

ZYGMUNT ŁUBKOWSKI

KILKA UWAG W SPRAWIE ORGANIZACJI DOŚWIADCZEŃ WIELOKROTNYCH

Wnioski z pojedynczego doświadczenia właściwe są tylko dla warunków, w jakich to doświadczenie wykonano. Jeśli powtórzymy doświadczenie dokładnie w tych samych warunkach i z tym samym materiałem, to powinniśmy otrzymać wyniki nie różniące się istotnie od poprzednich, a więc potwierdzające wnioski wyciągnięte z doświadczenia pierwszego. Wyniki pojedynczego doświadczenia będą w tym wypadku populacją próbną stanowiącą podstawę do charakterystyki badanej cechy populacji generalnej, którą będzie ogół wyników osiągalnych przy powtarzaniu doświadczenia w tych samych warunkach i z tym samym materiałem.

Warunkom tym odpowiadają niektóre doświadczenia z fizjologii roślin. W doświadczeniach wazonowych możemy bardzo dokładnie wyrównać glebę. W szklarniach możemy kontrolować temperaturę, wilgotność i skład powietrza; wprawdzie w dotychczasowych szklarniach wpływ tych czynników zmienia się w związku z niejednakową odległością wazonów od źródła ciepła, światła i dostępu powietrza, jednak nowoczesne pracownie tego typu, tzw. fytotrony, pozwalają zapewnić badanym obiektom dokładnie takie same warunki w ciągu całego trwania doświadczenia. Największą trudność stanowi sprawa jednolitości materiału. Gdyby się udało tę trudność pokonać oraz poznać właściwości badanego materiału, to z jednego doświadczenia można by wyciągnąć wnioski miarodajne dla tego ściśle określonego materiału i danych warunków.

W doświadczalnictwie polowym nie tylko nie możemy doświadczenia powtórzyć w zupełnie takich samych warunkach, ale zwykle nie potrafimy poznać dokładnie tych warunków, a tym bardziej ustalić, jaki wpływ wywarły one na wyniki. Na warunki składa się cały szereg czynników, które możemy podzielić na trzy główne grupy, mianowicie na czynniki związane z glebą, z przebiegiem pogody i z agrotechniką. Czynniki te działają zespołowo, a więc nawet gdybyśmy poznali, jaki wpływ na wyniki wywierają niektóre z nich, nie moglibyśmy liczbowo ująć ich efektu. A więc nie możemy i nie będziemy nigdy mogli opisać dokładnie populacji generalnej, której populację próbną

stanowi konkretne pojedyncze doświadczenie. Jednak wydawanie zaleceń dla produkcji musimy oprzeć na dokładnym określeniu warunków, dla jakich zalecenia są miarodajne, a więc czynników klimatycznych czyli przeciętnego z wielu lat przebiegu pogody, warunków glebowych i poziomu agrotechniki. Wyniki doświadczeń, na podstawie których będziemy wydawać zalecenia, powinny być reprezentatywne dla tych warunków. Jak wiadomo, populacja próbna reprezentuje dobrze generalną, jeśli jest z niej pobrana losowo. Należy przede wszystkim zastanowić się nad tym, czy i o ile doświadczalnictwo nasze spełnia ten warunek.

W doświadczeniu założonym jedną z metod poprawnych obiekty otrzymały poletka przydzielone losowo, wyniki reprezentują więc dobrze populację generalną, którą stanowią hipotetyczne plony, jakie można by uzyskać na danym polu doświadczalnym, gdyby je obsiano w całości równocześnie każdym z badanych obiektów. Ocena tej populacji generalnej będzie tym dokładniejsza, im więcej zastosujemy powtórzeń. Jednak celem naszych badań jest inna populacja generalna. Czy przez zastosowanie w poszczególnym doświadczeniu większej liczby powtórzeń wyniki tego doświadczenia będą ją lepiej reprezentowały? I czy dzięki temu będziemy mogli otrzymać tę samą ilość informacji o populacji generalnej przez przeprowadzenie mniejszej liczby doświadczeń wielopowtórzeniowych zamiast większej powtórzonych mniej razy? Tak byłoby, gdyby wynik każdego z doświadczeń stanowiących daną serię był pobrany losowo z populacji generalnej, którą pragniemy zbadać. W tym bowiem wypadku zmienność wewnątrzobektowa, a więc zmienność między plonami tego samego obiektu z różnych doświadczeń, miałaby charakter losowy. Byłaby ona tym mniejsza, im lepiej byłyby wyznaczone średnie obiektowe poszczególnych doświadczeń. Nie mamy jednak żadnych podstaw do twierdzenia, że nasza populacja próbna jest pobrana losowo z populacji generalnej.

