).

Kiedy wszystkie wymienione interakcje lub co najmniej jedna z nich są istotne, wówczas odpowiednim podejściem do badania reakcji odmian na środowiska jest zastosowanie dwóch metod, tj. analizy wariancji i analizy AMMI, podobnie jak te, oparte na modelu (10), jednakże oddzielnie dla każdej badanej intensywności agrotechniki. Zatem, model łącznej analizy wariancji, z wykorzystaniem multiplikatywnej postaci efektów interakcji $G \times L$, czyli model addytywno-multiplikatywny dla danych w trójczynnikowej klasyfikacji o postaci Odmiana \times Miejscowość \times Rok, oddzielnie dla każdej z badanych intensywności agrotechniki, ma następującą postać:

$$X_{ijk(l)} = m_l + Y_{i(l)} + L_{j(l)} + YL_{ij(l)} + G_{k(l)} + GY_{ik(l)} + \sum_{t=1}^T u'_{tk(l)} v'_{tj(l)} + GLY_{ijk(l)} + e_{1ijk(l)} \quad (11)$$

gdzie: $X_{ijk(l)}$ jest średnią poprawioną obserwacji plonu z n powtórzeń dla 3-czynnikowej kombinacji i -tego ($i=1,2, \dots, I$) roku, j -tej ($j=1,2, \dots, J$) miejscowości oraz k -tej ($k=1,2, \dots, K$) odmiany w obrębie l -tej ($l=1,2, \dots, L$) intensywności agrotechniki, m_l jest średnią ogólną w obrębie l -tej intensywności agrotechniki, $Y_{i(l)}$ jest losowym efektem głównym i -tego roku w obrębie l -tej intensywności agrotechniki, $L_{j(l)}$ jest stałym efektem głównym j -tej miejscowości w obrębie l -tej intensywności agrotechniki, $G_{k(l)}$ jest stałym efektem głównym k -tej odmiany w obrębie l -tej intensywności agrotechniki, $YL_{i(l)j}$ jest losowym efektem interakcji podwójnej i -tego roku z j -tą miejscowością w obrębie l -tej intensywności agrotechniki, $GY_{ik(l)}$ jest losowym efektem interakcji podwójnej k -tej odmiany z i -tym rokiem w obrębie l -tej intensywności agrotechniki,

$GL_{jk(l)} = \sum_{t=1}^T u'_{ik(l)} v'_{ij(l)}$ jest modelem AMMI dla efektów interakcji podwójnej k -tej odmiany

z j -tą miejscowością w obrębie l -tej intensywności agrotechniki, w którym $u'_{tk(l)}$ oraz $v'_{tj(l)}$ są odmianowymi i środowiskowymi parametrami t -tego składnika multiplikatywnego efektów interakcji $G \times L_{(l)}$ w obrębie l -tej intensywności agrotechniki,

$GLY_{ijk(l)}$ jest losowym efektem interakcji potrójnej k -tej odmiany z j -tą miejscowością oraz i -tym rokiem w obrębie l -tej intensywności agrotechniki,

$e_{1ijk(l)}$ jest średnim błędem 1 w układzie split-block, związanym z obserwacjami $X_{ijk(l)}$ w obrębie l -tej intensywności agrotechniki.

Uproszczenie interpretacji efektów interakcji GL_{jk} w modelu (10) oraz $GL_{jk(l)}$ w modelu (11), opartej na ich modelowaniu za pomocą parametrów multiplikatywnych, jest uzasadnione i skuteczne tylko wtedy, gdy efekty te mogą być wystarczająco dobrze przybliżone przez kilka (s) pierwszych istotnych, zwykle jedną lub dwie, spośród T niezerowych interakcyjnych składowych głównych (ang. interaction principal components, IPC) — (van Eeuwijk, 1995; Gauch i Zobel, 1997; Yan i Kang, 2003). Ograniczając się w interpretacji efektów interakcji Odmiana \times Miejscowość za pomocą analizy AMMI, do istotnych interakcyjnych składowych głównych, pozbywamy się tzw. szumu losowego w danych (ang. noise), czyli zakłócających efektów losowych, uwikłanych z pozostałymi, nieistotnymi interakcyjnymi składowymi głównymi (Gauch, 1992, 2006; Gauch i Zobel, 1996, 1997; Annicchiarico, 2002 b, 2009). Wobec tego, dużym walorem analizy AMMI, potwierdzonym w licznych badaniach teoretycznych i empirycznych (Gauch, 2006; Annicchiarico, 2002 a, b, Ebdon i Gauch, 2002), jest poprawa precyzji wnioskowania o reakcji odmian w środowiskach na podstawie poprawionych (odtworzonych) średnich plonów odmian w każdej badanej miejscowości z wykorzystaniem istotnych interakcyjnych składowych głównych (ang. AMMI-adjusted means lub AMMI-modeled means), w porównaniu ze zwykłymi średnimi plonu dla kombinacji Odmiana \times Miejscowość.

