Listy Biometryczne
Vol. XXIV (1987) No.2
Biometrical Letters

PRZEGLĄD MODELI DOŚWIADCZEŃ POLOWYCH UWZGLĘDNIAJĄCYCH SĄSIEDZTWO POLETEK

Anita Dobek, Hanna Kiełczewska

Akademia Rolnicza, Zakład Metod Matematycznych i Statystycznych,
Wojska Polskiego 28, 60-637 Poznań

Dobek A., Kiełczewska H., 1987. The review of models for field experiments in which neighbouring plots effects are included. Listy Biometryczne XXIV, z.2., Wydawnictwo Naukowe Universytetu im. Adama Mickiewicza, Poznań, (Adam Mickiewicz University Press) pp. 55-62, PL ISSN 0458-0036.

The paper presents a review of models for field trials taking into consideration the neighbourhood of plots. Those models are specially useful in two situations : when competition effects occur and when there is a high variability in field trials with large numbers of genotypes. Using those models some improvement in the process of estimation may be obtained.

1. WSTĘP

Zmienność żyzności gleby w doświadczeniach z dużą liczbą genotypów jest istotną przeszkodą w ocenie odmian w hodowli roślin. Szeroko stosowane w tych wypadkach układy bloków niekompletnych nie zawsze są poprawne z uwagi na to, że żyzność gleby i inne podobne zjawiska zmieniają się na polu zwykle w sposób ciągły, podczas gdy układy blokowe zakładają stały efekt żyzności w całym bloku. Pomocą w rozwiązywaniu tego typu problemów jest "statystyka przestrzenna" (spatial statistics), której szybki rozwój przypada na ostatnich dziesięć lat.

W pierwszej części pracy zawarty jest przegląd metod związanych z usuwaniem efektu żyzności (trendu) z oceny efektu obiektowego. Metody te bazują na założeniu, że efekt środowiskowy poletka jest ściśle związany z efektami jego najbliższych sąsiadów. Sąsiedztwo można rozpatrywać w wielu kierunkach. W przedstawionych metodach uwzględnia się tylko sąsiedztwo w jednym kierunku.

Innym problemem, z którym spotykają się hodowcy roślin we wczesnych

---

Słowa kluczowe: efekt konkurencji, estymacja parametrów obiektowych, sąsiedztwo poletek, żyzność gleby

etapach hodowli jest mała ilość materiału hodowlanego, co w doświadczeniach z replikacjami powoduje występowanie małych poletek, np. w wypadku buraków cukrowych poletek składających się z jednego rzędu roślin. W takich sytuacjach może wystąpić zjawisko konkurencji (competition effect) polegające na tym, że na plon danej odmiany wpływa agresywny sąsiad. Innym przypadkiem konkurencji jest współdziałanie (interference) pomiędzy poletkami w sytuacji, gdy działanie obiektu doświadczalnego nie jest precyzyjnie stosowane na poletko, np. podlewanie czy opryskiwanie. W drugiej części pracy zajmiemy się przedstawieniem metod uwzględniających te zjawiska, co w konsekwencji może doprowadzić do uzyskania dokładniejszych ocen efektów obiektowych.

2. METODY ANALIZY UWZGLĘDNIAJĄCE NIEJEDNORODNOŚĆ ŻYZNOŚCI GLEBY

Analizę doświadczeń z replikacjami, w której przeprowadza się poprawianie plonu z poletka uwzględniając plony poletek sąsiednich, co może doprowadzić do wyeliminowania zmienności związanej z żyznością, wprowadził Papadakis (1937). Wcześniej ten sam problem próbowano rozwiązać poprzez stosowanie poletek kontrolnych (patrz np. Wiancko 1914, Załęski 1927). Papadakis po raz pierwszy zastosował poprawianie bez wprowadzania poletek kontrolnych uwzględniając jedynie plon z poletek sąsiednich w doświadczeniach systematycznych z dużą liczbą obiektów. Jego analiza jest w istocie analizą kowariancji ze względu na plony z sąsiednich poletek, z których wyeliminowano efekty obiektowe. Papadakis wyznacza zatem zmienne $d_i = Y_i - \tau^*_{T_i}$, gdzie Y_i jest obserwacją z i-tego poletka (poletka ponumerowane są $1,\dots,n$), a $\tau^*_{T_i}$ jest dowolną oceną efektu obiektu występującego na i-tym poletku. Zmienną towarzyszącą jest suma $d_{i-1} + d_{i+1}$, $i=2,\dots,n-1$. Dla Y_1 zmienną towarzyszącą jest $2d_2$, a dla Y_n jest nią $2d_{n-1}$.

