

## 5 ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА ПРИ ПОИСКЕ ОПТИМАЛЬНЫХ УСЛОВИЙ

### 5.1 Основные понятия и определения

Эксперимент занимает центральное место в науке. А применение математических методов планирования эксперимента является одним из возможных путей повышения эффективности научных исследований.

Планирование эксперимента – это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью.

При планировании эксперимента существенно следующее:

- стремление к минимизации общего числа опытов;
- одновременное варьирование всеми переменными, определяющими процесс, по специальным правилам – алгоритмам;
- использование математического аппарата, формализующего многие действия экспериментатора;
- выбор четкой стратегии, позволяющей принимать обоснованные решения после каждой серии экспериментов.

В области механизации сельскохозяйственного производства планирование эксперимента широко применяется для определения оптимальных геометрических параметров и кинематического режима работы как новых, так и модернизируемых технических средств.

Эксперимент, который ставится для решения задач оптимизации, называется экстремальным. Это название связано с глубокой аналогией между оптимизацией и поиском экстремума некоторой функции.

При планировании эксперимента удобно пользоваться так называемым кибернетическим подходом, или методом «черного ящика».

Кибернетический метод относится к числу экспериментально-теоретических. Использование этого метода обусловлено сложностью рассматриваемых процессов и объектов, в них участвующих: многофакторностью исходных и определяющих условий, необходимостью многокритериальной оптимизации, неполнотой знаний об исследуемых объектах и т. п. Сущность этого подхода поясняется схемой на рисунке 5.1.

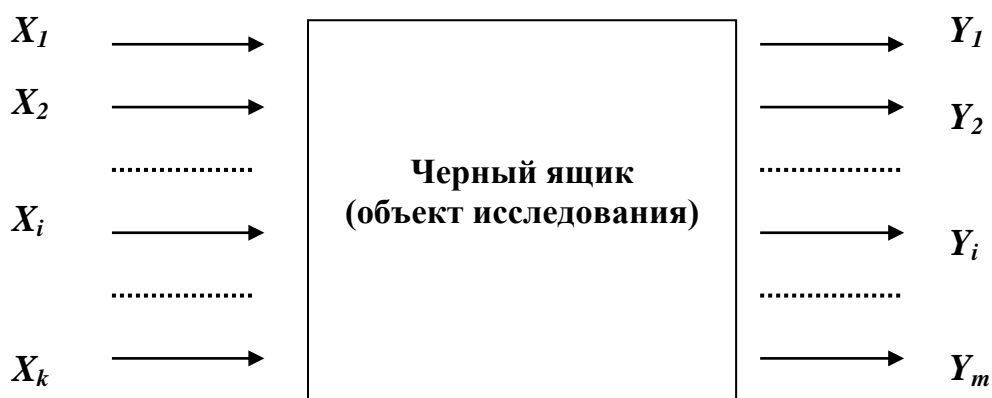


Рисунок 5.1 Схема кибернетического подхода к объекту исследования.

Стрелки справа изображают численные характеристики целей исследования. Мы обозначаем их буквой  $Y$  и называем параметрами оптимизации. В литературе встречаются и другие названия: критерий оптимизации, критерий эффективности, целевая функция, выход «черного ящика» и т. д.

Для проведения эксперимента необходимо иметь возможность воздействовать на поведение «черного ящика». Все способы такого воздействия мы обозначаем буквой  $X$  и называем факторами. Их называют также входами «черного ящика».

При решении задачи оптимизации будем использовать математические модели объекта исследования. Под математической моделью будем понимать уравнение, связывающее параметр оптимизации с факторами. В общем виде его можно записать следующим образом

$$Y=f(X_1, X_2, \dots, X_k). \quad (5.1)$$

Это уравнение называется функцией отклика.

Каждый фактор может принимать в опыте одно из нескольких значений. Такие значения называются уровнями.

При планировании эксперимента объект исследования обязательно должен обладать следующими двумя свойствами:

1. Воспроизводимостью результатов. При повторении эксперимента через неравные промежутки времени при одних и тех же значениях факторов разброс значений параметра оптимизации не должен превышать некоторой заранее заданной величины (наших требований к точности эксперимента).

2. Факторы должны быть управляемыми.

Планирование эксперимента предполагает активное вмешательство в процесс и возможность выбора в каждом опыте тех уровней факторов, которые представляют интерес. Поэтому такой эксперимент называется активным или алгоритмизированным.

Если один или несколько факторов неуправляемые, то воспроизводимость результатов, как правило, отсутствует и следует обращаться к активно-пассивному эксперименту.

В том случае, если все факторы неуправляемые, то эксперимент относится к категории пассивного эксперимента.