W wypadku powtarzania doświadczeń w czasie jest to nawet niemożliwe. Nie możemy z serii lat, w ciągu których mamy wyniki stosować w produkcji, wybrać losowo kilka i w nich przeprowadzić doświadczenia. Sprawa komplikuje się jeszcze przez to, że zmienność tzw. sezonowa nie wynika jedynie z odmiennego w poszczególnych latach przebiegu pogody; składają się na nią także czynniki glebowe — doświadczeń nie zakłada się przecież co roku na tym samym polu — i agrotechniczne, związane częściowo z przebiegiem pogody. Wnioskując z krótkiego cyklu narażamy się na popełnienie błędu systematycznego, mianowicie na ocenę populacji z próby, która jej nie reprezentuje. Przedłużanie cyklu zmniejsza tę obawę, powoduje jednak konieczność długiego oczekiwania na wyniki. W każdym razie lepsze wyzna-

czenie średnich poszczególnych doświadczeń zupełnie nie wpłynie na reagowanie obiektów na warunki, jakie zostały w poszczególnych latach, nie zmniejszy zatem zmienności współdziałania obiektowo-sezonnego.

Pewien wpływ na tę zmienność może mieć równoczesne powtarzanie wyników w czasie i przestrzeni oczywiście pod warunkiem, że będzie to układ w pełni ortogonalny, a więc że wszystkie doświadczenia będą założone co roku w tych samych miejscowościach i że miejscowości te, a raczej pola, będą reprezentatywne dla danego rejonu. Przebieg pogody wykazuje bowiem nieraz z roku na rok większą zmienność w tej samej miejscowości, niż na terenie całego rejonu. Powtarzanie wyników w przestrzeni pozwala też często zmniejszyć tę część zmienności sezonowej, której przyczyną są czynniki agrotechniczne. Wydaje się więc, że — nawet gdy pragniemy uzyskać wyniki miarodajne dla rejonu jednolitego pod względem gleby i klimatu — uzyskamy je szybciej, jeśli doświadczenia przeprowadzimy równocześnie w kilku punktach tego rejonu.

W praktyce mało jest rejonów zupełnie jednolitych pod względem położenia, gleby i klimatu. Ponadto w większości wypadków interesują nas bardziej przeciętne wyniki, jakie moglibyśmy uzyskać w rejonie większym, w którym istnieją mniejsze lub większe odchylenia od warunków uznanych za typowe dla tego rejonu. Jeśli weźmiemy np. Żuławy, rejon stosunkowo niewielki i bardzo specyficzny, o braku zróżnicowania wynikającego z rzeźby terenu, o jednakowych niemal warunkach klimatycznych i pozornie także jednolitych glebach, to przy bliższej analizie stwierdzimy, że istnieje szereg odmian tych mad żuławskich związanych z okresem ich powstania, głębokością poziomu wody gruntowej, składem mechanicznym utworów, z których się wytworzyły itp. A cóż dopiero mówić o rejonach większych, takich jak Pojezierze Mazurskie czy Wyżyna Małopolska. W niektórych wypadkach możemy otrzymać wyniki jednakowe, pomimo że rejon nie jest jednolity. Jeśli byśmy np. chcieli zbadać, czy Żuławy nadają się pod uprawę rzepaku ozimego i jaki jest optymalny termin siewu tego rzepaku, to prawdopodobnie w całym rejonie otrzymamy takie same wyniki. Wydaje się natomiast, że doświadczenia nawozowe nie wykazałyby jednolitego reagowania gleb żuławskich na poszczególne składniki.

