

饱和 D 最优设计方法在农业试验中的应用

张永成

(青海省农科院作物所 西宁 810016)

在农业试验中,饱和 D 最优设计是一种较好的设计方案。其特点是预测值精度较高,对选择最佳生产措施有较强的实用性。饱和 D 最优设计是回归方程中参数数目与试验处理组合数目相等的设计,回归自由度与总自由度相等达到饱和。采用这种设计试验处理数少,获得的信息量大,误差小,精确度高。

1 试验设计

1.1 选择试验因素,确定上下限水平

本例为 ABT5 号处理马铃薯的 2 因素试验,分别用 Z_1 和 Z_2 表示(表 1)。

表 1 因素水平

试 验 因 素	水 平	
	上 限	下 限
Z_1 浓度(ppm)	50	6
Z_2 浸种时间(h)	2	0.5

1.2 求试验因素的零水平和变化区间

根据饱和设计的要求规定:

某因素的零水平(Z_{0j})=

该因素的(上限水平+下限水平)
2

某因素的变化区间(Δ_j)=

该因素的(上限水平-零水平)
1

故①两因素的零水平分别为:

浓 度: $Z_{01}=\frac{50+6}{2}=28$

浸种时间: $Z_{02}=\frac{2+0.5}{2}=1.25$

②两因素的变化区间分别为:

浓 度: $\Delta_1=\frac{50-28}{1}=22$

浸种时间: $\Delta_2=\frac{2-1.25}{1}=0.75$

1.3 对各因素的水平进行编码

把各因素有量纲的自然变量 Z_j 变成无量纲的规范变量 x_j 。将浓度编码值用 x_1 表示,浸种时间编码值用 x_2 表示,进行线性变换,其公式为:

编码值 $x_j=\frac{Z_j-Z_{0j}}{\Delta_j}$

故: $Z_j=x_j\cdot\Delta_j+Z_{0j}$

根据以上已知数据,对浓度和浸种时间分别进行编码,求出各自的编码值 x_i 与自然变量 Z_i 的对应值,见表 2 和表 3。

表 2 浓度编码值对应的自然变量

x_1	Z_1 (浓度)
1	$Z_{11}=x_1\cdot\Delta_1+Z_{01}=1\times22+28=50$
u	$Z_{u1}=0.3944\times22+28\approx37.5$
λ	$Z_{\lambda1}=-0.1315\times22+28=25$
-1	$Z_{-11}=-1\times22+28=6$

注: $u=0.3944$ $\lambda=-0.1315$

表 3 浸种时间编码值对应的自然变量

x_2	Z_2 (浸种时间)
1	$Z_{12}=x_2 \cdot \Delta_2+Z_{02}=1 \times 0.75+1.25=2$
u	$Z_{u2}=0.3944 \times 0.75+1.25=1.5$
λ	$Z_{\lambda 2}=-0.1315 \times 0.75+1.25 \approx 1$
-1	$Z_{-12}=-1 \times 0.75+1.25=0.5$

注:u=0.3944 λ=-0.1315

1.4 试验方案

根据以上计算列出试验方案表,见表 4。

表 4 二因素二次饱和 D 最优设计实施方案

试验号	浓度(ppm)		浸种时间(h)	
1	—	6	-1	0.5
2	1	50	-1	0.5
3	-1	6	1	2
4	-0.1315	25	-0.1315	1
5	1	50	0.3944	1.5
6	0.3944	37.5	1	2

表 5 最优设计矩阵及结果

试 验 号	结 构 矩 阵						试验结果	
	x_0	x_1	x_2	x_1^2	x_2^2	x_1x_2	y	$\hat{y}$
1	1	-1(6)	-1(0.5)	1	1	1	126.35	126.87
2	1	1(50)	-1(0.5)	1	1	-1	129.35	128.93
3	1	-1(6)	1(2)	1	1	-1	133.40	132.97
4	1	λ(25)	λ(1)	0.0173	0.0173	0.0173	135.15	134.88
5	1	1(50)	u(1.5)	1	0.1556	0.3944	126.33	125.42
6	1	u(37.5)	1(2)	0.1556	1	0.3944	126.20	125.30
b	776.78	27.93	35.95	537.4	537.3	-34.46	R=0.99159**	