Koncepcja wykresu plonu nominalnego

Analiza AMMI dla efektów interakcji GL_{jk} oraz $GL_{jk(l)}$ może posłużyć w celu wykorzystania jej wyników do wizualizacji adaptacyjnej reakcji badanych odmian względem plonu na środowisko w agro-ekosystemach, za pomocą wykresu plonu nominalnego dla tych odmian. Taki wykres odzwierciedla graficznie model plonu

nominalnego odmian. Model plonu nominalnego odmian można rozpatrywać zarówno dla średnich odmianowych poprzez intensywność agrotechniki i lata (opierając się na modelu (10)), jak i dla średnich odmianowych poprzez lata, oddzielnie przy każdej intensywności agrotechniki (opierając się na modelu (11)). Do określenia wspomnianych modeli i wykresów plonu nominalnego odmian, posługujemy się wynikami zarówno łącznej analizy wariancji (tab. 2), opartej na modelu (10) lub (11), w celu oceny i testowania istotności (testem F) efektów głównych odmian, jak i analizy AMMI. Jak wiadomo, analiza AMMI efektów interakcji GxL pozwala na estymację skalowanych symetrycznie odmianowych oraz środowiskowych parametrów multiplikatywnych dla pierwszej interakcyjnej składowej głównej, czyli parametrów \hat{u}'_{1k} i \hat{v}'_{1j} oraz $\hat{u}'_{1k(l)}$ i $\hat{v}'_{1j(l)}$ (Gauch i Zobel, 1997; Annicchiarico, 2002 b, 2009; Pinnschmidt i Hovmøller, 2002; Annicchiarico i in., 2010 a, b).

Model plonu nominalnego odmiany jest liniową funkcją regresji jej plonu nominalnego, N_{ij} , czyli odchylen średniej plonu tej odmiany w miejscowościach od oceny głównego efektu środowiskowego, względem ocen środowiskowych multiplikatywnych parametrów pierwszej interakcyjnej składowej głównej, EIPC1 (Gauch i Zobel, 1997; Ebdon i Gauch, 2002; Annicchiarico i in., 2006). Model ten jest rezultatem pewnych przekształceń odpowiedniego modelu addytywno-multiplikatywnego. W naszych badaniach model plonu nominalnego odmiany jest wynikiem przekształceń modelu (10) lub (11). Dla modelu (10) model plonu nominalnego ma postać (Gauch i Zobel, 1997, Annicchiarico, 2002 b, 2009, Samonte i in., 2005):

$$N_{jk} = \hat{m} + \hat{G}_k + (\hat{u}'_{1k} \hat{v}'_{1j}) \quad (12)$$

gdzie:

N_{jk} jest plonem nominalnym dla k -tej odmiany oraz j -tej miejscowości, \hat{m} jest oceną średniej ogólnej, \hat{G}_k jest oceną stałego efektu głównego k -tej odmiany, u'_{1k} oraz v'_{1j} są symetrycznie skalowanymi, ocenami odpowiednio odmianowych (dla k -tej odmiany) i środowiskowych (dla j -tej miejscowości) parametrów multiplikatywnych dla pierwszej interakcyjnej składowej głównej.

Addytywno-multiplikatywny model (10) może być wykorzystany do konstrukcji modelu i wykresu plonu nominalnego odmian, jako narzędzie do analizy adaptacyjnej reakcji odmian na środowisko, przeciętnie poprzez poziomy intensywności agrotechniki oraz lata, na podstawie danych w klasyfikacji kompletnej o postaci Odmiana \times Agrotechnika \times Miejscowość \times Rok (Ma i in., 2004; Annicchiarico i in., 2010 a, Mohammadi i in., 2011). Natomiast, model (11) łącznej analizy wariancji i analizy AMMI dla efektów interakcji $G \times L_{jk(l)}$ może być wykorzystany do konstrukcji modelu i wykresu plonu nominalnego odmian, oddzielnie dla każdej intensywności agrotechniki. Takie rozwiązanie metodyczne jest uzasadnione wówczas, kiedy zostanie potwierdzona przynajmniej jedna istotna interakcja pomiędzy czynnikami $G \times M$, $L \times M$ lub $G \times M \times L$. Zatem, model plonu nominalnego oddzielnie dla każdej z intensywności agrotechniki ma następującą postać:

$$N_{jk(l)} = \hat{m}_{(l)} + \hat{G}_{k(l)} + (\hat{u}'_{1k(l)} \hat{v}'_{1j(l)}) \quad (13)$$

gdzie:

$N_{jk(l)}$ jest plonem nominalnym dla k -tej odmiany oraz j -tej miejscowości w obrębie l -tej intensywności agrotechniki, $\hat{m}_{(l)}$ jest oceną średniej ogólnej w obrębie l -tej intensywności agrotechniki, $\hat{G}_{k(l)}$ jest oceną stałego efektu głównego k -tej odmiany w obrębie l -tej intensywności agrotechniki, $\hat{u}'_{1k(l)}$ oraz $\hat{v}'_{1j(l)}$ są symetrycznie skalowanymi, ocenami odpowiednio odmianowych (dla k -tej odmiany) i środowiskowych (dla j -tej miejscowości) parametrów multiplikatywnych dla pierwszej interakcyjnej składowej głównej w obrębie l -tej intensywności agrotechniki.

Wykres plonu nominalnego jest oparty jedynie na pierwszej interakcyjnej składowej głównej (IPC1), niezależnie od tego czy pozostałe IPC są istotne (Annicchiarico i in., 2006; Annicchiarico i Iannucci, 2008; Gauch i in., 2008). Jest jasne, że wiarygodność i precyzja wizualizacji adaptacyjnej reakcji odmian w środowiskach rolniczych za pomocą wykresu plonu nominalnego odmian zależy od procentu wyjaśnionej zmienności efektów interakcji $G \times L$ przez IPC1. W badaniach empirycznych autorzy stosują rozpatrywany rodzaj wykresów przy różnym procencie wyjaśnionej zmienności efektów interakcji $G \times L$ przez pierwszą interakcyjną składową główną. W pionierskich badaniach Guacha i Zobela (1997), dotyczących koncepcji i wykorzystania wykresu plonu nominalnego w celu graficznego zobrazowania adaptacyjnej reakcji odmian na warunki środowiskowe, pierwsza interakcyjna składowa główna wyjaśniała 87% całkowitej sumy kwadratów interakcji $G \times L$. Mimo, że często druga interakcyjna główna także była istotna, autorzy zdecydowali się na zobrazowanie graficzne adaptacyjnej reakcji odmian na badane środowiska rolnicze za pomocą wykresu plonu nominalnego, opartego jedynie na IPC1 (Pinnschmidt i Hovmøller, 2002; Samonte in., 2005; Annicchiarico i Iannucci, 2008). W cytowanych badaniach IPC1 wyjaśniała od około 40% (Samonte in., 2005), do około 50% całkowitej zmienności efektów interakcji $G \times L$ (Annicchiarico i in., 2011).

Wykres plonu nominalnego powstaje w układzie współrzędnych prostokątnych w taki sposób, że na osi poziomej podane są środowiskowe parametry multiplikatywne dla pierwszej interakcyjnej składowej głównej (EIPC1), zaś na osi pionowej podane są wartości plonów nominalnych (Annicchiarico, 2002 b; Zhang i in., 2006; Mohamadi i Amri, 2013). Liniowość tego wykresu w znaczący sposób ułatwia porównanie adaptacyjnej reakcji odmian w środowiskach rolniczych, głównie poprzez ocenę rankingu plonu nominalnego odmian oraz jego ilościowego zróżnicowania międzyodmianowego w każdym rozpatrywanym agro-ekosystemie. Zauważmy, że różnice plonu nominalnego odmian w danym środowisku są równe różnicom faktycznej średniej odmian. Jest tak dlatego, że w modelu plonu nominalnego odmian [modele (12) i (13)] wyeliminowane są efekty główne środowiska, które nie zmieniają różnic średniego plonu odmian w każdym środowisku (Gauch, 1992; Annicchiarico i in., 2010 a, 2011) Zatem, na podstawie wykresu plonu nominalnego można wskazać odmiany o relatywnie najwyższym plonie w środowiskach, czyli odmiany relatywnie (wśród badanych) najlepiej przystosowane, inaczej o najlepszej adaptacji, do konkretnych warunków lokalnych (ang. top-ranking cultivars) — (Gauch i Zobel, 1997; Annicchiarico i Iannucci, 2008; Pecetti i in., 2011).