Sprawę komplikuje fakt, że granice rejonów nie dadzą się na ogół ściśle wyznaczyć, że istnieje zawsze pas przejściowy, nieraz znacznej szerokości. Można postępować dwojako: albo interesuje nas tylko jeden rejon, którego granice wyznaczamy na podstawie mapy gleb i rzeźby terenu oraz danych klimatycznych, albo też chcemy sprawdzić, czy wyniki doświadczeń potwierdzą podział na rejonny przeprowadzony na

podstawie danych przyrodniczych i czy pozwolą wyznaczyć możliwie dokładnie granice między tymi rejonami.

W pierwszym przypadku powinniśmy dążyć do tego, by uzyskana seria wyników, jako populacja próbna, była reprezentatywna dla populacji generalnej, którą otrzymalibyśmy, gdybyśmy badany zabieg zastosowali pod daną roślinę we wszystkich gospodarstwach rejonu. Należałoby próbę tę pobrać losowo. Sprawa nie przedstawiałaby większych trudności, gdyby każde gospodarstwo mogło przeprowadzić doświadczenie z odpowiednią dokładnością. W tym wypadku można by doświadczenie rozmieścić losowo przez zastosowanie losowania dwustopniowego, a więc wylosowanie najpierw miejscowości w liczbie odpowiadającej liczebności danej serii, a następnie jednego gospodarstwa w każdej miejscowości. Wybrane losem gospodarstwa miałyby założyć doświadczenia w warunkach odpowiadających instrukcji. W ten sam sposób należałoby rozmieszczać doświadczenia w każdym roku danego cyklu.

W naszych warunkach nie możemy w ten sposób uzyskać populacji próbnej. Musimy przeprowadzać doświadczenia tam, gdzie — jak się spodziewamy — otrzymamy wyniki poprawne. Powinniśmy więc w inny sposób zapewnić naszej próbie możliwie wysoki stopień reprezentatywności. Podstawową zasadą będzie dążenie do unikania błędów systematycznych, które wypaczyłyby wyniki. Jeśli więc np. w terenie falistym przeprowadzimy doświadczenia tylko w dolinach, to wyniki nie będą miarodajne dla stoków i grzbietów. Podobnym wypaczeniem byłoby umieszczenie doświadczeń np. w większości na stokach południowych. Przed rozplanowaniem doświadczeń należałoby więc przeprowadzić szczególnie wnikliwą analizę terenu. Niemniej jednak nie mamy sposobu na stwierdzenie, że uzyskana seria wyników stanowi próbę reprezentatywną dla populacji generalnej.

Gdyby ta próba była pobrana losowo, to liczebność jej nie odgrywałaby większej roli. Przed przecenianiem istotności wyników uzyskanych z nielicznej próby chroniłyby nas odpowiednio wyższe wartości kryteriów statystycznych. Inaczej przedstawia się sprawa, jeśli doświadczenia rozmieszczamy nie losowo, lecz na podstawie znajomości warunków. W tym wypadku ryzyko wypaczenia wyników będzie tym większe, im mniejsza będzie liczebność serii. Im liczniejsze będziemy zakładali doświadczenia tym większa będzie możliwość, że rozmieścimy je we wszystkich mikrorejonach danego rejonu, a więc np. w terenie falistym na grzbietach, stokach i podnóżach wzgórz, w dolinkach i kotlinkach, na stokach o różnej wystawie i nachyleniu, a także będziemy mogli uwzględnić różny poziom agrotechniki gospodarstw, dla których będą przeznaczone nasze zalecenia.

Podstawowym warunkiem poprawności — w znaczeniu praktycz-

nym — wnioski z doświadczeń polowych jest więc duża liczebność serii. Liczebności tej nie możemy zmniejszyć przez zakładanie poszczególnych doświadczeń w większej liczbie powtórzeń. Spowodowałyby to bowiem tylko lepsze wyznaczenie średnich obiektowych z poszczególnych doświadczeń i w związku z tym zmniejszyłyby błędy, jakimi byłyby obciążone poszczególne średnie, nie wpłynęłyby jednak zupełnie na reprezentatywność tych wyników. Wynika z tego bardzo ważny problem, a mianowicie, czy celowe jest zakładanie doświadczeń polowych o większej liczbie powtórzeń, czy nie byłoby słuszne zmniejszenie stosowanej u nas dotąd liczby 6 powtórzeń i przeznaczenie zaoszczędzonych dzięki temu sił i środków na odpowiednie zwiększenie liczebności serii.