Model obserwacji zaproponowany przez Papadakisa (1937) można więc zapisać w postaci

$$Y = T\tau - WY + Z \quad , \tag{1}$$

gdzie Y jest n-wymiarowym wektorem obserwacji, τ jest p-wymiarowym wektorem nieznanych parametrów obiektowych, T jest nxp-wymiarową macierzą układu, W jest nxn-wymiarową macierzą tworzącą zmienne towarzyszące, a Z jest n-wymiarowym wektorem błędów losowych z wartością oczekiwaną 0 i wariancją $\sigma^2 I$.

Bartlett (1938) podejmując również zagadnienie poprawiania plonu z poletka, jako zmienną towarzyszącą zastosował średnią $(d_{i-1} + d_{i+1})/2$ oraz w modelu uwzględnił bloki. W 1970r. Papadakis zmodyfikował swoją metodę poprzez wprowadzenie iteracji ze względu na estymowane parametry obiektowe. Z kolei w pracy z 1978r. Bartlett zrewidował teoretycznie metodę Papadakisa zarówno dla jedno jak i dwukierunkowych układów sugerując wprowadzenie współczynnika regresji przy zmiennej towarzyszącej. Estymacja tego współczynnika odbywa się metodą iteracyjną. Bartlett swoją metodę rekomenduje jako pomocnicze narzędzie w poprawianiu dokładności estymatorów obiektowych, w wypadku gdy zawiodły klasyczne układy doświadczalne, szczególnie gdy bloki są duże i nie można oczekiwać wyeliminowania efektu żyzności poprzez eliminację efektu blokowego.

Metoda poprawiania estymatorów parametrów obiektowych zwana metodą NN (nearest neighbour), a opierająca się na tzw. drugich różnicach, została po raz pierwszy opisana w pracy Wilkinsona i in. (1983). W wyniku krytycznych uwag skierowanych pod jej adresem w dyskusji, która odbyła się po przedstawieniu metody, autorzy zmodyfikowali swoją metodę, a jej nową wersję przedstawił Wilkinson (1984).

Metodą NN, o czym pisze między innymi Wilkinson (1984), uzyskuje się takie same estymatory parametrów obiektowych jak iteracyjną metodą Papadakisa analizy kowariancji ze stałym współczynnikiem regresji. Różnice występują oczywiście w koncepcji poprawiania, a także w analizie wariancji i w sposobie wyznaczania błędów standardowych.

Istotą metody NN jest eliminacja tzw. "gładkiego trendu" (smooth trend), czyli ciągłej zmienności w żyzności gleby. Realizuje się to poprzez odjęcie od obserwacji z danego poletka średniej z niepoprawionych plonów z sąsiednich poletek zamiast odejmowania ustalonych średnich blokowych, jak to się dzieje w klasycznym układzie blokowym.

Plan doświadczenia, dla którego Wilkinson i in. (1983) opisują swą metodę, to plan r bloków kompletnych tworzących jeden lub kilka pasów. Istotnym elementem metody NN są dodatkowe poletka umieszczone na skraju każdego pasa, wzdłuż którego następuje poprawianie.

Dla obserwacji z doświadczenia przyjmuje się następujący model liniowy

$$Y = T\tau + X + Z \quad , \qquad (2)$$

gdzie Y, T, τ i Z są jak w modelu (1), a X jest n-wymiarowym wektorem losowych efektów trendu. O efektach trendu zakłada się, iż są "gładkie" w tym sensie, że wariancja ich drugich różnic jest mała w porównaniu z wariancją błędu σ^2.

Jak już wspomniano wcześniej, analizę NN prowadzi się opierając się na tzw. drugich różnicach, co dla pojedynczej obserwacji oznacza przekształcenie

$$Y_i' = Y_i - \frac{b}{2}(Y_{i-1} + Y_{i+1}) \quad ,$$

natomiast w postaci macierzowej zapisuje się jako

$$Y' = (I - \frac{b}{2}A)Y \quad ,$$

gdzie A jest macierzą wybierającą odpowiednich sąsiadów.