Dlatego, taki sposób oceny relatywnej przydatności rolniczej odmian do różnych warunków środowiskowych nazywany jest adaptacyjną reakcją odmian w środowiskach rolniczych (Pecetti i in., 2011; Annicchiarico i in., 2011). W opisywanych badaniach nad adaptacyjnością można określić każdą odmianę, jako odznaczającą się wąską lub szeroką adaptacją, albo relatywnie nie wykazującą adaptacji w badanym zakresie różnorodności środowisk w agro-ekosystemach (Gauch i Zobel, 1997; Annicchiarico, 2002 a). Ponadto, wykres plonu nominalnego mimo iż oparty jest jedynie na pierwszej interakcyjnej składowej głównej, umożliwia wskazanie odmian o stabilnym plonowaniu w obrębie rozpatrywanych środowisk rolniczych (Gauch, 2006; Annicchiarico i in., 2006; De Vita i in., 2010; Mohamadi i Amri, 2013).

GRUPOWANIE ODMIAN O PODOBNEJ REAKCJI NA ŚRODOWISKO ZA POMOCĄ ANALIZY SKUPIEŃ

Jednym ze sposobów podejścia do analizy danych z wielokrotnych serii wieloletnich doświadczeń odmianowych i odmianowo-agrotechnicznych jest podział odmian na grupy jednorodne, w których zmienność efektów interakcji Odmiana \times Miejscowość jest odpowiednio zmniejszona w stosunku do zmienności tych efektów w całej serii doświadczeń (Gauch i Zobel, 1997; de la Vega i Chapman, 2006). Wobec tego, efekty interakcji $G \times L$ dla odmian w każdej takiej grupie jednorodnej są podobne (Gauch i Zobel, 1997; Annicchiarico, 2002 b). Takie podejście może być bardzo przydatne i efektywne w praktyce hodowli i oceny odmian do badania różnych aspektów reakcji odmian na zmienne warunki środowiskowe (Annicchiarico, 2002 b; Annicchiarico i Iannucci, 2008; Paderewski, 2008). Jednakże, wymienione rozwiązanie nie pozwala sprawnie i skutecznie wydzielić jednorodnych grup odmian pod względem rodzaju reakcji na środowisko, albowiem nie rozpatruje się w nim efektów głównych (średnich badanej cechy) dla odmian, które decydują, obok efektów interakcji $G \times L$, o środowiskowej reakcji odmian. Z drugiej strony, wydzielenie grup odmian o podobnej reakcji na środowiska jest koniecznym postępowaniem w metodyce statystycznej oceny odmian, zalecanym i stosowanym przez wielu biometryków i badaczy odmian (Basford i Cooper, 1998; Cooper i in., 1996). Takie podejście znacząco upraszcza i ułatwia oraz usprawnia graficzną prezentację różnych rodzajów reakcji odmian na środowiska, zwłaszcza w przypadku, kiedy liczba badanych odmian jest dość duża (Sivapalan i in., 2000; Cooper i in., 2001; Zhang i in., 2006; Mądry i in., 2011, 2012).

Kierując się tą dobrą zasadą metodologiczną, a także mając do czynienia z dużą liczbą badanych odmian w PDOiR z pszenicą ozimą oraz z innymi gatunkami roślin, można stosować podział odmian na grupy jednorodne pod względem plonu nominalnego odmian, który określa rodzaj adaptacyjnej reakcji badanych odmian w środowiskach rolniczych. Dlatego, każda taka grupa jednorodna będzie zawierała odmiany o podobnej adaptacyjnej reakcji na środowiska. Do wydzielenia wspomnianych grup jednorodnych odmian stosuje się różne metody statystyczne. Jedną z częściej używanych w podobnych badaniach empirycznych, jest analiza skupień z wykorzystaniem hierarchicznej metody aglomeracji Warda, z odległością Euklidesa lub kwadratem odległości Euklidesa (Cooper i in., 1996;

Zhang i in., 2006; Paderewski i in., 2011). Ta metoda służąca do oszacowania odległości między skupieniami (grupami) oparta jest na podejściu analizy wariancji, zmierzając do minimalizacji sumy kwadratów odchyłeń dowolnych dwóch grup jednorodnych, w tym konkretnym przypadku, odmian o podobnej adaptacyjnej reakcji na środowiska. Szczegółowe opisy tej metody znajdują się w pracy Warda (1963), a także między innymi w publikacji Crossa i Franco (2004) i innych opracowaniach monograficznych (Krzanowski, 1988; Johnson i Wichern, 2002).