В некоторых случаях отсутствие воспроизводимости результатов объясняется действием фактора, систематически изменяющегося (дрейфующего) во времени. Тогда нужно обращаться к специальным методам исследования.

Последние три случая мы не будем рассматривать. Наша цель – рассмотрение методов планирования экстремального эксперимента для статистически воспроизводимых объектов.

*Резюме.* Результаты эксперимента используются для получения математической модели объекта исследования, которая представляет собой уравнение, связывающее пара-

метр оптимизации и факторы. Такое уравнение называется функцией отклика.

Использование для получения модели всех возможных опытов приводит к абсурдно большим экспериментам. Задача выбора необходимых для эксперимента опытов, методов математической обработки их результатов и принятия решений – это и есть задача планирования эксперимента. Частный случай этой задачи – планирование экстремального эксперимента, т. е. эксперимента, поставленного с целью поиска оптимальных условий функционирования объекта. Планирование экстремального эксперимента – метод выбора минимального количества опытов, необходимых для отыскания оптимальных условий.

## **5.2 Постановка и проведение эксперимента**

Постановка и проведение активного эксперимента включает ряд обязательных этапов:

1. Выбор объекта и формулировка цели исследования
2. Выбор параметра оптимизации
3. Определение факторного пространства, установление уровней и интервала варьирования факторов (выбор условий проведения опытов)
4. Анализ априорной информации и выбор плана эксперимента
5. Проведение опытов
6. Обработка полученных результатов и построение математической модели
7. Интерпретация результатов и принятие решений после построения модели

Выбор объекта и формулировка цели исследования нами не рассматриваются.

Проанализируем последовательно все остальные этапы.

### Выбор параметра оптимизации

В качестве параметров оптимизации могут использоваться агротехнические, эксплуатационные, технологические, экономические или иные оценочные показатели.

Параметр оптимизации – это признак, по которому оптимизируется процесс. Он должен быть количественным, задаваться числом. Множество значений, которые может принимать параметр оптимизации, называется областью его определения.

Параметр оптимизации должен быть:

- эффективным с точки зрения достижения цели;
- универсальным;
- количественным и выражаться одним числом;
- статистически эффективным;
- имеющим физическим смысл, простым и легко вычисляемым;
- существующим для всех различных состояний объекта.

Заданному набору значений факторов должно соответствовать одно с точностью эксперимента значение параметра оптимизации. (Однако обратное положение неверно: одному и тому же значению параметра могут соответствовать разные наборы значений факторов).

Задачи с одним выходным параметром имеют очевидные преимущества. Но на практике часто приходится учитывать несколько выходных параметров. Математические модели можно построить для каждого из параметров, но одновременно оптимизировать несколько функций невозможно.

Обычно оптимизируется одна функция, наиболее важная с точки зрения цели исследования, при ограничениях, налагаемых другими функциями. Поэтому из многих выходных параметров выбирается один в качестве параметра оптимизации, а остальные служат ограничениями.

Определение факторного пространства, установление уровней и интервала варьирования факторов (выбор условий проведения опытов).

Фактор – измеряемая переменная величина, принимающая в некоторый момент времени определенное значение. Факторы соответствуют способам воздействия на объект исследования.

Каждый фактор имеет область определения. Под областью определения понимается совокупность всех значений, которые может принимать данный фактор. Совокупность значений фактора, которая используется в эксперименте, является подмножеством из множества значений, образующих область определения.

Область определения может быть непрерывной и дискретной. В рассматриваемом типе задач всегда используются дискретные области определения.

Если число факторов больше пятнадцати, нужно использовать методы отсеивания несущественных факторов.

Факторы должны быть управляемы. Управлять фактором – это значит установить нужное значение и поддерживать его постоянным в течение опыта или менять по заданной программе. Кроме того, факторы должны быть однозначны. Трудно управлять фактором, который является функцией других факторов. Но в планировании могут участвовать сложные факторы, такие, как соотношения между компонентами, их логарифмы и т. п.

При планировании эксперимента обычно одновременно изменяется несколько факторов. Поэтому очень важны требования, которые предъявляются к совокупности факторов. Во-первых, факторы должны быть совместимы. Совместимость факторов означает, что все их комбинации осуществимы и безопасны. Во-вторых, факторы должны быть независимы, т. е. любой фактор должен устанавливаться на любом уровне вне зависимости от уровней других факто-

ров. Если это условие невыполнимо, то невозможно планировать эксперимент.

Выбранное множество факторов должно быть достаточно полным. Если какой-либо существенный фактор пропущен, это приведет к неправильному определению оптимальных условий или к большой ошибке опыта. Факторы могут быть как количественными, так и качественными.