W ostatnich latach sprawa stosunku między liczbą lat, miejscowości i powtórzeń w doświadczeniach polowych bywa dość często omawiana w literaturze. W krajach zachodnich, przede wszystkim w Stanach Zjednoczonych A. P., podstawę takich rozważań stanowi ekonomiczny punkt widzenia. Chodzi mianowicie o to, w jaki sposób można — bez straty informacji — zmniejszyć nakład kosztów na doświadczalnictwo, względnie jak je zorganizować, by tymi samymi funduszami uzyskać jak najwięcej informacji. Głosy najdalej idące opowiadają się za bardzo liczebnymi seriami doświadczeń bezpowtórzeńowych, co w Niemczech propagował Mudra, a u nas przed 10 laty Saloni w — o ile wiem — nie opublikowanym referacie. Kempthorne zaleca zakładanie doświadczeń dwupowtórzeńowych, co umożliwia wydzielenie ze zmienności ogólnej bodaj części nieściśłości poszczególnych doświadczeń i zorientowanie się w przybliżeniu, jakimi błędami obciążone są poszczególne średnie. Wprawdzie w pojedynczych doświadczeniach ocena nieściśłości będzie oparta na małej liczbie stopni swobody, jednak w całej serii liczba ta będzie wystarczająco duża.

Ostatnio zajął się tą sprawą Rundfeldt, przy czym oparł się on przede wszystkim na pracy Yatesa. Rozważa on wpływ, jaki na ilość uzyskiwanych informacji wywiera zwiększanie liczby powtórzeń, najpierw z punktu widzenia teoretycznego, a następnie na przykładzie konkretnej serii doświadczeń z 6 odmianami jarej pszenicy porównywanymi w 6 punktach przez 3 lata, w każdym doświadczeniu w 6 powtórzeń. Autor podaje pełną analizę zmienności, z czego do naszych celów potrzebne są wartości stosunków niektórych zmienności. Wynoszą one: stosunek zmienności współdziałania odmian z miejscowościami w latach do sumarycznego błędu — 3,21, istotny przy P większym od 0,99, oraz stosunki zmienności współdziałań obiektów z latami i obiektów z miejscowościami do potrójnego współdziałania; pierwszy z nich jest równy 1,97, drugi 1,80; prawdopodobieństwo istotności wpływu obu tych

współdziałań wynosi między 0,90 a 0,95. Analizując uzyskane liczby autor dochodzi do wniosku, że w doświadczeniach o podobnym rozkładzie zmienności zwiększanie liczby lat pozwala uzyskać przeszło dwukrotnie więcej informacji, a liczby miejscowości blisko dwu i półkrotnie, niż zwiększanie liczby powtórzeń. Tak bywa zawsze, jeśli współdziałania są większe od zera. Gdyby analiza zmienności nie wykazała wpływu współdziałań, gdybyśmy zatem mogli przyjąć, że powtórzenie doświadczeń w czasie i przestrzeni nie spowodowało dodatkowej zmienności, to moglibyśmy całą serię traktować jako jedno doświadczenie, a liczba stopni swobody sumarycznego błędu takiej serii równałaby się: $(lmn-1)(k-1)$, gdzie l , m i n symbolizują liczby lat, miejscowości i powtórzeń a k liczbę obiektów. W tym wypadku byłoby oczywiście obojętne czy zwiększymy liczbę lat, miejscowości czy powtórzeń; można by nawet poprzestać na jednym doświadczeniu o lmn powtórzeniach. W większości serii doświadczeń zaznacza się jednak istotny wpływ współdziałań, przeważnie bardzo znaczny.