Wilkinson (1984) proponuje następujący sposób wyznaczania współczynnika b

$$b = 1 - P(F>F_{trend}|F>1) \quad ,$$

gdzie F jest wartością statystyki F Snedecora wyznaczoną przy testowaniu hipotezy o równości średnich obiektowych gdy nie przeprowadza się poprawiania obserwacji, tzn. zakłada się, że trend nie występuje, a F_{trend} jest wartością tej samej statystyki dla poprawionych obserwacji. Cykl iteracyjny rozpoczyna się przyjmując b=1 i wyznaczając odpowiednią wartość F_{trend}.

W metodzie tej estymatory parametrów obiektowych wyznacza się poprzez przyrównanie sum obiektowych dla poprawionych obserwacji do ich wartości oczekiwanych.

Ponieważ błędy standardowe różnic obiektowych zależą w istotny sposób od rozmieszczenia obiektów w doświadczeniu, Wilkinson i in. (1983) wprowadzają pojęcie NN zrównoważenia, które zapewnia optymalną efektywność estymatora.

Analiza BK (Besag i Kempton 1984, Kempton 1984, Besag i Kempton 1986) nie wymaga specjalnego układu doświadczalnego. Może być zatem stosowana w każdym eksperymencie, w którym żyzność gleby wykazuje dużą zmienność, co spotyka się szczególnie w ostrych klimatach. Autorzy oczekują, że proponowana aproksymacja poprawi estymatory obiektowe w wypadku, gdy nie zastosowano właściwego układu klasycznego.

Besag i Kempton przedstawiają swoją metodę dla pojedynczego pasa poletek, choć zaznaczają, że może być ona łatwo rozszerzona na przypadek wielu niezależnych pasów.

Niech więc Y_i oznacza plon z i-tego poletka, a Y odpowiedni wektor złożony z n obserwacji. Model dla wektora losowego Y ma postać

$$Y = \gamma 1 + T\tau + X + Z \quad ,$$

gdzie γ jest efektem ogólnym, 1 oznacza n-wymiarowy wektor złożony z jedynek, a T, τ, X i Z są jak w (2). Usunięcie liniowego lokalnego trendu żyzności gleby uzyskuje się poprzez analizę pierwszych różnic wartości obserwowanych. Wprowadza się więc zmienne $U_i = Y_i - Y_{i+1}$, co w notacji macierzowej zapisać można jako $U = \Delta Y$, gdzie Δ jest $(n-1)$xn-wymiarowym operatorem pierwszych różnic. Zakłada się, że różnice $X_i - X_{i+1}$ są nieskorelowanymi zmiennymi losowymi z wartością oczekiwaną $E(\Delta X) = 0$ oraz wariancją $Var(\Delta X) = \sigma_x^2 I$, natomiast $E(Z) = 0$ i $Var(Z) = \sigma^2 I$, a stąd $E(U) = \Delta T\tau$ i $Var(U) = \sigma_x^2 I + \sigma^2 \Delta\Delta'$.

Rozpatrując dwa skrajne przypadki, $\sigma_x^2 = 0$ i $\sigma^2 = 0$, zauważyć można, że pierwszy z nich prowadzi do klasycznego modelu z nieskorelowanymi błędami Z, a drugi pozwala na uzyskanie estymatora τ opartego na $\Delta'\Delta Y$, co zapewnia eliminację lokalnego liniowego trendu.

Estymator wektora parametrów obiektowych w analizie BK uzyskuje się metodą iteracyjną. W pierwszym kroku wyznacza się estymator $\hat{\tau}$ uogólnioną metodą najmniejszych kwadratów przyjmując $\phi = \sigma_x^2/\sigma^2 = 0$. Następnie estymator ten podstawia się do modelu wyznaczając błąd $e = U-\Delta T\hat{\tau}$. W kolejnym kroku metodą największej wiarogodności oblicza się nową wartość ϕ minimalizując odpowiednią funkcję e. Z tak wyznaczoną wartością ϕ wraca się do kroku pierwszego.