### Анализ априорной информации и выбор плана эксперимента

Как правило, оптимизация начинается в условиях, когда объект уже подвергался некоторым исследованиям. Или же имеется информация об исследованиях аналогичных объектов. Такая информация называется априорной (т. е. полученной до начала эксперимента).

Априорная информация анализируется для получения представления о параметре оптимизации, о факторах, о наилучших условиях ведения процесса и характере поверхности отклика.

Следовательно, выбор экспериментальной области факторного пространства связан с тщательным анализом априорной информации.

По каждому фактору вначале выбирается *основной уровень*. Он устанавливается либо в окрестности лучшей точки (если она известна), либо выбирается произвольно и уточняется в ходе проведения опытов.

Далее для каждого фактора выбирается верхний и нижний уровни, на которых он будет варьироваться в эксперименте. Обычно за верхний уровень принимается тот, который соответствует большему значению фактора, хотя это и не обязательно. Разность между верхним и основным уровнями называется интервалом варьирования фактора. Таким образом, нижний уровень определяется вычитанием интервала варьирования из основного уровня.

Величина интервала варьирования фактора должна превышать погрешность его измерения.

Для упрощения записи условий эксперимента и обработки экспериментальных данных масштабы осей выбираются так, чтобы верхний уровень соответствовал +1, нижний –1, а основной – нулю. Для перевода факторов из натуральных единиц в кодированные значения используется следующее выражение

$$X_j = \frac{\tilde{X}_j - \tilde{X}_{j0}}{I_j}. \quad (5.2)$$

где  $X_j$  – кодированное значение фактора;

$\tilde{X}_j$  – натуральное значение фактора;

$\tilde{X}_{j0}$  – натуральное значение основного уровня;

$I$  – интервал варьирования;

$j$  – номер фактора.

Выбор плана эксперимента зависит от выбора математической модели. Выбрать модель – означает выбрать вид функции отклика (см. выражение 5.1). Тогда остается спланировать и провести эксперимент для оценки численных значений коэффициентов этого уравнения.

Для рассматриваемого класса задач используются алгебраические полиномы первой или второй степени. Для трехфакторного эксперимента они имеют вид

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3. \quad (5.3)$$

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_{12} X_1 X_2 + b_{13} X_1 X_3 + b_{23} X_2 X_3 + b_{11} X_1^2 + b_{22} X_2^2 + b_{33} X_3^2. \quad (5.4)$$

где  $Y$  – отклик;

$X_1, X_2, X_3$  – управляемые факторы;

$b_0$  – свободный член;

$b_1, b_2, b_3$  – коэффициенты при факторах первой степени;

$b_{11}, b_{22}, b_{33}$  – коэффициенты при факторах второй степени;

$b_{12}$ ,  $b_{13}$ ,  $b_{23}$  - коэффициенты эффектов взаимодействия факторов.

Выбор типа плана зависит от выбранной математической модели, числа выбранных факторов и цели исследования.

Рассмотрим некоторые планы для трех факторов.

1. Полнофакторный план  $3^3$  – традиционный план технико-экономического эксперимента, получаемый при исследовании воздействия отдельно каждого из трех факторов на трех уровнях. Единственный недостаток – максимально возможное число опытов  $N=27$ , т. е. перерасход ресурса.

2. План Хартли  $H_{a3}$  имеет всего 11 точек опытов, но плохие статистические характеристики. Крупный недостаток – высокая корреляция между коэффициентами регрессии при членах первого порядка и эффектов взаимодействий.

3. План Рехтшафнера имеет минимальное число точек 10, но имеет плохие статистические характеристики.

4. Ротатабельный план Бокса РКЦП-3 позволяет предсказывать значения отклика с равной точностью во всем исследуемом факторном пространстве. Но шесть точек этого плана необходимо брать с интервалом варьирования большим чем единица, что создает дополнительные трудности при его реализации. Кроме того, число опытов достаточно велико -  $N=20$ .

5. Один из лучших планов для трех факторов – план Бокса  $B^3$ .

Матрица планирования для плана второго порядка Бокса  $B_3$  приведена в таблице 5.1.

Здесь следует отметить, что все приведенные выше сравнительные оценки относятся к планам только для трех факторов. Для другого числа факторов эти оценки могут измениться. Так, например, при пяти факторах план Хартли  $H_{a5}$  будет относиться уже к числу лучших.