Sprawę można rozważyć z innego jeszcze punktu widzenia. W podanym przez Rundfeldta przykładzie każdy obiekt powtórzony był $6 \cdot 6 \cdot 3$, a więc 108 razy; łączna liczba stopni swobody zmienności mogących stanowić podstawę oceny istotności różnic międzyobiektowych wynosi więc: $107 \cdot 5 = 535$ i rozkłada się na zmienności: sumarycznego błędu, 450 stopni swobody, potrójnego współdziałania 50 stopni swobody, współdziałania odmian z miejscowościami — 25 stopni swobody, i z latami — 10 stopni swobody. Jeśli bodaj jedno współdziałanie będzie istotne, to odpada nam przy ocenie istotności różnic międzyobiektowych blisko 90% ogólnej liczby stopni swobody. Jeśli zmienność współdziałania odmian z latami będzie istotna w stosunku do potrójnego współdziałania, to sprawdzianem istotności różnic będzie zmienność oparta tylko na 10 stopniach swobody. Zmienność ta jako istotna będzie przy tym stosunkowo duża, a więc nie będziemy mogli liczyć na uzyskanie z tej serii większej ilości informacji.

Gdybyśmy zmniejszyli liczbę powtórzeń do 3, a zamiast tego wykonali doświadczenie w 9 punktach i przedłużyli cykl o 1 rok, to każdy obiekt byłby też powtórzony 108 razy, ale z łącznej liczby 535 stopni przypadłoby na zmienność sumarycznego błędu 360 stopni swobody, na potrójne współdziałanie 120 stopni, na współdziałanie obiektowo-lokalne 40 stopni, a obiektowo-sezonowe 15 stopni. Liczba stopni swobody sumarycznego błędu byłaby więc mniejsza, ale dostateczna dla oceny istotności wpływu współdziałań.

W przeprowadzonych u nas doświadczeniach istotność zmienności współdziałań występuje często jeszcze wyraźniej. Obrazują to zestawione w tabeli 1 fragmenty analiz zmienności serii doświadczeń rejo-

nizacyjnych z odmianami owsa, wstępnych z odmianami owsa i rejonizacyjnych z odmianami jęczmienia wykonanych w latach 1953—1955 (tabela 1).

Tabela 1

Fragmenty analiz zmienności serii doświadczeń odmianowych z lat 1953—1955

Zmienność	Rejonizacyjne z owsem		Wstępne z owsem		Rejonizacyjne z jęczmieniem		Symbol zmienności
	liczba stop-ni swobody	przec. zmienność	liczba stop-ni swobody	przec. zmienność	liczba stop-ni swobody	przec. zmienność	
Współdziałania odmian z latami	16	12,74	18	13,70	22	21,04	S ₁
Współdziałania odmian z miejscowościami	176	3,67	144	6,66	363	4,17	S ₂
Współdziałania odmian z miejscowościami w latach	342	4,36	279	5,49	720	5,03	S ₃
Sumarycznego błędu	2710	1,255	2243	1,388	5620	1,15	S ₄
$\frac{S_3}{S_4} =$	3,47 F (P = 0,99) < 1,28		3,96 F (P = 0,99) < 1,28		4,37 F (P = 0,99) < 1,19		
$\frac{S_1}{S_3} =$	2,92 F (P = 0,99) < 2,09		2,50 F (P = 0,99) < 2,09		4,18 F (P = 0,99) < 1,90		

Wpływ współdziałań zaznaczył się we wszystkich tych seriach bardzo silnie, znacznie wyraźniej niż w serii omawianej przez Rundfeldta. Można by do liczb tych zastosować podany przez Rundfeldta wzór Yatesa i obliczyć, w jakim stosunku na ilość informacji wpłynęłaby w każdej z tych serii zmiana liczby lat, miejscowości i powtórzeń. Spróbowałem jednak ująć tę sprawę inaczej, mianowicie przeliczyć serię doświadczeń wstępnych z odmianami owsa w ten sposób, że do obliczeń wzięto z każdego doświadczenia zamiast 6 tylko pierwsze 2 powtórzenia. Odpowiadający tabeli 1 fragment analizy zmienności zestawiono w tabeli 2.

Przy wyciągnięciu wniosków z porównania odpowiednich 2 fragmentów analiz zmienności należałoby wziąć pod uwagę następujące wyjaśnienia:

1. Celem ostrzejszego naświetlenia problemu wzięłem wypadek krańcowy, a więc zmniejszyłem liczbę powtórzeń z 6 na 2, a nie na 3 lub 4.