Praca Williamsa (1986) zawiera jeszcze jedno podejście do problemu analizy uwzględniającej sąsiedztwo poletek. Williams zakłada, że porównywane obiekty rozmieszczone są w kompletnych replikacjach, które tworzą rozkładalny układ o b blokach niekompletnych. Dla obserwacji przyjmuje następujący model liniowy

$$Y = \gamma 1 + R\pi + T\tau + X \quad ,$$

gdzie γ jest efektem ogólnym, π i τ są odpowiednio r i v-wymiarowymi wektorami efektów replikacji i efektów obiektów, R i T są odpowiadającymi im macierzami układu, X jest n-wymiarowym wektorem efektów losowych o zerowej wartości oczekiwanej, $E(X) = 0$ i macierzy dyspersji $Var(X) = \sigma^2(I + \alpha P - \phi F)$. W macierzy dyspersji, $Var(X)$, parametr σ^2 jest wariancją błędu losowego, $\alpha = \sigma_\beta^2/\sigma^2$, gdzie σ_β^2 jest wariancją dla bloków niekompletnych, $P = B(B'B)^{-1}B'$, gdzie B jest nxb-wymiarową macierzą układu dla bloków, $\phi = \sigma_z^2/\sigma^2$, gdzie σ_z^2 jest komponentem wariancji związanym z żyznością, a F jest blokową macierzą diagonalną o elementach $3L_{k_i}/(k_i^2 - 1)$, gdzie k_i jest wielkością i-tego bloku, a L_{k_i} jest $k_i x k_i$-wymiarową macierzą, której (j,l)-ty element jest równy $|j-l|$.

Williams proponuje następujący tok postępowania. Wyestymować parametry σ^2, α, ϕ poprzez przyrównanie odpowiednich form kwadratowych do ich wartości oczekiwanych, a następnie wyznaczyć $\hat{\tau}$ uogólnioną metodą najmniejszych kwadratów. Możliwe są przy tym następujące modyfikacje rozważanego modelu. Jeśli $\phi=0$, to model redukuje się do modelu dla bloków niekompletnych. Ponadto, gdy $\alpha=0$, wtedy model redukuje się do modelu dla bloków kompletnych.

3. METODY ANALIZY UWZGLĘDNIAJĄCE WSPÓŁDZIAŁANIE MIĘDZY POLETKAMI

We wczesnych fazach hodowli istotnym problemem, na jaki natrafia hodowca, jest mała ilość materiału doświadczalnego (np. nasion). W tej sytuacji zwykle stosuje się małe poletka, co pozwala zwiększyć liczbę replikacji uwzględniając tym samym warunki środowiskowe. Jednakże zastosowanie małych poletek może spowodować wystąpienie zjawiska konkurencji międzyodmianowej związanej z bliskim sąsiedztwem w jakim znajdują się różne odmiany na polu. Jednym sposobem eliminacji tego zjawiska jest stosowanie poletek ochronnych otaczających każde z poletek doświadczalnych. Jest to jednak postępowanie bardzo kosztowne. Innym rozwiązaniem jest uwzględnienie efektu konkurencji przy modelowaniu obserwacji. Zagadnieniem tym zajmują się między innymi Kempton i Howes (1981), Kempton (1982), Kempton i Lockwood (1984), Kempton (1985), Green, Jennison i Seheult (1985) oraz Besag i Kempton (1986). Rozważają oni dwa typy modeli. Pierwszy, który zawiera pojedyncze efekty konkurencji każdej z odmian i który stosuje się przy badaniu konkurencji międzypoletkowej w specjalnych układach doświadczalnych, a mianowicie w takich, w których występuje zrównoważenie ze względu na sąsiedztwo. Drugi z modeli można stosować w prostych doświadczeniach, w których nie musi występować zrównoważenie. Jest to model analizy kowariancji.

Jak podają Besag i Kempton (1986) pierwszy z modeli można zapisać w postaci

$$Y = B\eta + T\tau + RT\phi + Z \quad ,$$

gdzie η jest b-wymiarowym wektorem efektów blokowych, τ i ϕ są p-wymiarowymi wektorami efektów obiektowych bezpośrednich i pośrednich (sąsiedzkich), B jest nxb-wymiarową macierzą układu dla bloków, T jest nxp-wymiarową macierzą układu dla obiektów, R jest nxn-wymiarową macierzą układu sąsiadów, a Z jest n-wymiarowym wektorem błędów losowych. Draper i Guttman (1980) omawiają specjalny przypadek tego modelu, gdy $\phi = \alpha\tau$, gdzie α jest współczynnikiem oddziaływania. Model ten można stosować także w wypadku doświadczeń, w których obiekt nie jest precyzyjnie umieszczony na poletku (np. spryskiwanie).

Estymację parametrów obiektowych w modelu Besaga i Kemptona przeprowadza się stosując metodę najmniejszych kwadratów.