W każdej analizie skupień rozpatruje się wartości miar niepodobieństwa obiektów pod względem wielu zmiennych jednocześnie. W rozpatrywanych badaniach zmiennymi są plony nominalne danej odmiany w różnych miejscowościach. Zatem, odległość Euklidesa dla każdej pary odmian jest określona jako pierwiastek kwadratowy sumy kwadratów różnic plonów nominalnych tych odmian w poszczególnych miejscowościach. Im ta odległość jest większa, tym większe jest niepodobieństwo odmian pod względem plonów nominalnych w miejscowościach, czyli odmiany odznaczają się bardziej odmienną adaptacyjną reakcją na środowiska. Wobec tego, w grupach jednorodnych powinny znaleźć się odmiany o relatywnie mniejszych wartościach odległości Euklidesa, niż odległości między grupami. Liczba wydzielanych grup obiektów (w tym przykładzie odmian) za pomocą analizy skupień może być różna, jednak powinna ona stanowić kompromis między uzyskaniem małej liczby grup oraz ich dużej zwięzłości (Crossa i Franco, 2004; Paderewski, 2008).

LITERATURA

- Annicchiarico P. 2002 a. Defining adaptation strategies and yield stability targets in breeding programmes In: Kang M. S. (Ed.) *Quantitative genetics, genomics and plant breeding*. CAB, Wallingford, UK: 165 — 183.
- Annicchiarico P. 2002 b. *Genotype-environment interactions: challenges and opportunities for plant breeding and cultivar recommendations*. FAO Plant Production and Protection Paper No. 174. Food and Agriculture Organization, Rome.
- Annicchiarico P. 2009. Coping with and exploiting genotype \times environment interactions. In: Ceccarelli S., Guimarães E. P., Weltzien E. (eds). *Plant Breeding and Farmer Participation*. Food and Agricultural Organization, Rome: 519 — 564.
- Annicchiarico P., Bellah F., Chiari T. 2006. Repeatable genotype \times location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria. *Eur. J. Agron.* 24: 70 — 81.
- Annicchiarico P., Chiapparino E., Perenzin M. 2010 a. Response of common wheat varieties to organic and conventional production systems across Italian locations and implications for selection. *Field Crops Res.* 116: 230 — 238.
- Annicchiarico P., Scotti C., Carelli M., Pecetti L. 2010 b. Questions and avenues for lucerne improvement. *Czech J. Genet. Plant Breed.* 46: 1 — 13.
- Annicchiarico P., Iannucci A. 2008. Adaptation strategy, germplasm type and adaptive traits for field pea improvement in Italy based on variety responses across climatically contrasting environments. *Field Crops Res.* 108: 133 — 142.
- Annicchiarico P., Pecetti L., Abdelguerfi A., Bouizgaren A., Carroni A. M., Hayek T., M'Hammadi Bouzina M., Mezni M. 2011. Adaptation of landrace and variety germplasm and selection strategies for lucerne in the Mediterranean basin. *Field Crops Res.* 120: 283 — 291.