2. W danej serii nie zrekompensowano zmniejszenia liczby powtó-

rzeń pojedynczego doświadczenia odpowiednim przedłużeniem cyklu i zwiększeniem liczby miejscowości.

Tabela 2

Fragment analizy zmienności

Zmienność	Liczba stopni swobody	Przeciętna zmienność
Współdziałanie odmian z latami	18	11,98
Współdziałanie odmian z miejscowościami	144	10,35
Współdziałanie odmian z miejscowościami w latach	279	9,69
Sumarycznego błędu	450	4,046

3. W roku 1953 w niektórych Stacjach Doświadczalnych Oceny Odmian doświadczenia były obarczone bardzo dużym błędem. Był to pierwszy rok działalności Stacji, doświadczenia te założono prawdopodobnie na zupełnie nieodpowiednich polach. Skutkiem tego błąd niektórych doświadczeń był szczególnie duży. Wystąpiło to znacznie wyraźniej, jeśli do obliczenia wyników wzięto tylko 2 powtórzenia; blisko 50% sumy kwadratów sumarycznego błędu wypada na 8 doświadczeń. Gdybyśmy te doświadczenia zdyskwalifikowali na podstawie kryterium podanego przez Barbackiego, Bilskiego i Rosnowskiego to przeciętna zmienność sumarycznego błędu zmniejszyłaby się z 4,046 na 2,42 przy 378 stopniach swobody.

Jednak nawet tak znaczne, blisko potrójne zwiększenie przeciętnego błędu nie wpłynęło istotnie na interpretację wyników. W szczególności nie spowodowało ono zmiany wysokości przeciętnego współdziałania odmian z latami. Ponieważ w serii doświadczeń wstępnych zmienność współdziałania odmian z latami była istotna w stosunku do potrójnego współdziałania, a więc ona stanowiła sprawdzian istotności różnic międzyodmianowych, seria dwupowtórzeniowa pozwoliłaby stwierdzić tę istotność z taką samą dokładnością co sześciopowtórzeniowa.

Ograniczenie liczby powtórzeń do $\frac{1}{3}$ umożliwiło by jednak zmniejszenie, co najmniej o połowę, nakładu pracy i kosztów związanych z przeprowadzeniem omawianych doświadczeń. Odpowiednie zwiększenie liczebności serii uczyniłoby ją bardziej reprezentatywną i pozwoliłoby uzyskać więcej informacji, mających stanowić podstawę zaleceń dla produkcji. W szczególności w doświadczeniach odmianowych przeprowadzenie rejonizacji, a więc ustalenie zasięgu rejonów i doboru odmian właściwych dla każdego rejonu, możliwe jest tylko na podstawie serii o dużej liczebności.

W naszych warunkach ograniczenie liczby powtórzeń do 2 byłoby zapewne zbyt daleko idące tym bardziej, jeśli nie zwiększono by równocześnie mniej więcej 2-krotnie liczby punktów doświadczalnych. Wydaje się jednak, że stosowane u nas dotychczas wnioskowanie z małej serii doświadczeń 6-powtórzeniowych nie jest właściwe i że należałoby dążyć do uzyskiwania liczniejszych serii wyników i opracowywać w miarę możliwości dłuższe cykle przy równoczesnym zmniejszeniu liczby powtórzeń na cztery, a nawet na trzy.

LITERATURA

1. Barbacki S., Bilski E., Rosnowski St.: Wyniki doświadczeń z odmianami buraków cukrowych wykonanych w latach 1947—1951. Warszawa 1956.
2. Kempthorn O.: The design and analysis of experiments. New York 1953.
3. Łubkowski Z.: Wyniki doświadczeń z odmianami owsa wykonane w latach 1953—1955 (w druku).
4. Mudra A.: Einführung in die Methodik der Feldversuche. Hirzel, Leipzig.
5. Rundfeldt H.: Zur Berechnung eines optimalen Verhältnisses der Anzahl der Prüffahre, der Prüforten und der Vergleichteilstücke bei Feldversuchen. Ztschr. f. Pflanzenzüchtung, tom 37, z. 2.
6. Yates F.: Modern experimental design and its function in plant selection (cyt. za H. Rundfeldtem).