W drugim podejściu model obserwacji ma postać

$$Y = B\eta + T\tau + \beta WY + Z \quad ,$$

gdzie W jest nxn-wymiarową macierzą wag o elementach (i,i+1) oraz (i,i-1) równych 1/2 i pozostałych równych zero, β (<0) jest współczynnikiem konkurencji dla najbliższych sąsiadów, a pozostałe elementy są takie jak w modelu pierwszym.

W wypadku układów zrównoważonych ze względu na sąsiedztwo istnieje również możliwość wprowadzenia różnych współczynników konkurencji dla różnych sąsiadów. Pozwala to ocenić "czułość" (sensivity) odmiany na konkurencję i agresywność odmiany. Wtedy β_{rs} czyli współczynnik oddziaływania poletka s na poletko r zależy od odmian znajdujących się na tych poletkach, $\beta_{rs} = \delta_r\gamma_s$, gdzie δ_r jest czułością odmiany r, a γ_s jest agresywnością odmiany s.

Przedstawiony model analizy kowariancji jest wg autorów właściwy dla roślin okopowych. Dla fasoli i zbóż autorzy zalecają przeprowadzenie analizy kowariancji, w której cechą towarzyszącą jest nie plon lecz inna cecha ściśle związana z plonem (np. wysokość roślin na sąsiednich poletkach).

LITERATURA

Bartlett, M.S. (1938). The approximate recovery of information from field experiments with large blocks. *J.Agric.Sci.*, **28**, 418-427.

Bartlett, M.S. (1978). Nearest neighbour models in the analysis of field experiments. *J.R.Statist.Soc. B*, **40**, 147-174.

Besag, J.E. (1984). A method for analysis of field experiments based on first differences, including a GENSTAT macro for its implementation RAK. W: *Spatial Methods in Field Experiments.* University of Durham, England.

Besag, J.E., Kempton, R.A. (1984). Spatial methods in the analysis of agricultural field trials. *Proc. 12th International Biometric Conference*, Tokyo, Japan.

Besag, J.E., Kempton, R.A. (1986). Statistical analysis of field experiments using neighbouring plots. *Biomertics* **42**, 231-251.

Draper, N.R., Guttman, I. (1980). Incorporating overlap effects from neighbouring units into response surface models. *Applied Statistics* **29**, 128-134.

Green, P.J., Jennison, C., Seheult, A.H. (1985). Analysis of field experiments by least squares smoothing. *J.R. Statist. Soc.,B*, **47**, 299-315.

Kempton, R.A. (1982). Adjustment for competition between varietes in plant breeding trials. *J.Agric.Sci.* **98**, 599-611.

Kempton, R.A., Howes, C.W. (1981). The use of neighbouring plot values in the analysis of variety trials. *Appl. Statist.*, **30**, 1, 59-70.

Kempton, R.A., Lockwood, G. (1984). Interplot competition in variety trials of field beans *Vicia faba L.*. *J.Agric.Sci.* **103**, 293-302.

Kempton, R.A. (1984). Nearest neighbour analysis of interplot competition. W: *Spatial Methods in Field Experiments.* Univeristy of Durham, England.

Kempton, R.A. (1985). Statistical models for interplot competition. *Aspects of Appl.Biology* **10**.

Papadakis, J.S. (1937). Methode statistique pour les experiences en champ. *Bull.Inst.Amel.Plantes a Salonique*, **23**.

Papadakis, J.S. (1970). *Agricultural Research.* Buenos Aires.

Wiancko, A.T. (1914). Use and management of check plots in soil fertility investigations. *J.Amer.Soc.Agron.*, 6, 122-124.

Wilkinson, G.N., Eckert, R.S., Hancock, T.W., Mayo, O. (1983). Nearest neighbour NN analysis of field experiments. *J.R.Statist.Soc., B*, **45**, 151-211.

Wilkinson, G.N. (1984). Nearest neighbour methodology for design and analysis of field experiments. *Proc. 12th International Biometrics Conference*, Tokyo, Japan.

Williams, E.R. (1986). A neighbour model for field experiments. *Biometrika*, 73, 2, 279-287.

Załęski, E. (1927). *Metodyka Doswiadczen Rolniczych*. Piller-Neumann, Lwów.

Praca wpłynęła 15 czerwca 1987 :
w wersji ostatecznej 11 grudnia 1987.