- Basford K. E., Cooper M. 1998. Genotype \times environment interactions and some considerations of their implications for wheat breeding in Australia. *Aust. J. Agric. Res.* 49: 153 — 174.
- Brancourt-Hulmel M., Doussinault G., Lecomte C., Berard P., Le Buaec B., Trottet M. 2003. Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. *Crop Sci.* 43: 37 — 45.
- Bujak H., Tratwa G. 2011. Ocena stabilności plonowania odmian pszenicy ozimej na podstawie doświadczeń porejestrowych w Polsce. *Biul. IHAR* 260/261: 69 — 79.
- Caliński T., Czajka S., Kaczmarek Z., Krajewski P. 1998. SERGEN 4 — Analysis of series of variety trials and plant genetic or breeding experiments. Institute of Plant Genetics, Polish Academy of Sciences, and Department of Mathematical and Statistical Methods, Agricultural University, Poznań, Poland.
- Cooper, M., DeLacy I. H., Basford K. E. 1996. Relationships among analytical methods used to study genotypic adaptation in multienvironment trials. in M. Cooper and G. L. Hammer (ed.) *Plant adaptation and crop improvement*. CABI, Wallingford, UK: 193 — 224.
- Cooper M., Woodruff D. R., Phillips I. G., Basford K. E., Gilmour A. R. 2001. Genotype-by-management interactions for grain yield and grain protein concentration of wheat. *Field Crops Res.* 69: 47 — 67.
- Crossa J., Cornelius P. L. 2002. Linear-bilinear models for the analysis of genotype-environment interaction In: Kang M.S. (Ed.) *Quantitative genetics, genomics and plant breeding*. CAB, Wallingford, UK: 305 — 322.
- Crossa J., Cornelius P. L., Yan W. 2002. Biplots of linear-bilinear models for studying crossover genotype \times environment interaction *Crop Sci.* 42: 619 — 633.
- Crossa J., Franco J. 2004. Statistical methods for classifying genotypes. *Euphytica* 137: 19 — 37.
- Crossa J., Vargas M., Joshi A.K. 2010. Linear, bilinear, and linear-bilinear fixed and mixed models for analyzing genotype \times environment interaction in plant breeding and agronomy. *Can. J. Plant Sci.* 90:561 — 574.
- de la Vega A. J., Chapman S. C. 2006. Defining sunflower selection strategies for a highly heterogeneous target population of environments. *Crop Sci.* 46: 136 — 144.
- De Vita P., Mastrangelo A. M., Matteu L., Mazzucotelli E., Virzi N., Palumbo M., Lo Storto M., Rizza F., Cattivelli L. 2010. Genetic improvement effects on yield stability in durum wheat genotypes grown in Italy. *Field Crops Res.* 119: 68 — 77.
- Derejko A., Mądry W., Gozdowski D., Rozbicki J., Golba J., Piechociński M., Studnicki M. 2011. Wpływ odmian, miejscowości i intensywności uprawy oraz ich interakcji na plon pszenicy ozimej w doświadczeniach PDO. *Biul. IHAR* 259: 131 — 146.
- Ebdon J. S., Gauch H. G. 2002. Additive main effect and multiplicative interactions analysis of national turfgrass performance trials. Interpretation of genotype \times environment interactions. *Crop Sci.* 42: 489 — 496.
- Elandt R. 1964. *Statystyka matematyczna w zastosowaniu do doświadczalnictwa rolniczego*. PWN, Warszawa.
- Gauch H. G. 1992. *Statistical analysis of regional yield trials. AMMI analysis of factorial designs*. Elsevier Science, New York.
- Gauch H. G. 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46: 1488 — 1500.
- Gauch H. G., Piepho H.P., Annicchiarico P. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Sci.* 48: 866 — 889.
- Gauch H. G., Zobel R. W. 1996. AMMI analysis of yield trials. In: M. S. Kang, H. G. Gauch (Ed.) *Genotype by environment interaction*. CRC Press, Boca Raton: 85 — 122.
- Gauch H. G., Zobel R. W. 1997. Identifying mega-environments and targeting genotypes. *Crop Sci.* 37: 311 — 326.
- Johnson R. A., Wichern D. W. 2002. *Applied multivariate statistical analysis*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA.
- Kang M. S. 1993. Simultaneous selection for yield and stability in crop performance trials: Consequences for growers. *Agron. J.* 85: 754 — 57.
- Kang M. S. 1998. Using genotype-by environment interaction for crop cultivar development. *Adv. Agron.* 62: 199 — 253.