Таблица 5.1 Матрица планирования плана Бокса  $B_3$  и результаты опытов

№ опыта	Порядок проведения опытов			Матрица планирования			Повторности			Среднее значение $\bar{Y}$	Расчетное значение $\hat{Y}$
				$X_1$	$X_2$	$X_3$	$Y_1$	$Y_2$	$Y_3$		
1	4	12	17	+1	+1	+1					
2	15	22	39	-1	+1	+1					
3	31	5	26	+1	-1	+1					
4	28	36	2	-1	-1	+1					
5	6	19	33	+1	+1	-1					
6	21	41	13	-1	+1	-1					
7	10	29	40	+1	-1	-1					
8	24	35	9	-1	-1	-1					
9	1	38	32	+1	0	0					
10	30	20	16	-1	0	0					
11	27	8	42	0	+1	0					
12	11	37	3	0	-1	0					
13	18	25	23	0	0	+1					
14	7	34	14	0	0	-1					

### Проведение опытов

При проведении опытов очень важно избежать систематических ошибок. Это достигается путем рендомизации опытов, т. е. проведением их в случайном порядке – см. таблицу 5.1. Для определения последовательности проведения опытов можно использовать таблицы или генераторы случайных чисел.

Кроме того, исключительно важно, чтобы измерения, как факторов, так и параметра оптимизации в течение опыта проводились с одинаковой точностью.

Последние два этапа, посвященные обработке и интерпретации полученных результатов рассмотрим более подробно

### 5.3 Обработка полученных результатов и построение математической модели

При обработке матриц планирования экстремального эксперимента проводится дисперсионный и регрессионный анализ. При этом определяются следующие статистические оценки: ошибка параллельных опытов каждой серии, дисперсия параметра оптимизации, коэффициенты регрессии, адекватность полученных моделей.

Ошибка параллельных опытов одной серии оценивается по критерию Стьюдента

$$t_p = \frac{Y_i - \bar{Y}}{S} \langle t_T, \quad (5.5)$$

где  $t_p$  – расчетное значение критерия Стьюдента;

$Y_i$  – результат  $i$ -го опыта;

$\bar{Y}$  – среднеарифметическое значение параметра оптимизации;

$S$  – среднее квадратическое отклонение;

$t_T$  – табличное значение критерия Стьюдента.

Однородность дисперсии полученных результатов в каждой точке плана эксперимента проверяется по критерию Кохрана. При однородности дисперсий дисперсия воспроизводимости вычисляется по выражению

$$S_v^2 = \frac{1}{N} \sum_{j=1}^N S_j^2, \quad (5.6)$$

где  $S_v^2$  – дисперсия воспроизводимости;

$N$  - число опытов;

$j$  - текущий номер опыта;

$S_j^2$  - дисперсия  $j$ -го опыта.

Коэффициенты регрессии определяются по выражению

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{Y}, \quad (5.7)$$

где  $\mathbf{b}$  - коэффициенты регрессии;

$\mathbf{X}$  - матрица независимых переменных;

$\mathbf{X}'$  - обратная матрица независимых переменных;

$\mathbf{Y}$  - вектор наблюдений.

Ошибка, с которой математическая модель описывает опытные данные, оценивается дисперсией адекватности

$$S_{ад}^2 = \frac{1}{N - m'} \sum_{j=1}^N (\bar{Y}_j - \hat{Y}_j)^2 \cdot n_j, \quad (5.8)$$

где  $S_{ад}^2$  - дисперсия адекватности;

$\bar{Y}_j$  - среднее арифметическое значение параметра оптимизации в  $j$ -ом опыте;

$\hat{Y}_j$  - расчетное значение параметра оптимизации для условий  $j$ -го опыта;

$n_j$  - число параллельных опытов в отдельных точках плана эксперимента;

$m'$  - число коэффициентов в уравнении регрессии.

Степень свободы дисперсии адекватности равна

$$K_{ад} = N - m', \quad (5.9)$$

Адекватность полученных математических моделей оценивается критерием Фишера

$$F_p = \frac{S_{ад}^2}{S_B^2}, \quad (5.10)$$

где  $F_p$  - величина, имеющая  $F$ -распределение со степенями свободы

$$K_1 = K_{ад} = N - m' \text{ и } K_2 = K_B = \sum_{j=1}^n n_j - N \text{ при верной нуле-}$$

вой гипотезе.

Если  $F_{кр} = F_{\alpha, K_1, K_2} \geq F_p$ , то проверяемая модель считается адекватной, если  $F_p \geq F_{кр}$  - неадекватной.

Для проверки значимости коэффициентов при членах уравнения рассчитывается дисперсия коэффициентов регрессии

$$S_{vj}^2 = \frac{S_B^2}{N}, \quad (5.11)$$

где  $S_{b_j}^2$  - дисперсия коэффициентов регрессии.