- Kang M. S. 2002. Genotype-environment interaction: Progress and prospects In: Kang M.S. (Ed.), Quantitative genetics, genomics and plant breeding, CAB International Wallingford, UK: 221 — 243.
- Krzanowski W. J. 1988. Principles of multivariate analysis: a users' perspective. Oxford University Press, Oxford.
- Ma B. L., Yan W., Dwyer L. M., Fregeau-Reid J., Voldeng H. D., Dion Y., Nass H. 2004. Graphic analysis of genotype, environment, nitrogen fertilizer, and their interactions on spring wheat yield. *Agron. J.* 96: 169 — 180.
- Mądry W. 2003. Analiza statystyczna miar stabilności na podstawie danych w klasyfikacji genotypy × środowiska. Część II Model mieszany Shukli i model regresji łącznej. *Coll. Biom.* 33: 207 — 220.
- Mądry W., Paderewski J., Gozdowski D., Drzazga T. 2011. Adaptive yield response of winter wheat cultivars across environments in Poland using joint AMMI and cluster analyses. *Intern. J. Plant Prod.* 5: 299 — 310.
- Mądry W., Paderewski J., Rozbicki J., Gozdowski D., Golba J., Piechociński M., Studnicki M., Derejko A. 2012. Plonowanie odmian pszenicy ozimej w różnych środowiskach- jednoroczna seria PDOiR. *Biul. IHAR* 263:189 — 204.
- McIntosh M. S. 1983. Analysis of combined experiments. *Agron. J.* 75: 153 — 155.
- Mejza I. 1999. Planowanie serii doświadczeń dwuczynnikowych z rozszczepionymi jednostkami i analiza wyników. *Roczniki Akademii Rolniczej w Poznaniu*, vol. 301.
- Mintenko A. S., Smith S. R., Cattani D. J. 2002. Turfgrass evaluation of native grasses for the Northern Great Plains Region. *Crop Sci.* 42: 2018 — 2024.
- Mohammadi R., Amri A. 2013. Genotype × environment interaction and genetic improvement for yield and yield stability of rainfed durum wheat in Iran. *Euphytica* 192: 227 — 249.
- Mohammadi R., Sadeghzadeh D., Armion M., Amri A. 2011. Evaluation of durum wheat experimental lines under different climate and water regime conditions of Iran. *Crop Pasture Sci.* 62: 137 — 151.
- Quinn G. P., Keough M. J. 2003. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.
- Paderewski J. 2008. Przydatność modelu AMMI do badania reakcji roślin rolniczych na warunki środowiskowe. Praca doktorska, Wydział Rolnictwa i Biologii, SGGW.
- Paderewski J., Gauch H. G., Mądry W., Drzazga T., Rodrigues P. C. 2011. Yield response of winter wheat to agro-ecological conditions using additive main effects and multiplicative interaction and cluster analysis. *Crop Sci.* 51: 969 — 980.
- Pecetti L., Annicchiarico P., Abdelguerfi A., Kallida R., Mefti M., Porqueddu C., Simoes N., Volaire F., Lelievre F. 2011. Response of Mediterranean tall fescue cultivars to contrasting agricultural environments and implications for selection. *J. Agron. Crop Sci.* 197: 12 — 20.
- Pinnschmidt H. O., Hovmöller M. S. 2002. Genotype × environment interactions in the expression of net blotch resistance in spring and winter barley varieties. *Euphytica* 125: 227 — 243.
- Samonte S. O. Pb., Wilson L. T., McClung A. M., Medley C. 2005. Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analyses. *Crop Sci.* 45: 2414 — 2424.
- Sivapalan S., O'Brien L., Ortiz-Ferrera G., Hollamby G.J., Barclay I., Martin P.J. 2000. An adaptation analysis of Australian and CIMMYT/ICARDA wheat germplasm in Australian production environments. *Aust. J. Agric. Res.* 51: 903 — 915.
- Smith A. B., Cullis B.R., Thompson R. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* 143: 449 — 462.
- Steinberg W. J. 2011. *Statistics Alive*. SAGE Publications Inc.: 527 — 529.
- Stiller W. N., Reid P. E., Constable G. A. 2004. Maturity and leaf shape as traits influencing cotton cultivar adaptation to dryland conditions. *Agron. J.* 96: 656 — 664.
- van Eeuwijk F. A., Keizer L. C. E., Bakker J. J. 1995. Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch Maize Variety Trials. *Euphytica* 84: 9 — 22.
- Welham S. J., Gogel B. J., Smith A. B., Thompson R., Cullis B. R. 2010. A comparison of analysis methods for late stage variety evaluation trials. *Aust. N. Z. J. Stat.* 52: 125 — 149.
- Wójcik A. R., Laudański Z. 1989. Planowanie i wnioskowanie statystyczne w doświadczalnictwie. PWN.

- Yan W., M.S. Kang. 2003. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.
- Zhang Y., He Z., Zhang A., van Ginkel M., Pena R. J., Ye G. 2006. Pattern analysis on protein properties of Chinese and CIMMYT spring wheat cultivars sown in China and CIMMYT. *Austr. J. Agric. Res.* 57: 811 — 822.