Доверительный интервал определяется по выражению

$$\Delta b_j = \pm t S_{b_j}, \quad (5.12)$$

где  $t$  - табличное значение критерия Стьюдента при числе степеней свободы, с которым определялась дисперсия воспроизводимости при выбранном уровне значимости;

$S_{b_j}$  – квадратичная ошибка коэффициента регрессии.

Коэффициент регрессии считается значимым, если его абсолютная величина больше доверительного интервала.

После получения адекватной математической модели находятся координаты особой точки  $S$ . Координаты этой точки определяют оптимальные геометрические параметры и кинематический режим работы разрабатываемых технических средств.

Для определения значений факторов, при которых  $Y_1$  имеет экстремальное значение (особая точка  $S$ ), берутся частные производные по  $X_i$ . Решением полученной системы уравнений являются текущие значения факторов  $X_i$ , при которых  $Y_1$  достигает экстремального значения.

Для определения характера поверхности отклика вблизи особой точки  $S$  строятся двумерные сечения поверхности отклика или линии равного выхода – рисунок 5.2.

На рисунке приведены линии равного выхода, полученные при оптимизации параметров селекционной кукурузокалибровочной машины.

Сечения строятся методом подстановки значений факторов  $X_1$ ,  $X_2$  и  $X_3$  в функцию отклика (5.4) при фиксированных значениях  $Y$ . Так как функция (5.4) трехмерна, то при подстановке один из факторов необходимо фиксировать на уровне, соответствующем его оптимальному значению.

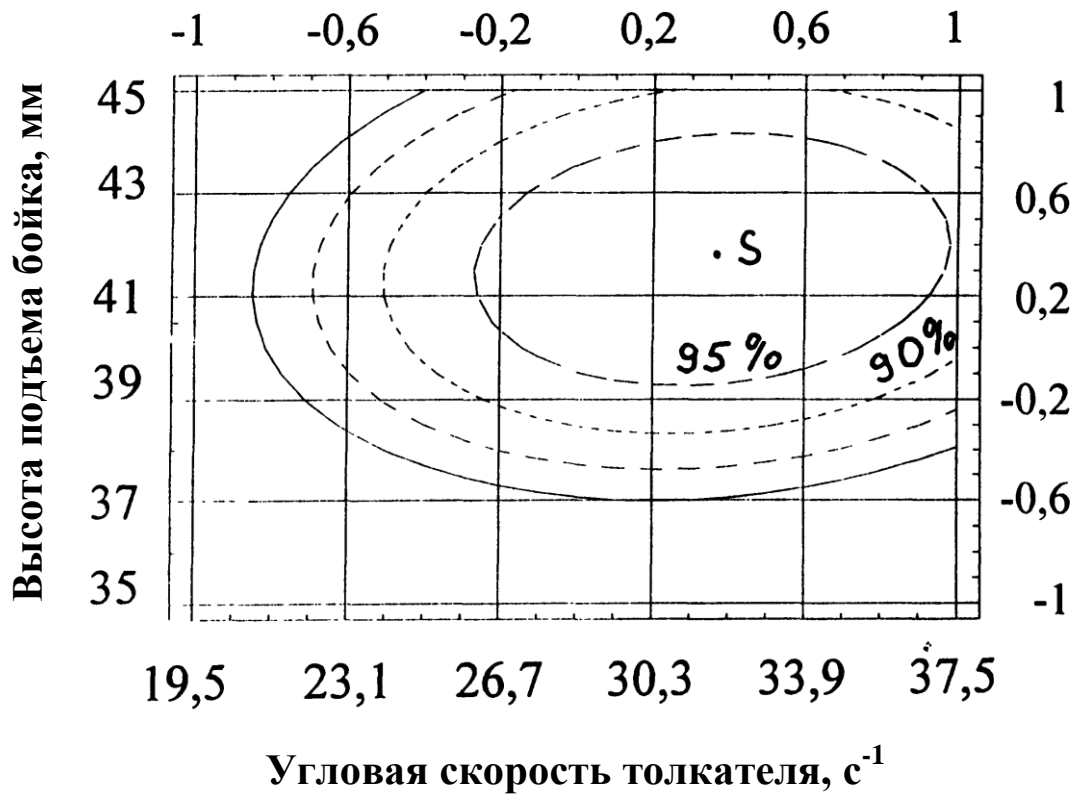


Рисунок 5.2 Линии равного выхода семян гибрида кукурузы Росс 149 МВ при сепарации на решетке с отверстиями 8 мм.

Пример проведения спланированного (алгоритмизированного) эксперимента приведен в приложении.