注:λ=-0.1315 u=0.3944 -1

$$C=[x' \ x]^{-1}$$
$$= \left[\begin{array}{cccccc} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -0.1315 & 1 & 0.3944 \\ -1 & -1 & 1 & -0.1315 & 0.3944 & 1 \\ 1 & 1 & 1 & 0.0173 & 1 & 0.1556 \\ 1 & 1 & 1 & 0.0173 & 0.1556 & 1 \\ 1 & -1 & -1 & 0.0173 & 0.3944 & 0.3944 \end{array} \right] \left[\begin{array}{cccccc} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -0.1315 & -0.1315 & 0.0173 & 0.0173 & 0.0173 \\ 1 & 1 & 0.3944 & 1 & 0.1556 & 0.3944 \\ 1 & 0.3944 & 1 & 0.1556 & 1 & 0.3944 \end{array} \right]^{-1}$$

针对

2 试验结果及分析

2.1 数学模型的建立

二次饱和 D 最优设计的回归模型为:

$$E(y)=B_0+\sum_{j=1}^pB_jx_j+\sum_{i<j}^pB_{ij}x_ix_j$$

(2-1)

本试验选择的是 2 因素 4 水平设计,其结构矩阵(x)如表 5。将各处理的株高填入试验结果栏中的 y 列。

表中的数字是编码值,括号内的数字是对应的自然变量。试验时,仅以括号内的数字实施(即实施方案)。y 值为实测值,y 为理论值。

计算相关矩阵 C

①解矩阵

$$C=[x' \ x]^{-1}$$

式中:C 表示相关矩阵,x 表示结构矩阵,x'为 x 的转置矩阵。

$$= \begin{pmatrix} 0.963 & -0.139 & -0.139 & -0.568 & -0.568 & -0.168 \\ -0.139 & 0.276 & 0.026 & 0.002 & 0.181 & 0.063 \\ -0.139 & 0.026 & 0.276 & 0.181 & 0.002 & 0.063 \\ -0.568 & 0.002 & 0.181 & 1.171 & -0.360 & 0.125 \\ -0.568 & 0.181 & 0.002 & -0.360 & 1.171 & 0.125 \\ -0.168 & 0.063 & 0.063 & 0.125 & 0.125 & 0.350 \end{pmatrix} \quad (2-2)$$

②求常数项矩阵 B

各列 B 值 = $\sum$ (该列每一编码值 $\times$ 对应的试验指标值)

$$\begin{aligned} \text{如, 第一列 } B_0 &= \sum x_{0y} \\ &= 1 \times 126.35 + 1 \times 129.35 + 1 \times 133.40 + \\ &1 \times 135.15 + 1 \times 126.33 + 1 \times 126.20 \\ &= 776.78 \end{aligned}$$

$$\text{同理: } B_1 = \sum x_{1y}$$

$$B_2 = \sum x_{2y}$$

$$B_{11} = \sum x_1^2 y$$

$$B_{22} = \sum x_2^2 y$$

$$B_{12} = \sum x_1 x_2 y$$

即:

$$B = \begin{pmatrix} 776.78 \\ 27.93 \\ 35.95 \\ 537.40 \\ 537.30 \\ -34.46 \end{pmatrix} \quad (2-3)$$

③计算回归系数 b

$$b = C \cdot B$$

$$= \begin{pmatrix} 0.963 & -0.139 & -0.139 & -0.568 & -0.568 & -0.168 \\ -0.139 & 0.276 & 0.026 & 0.002 & 0.181 & 0.063 \\ -0.139 & 0.026 & 0.276 & 0.181 & 0.002 & 0.063 \\ -0.568 & 0.002 & 0.181 & 1.171 & -0.360 & 0.125 \\ -0.568 & 0.181 & 0.002 & -0.360 & 1.171 & 0.125 \\ -0.168 & 0.063 & 0.063 & 0.125 & 0.125 & 0.350 \end{pmatrix} \times \begin{pmatrix} 776.78 \\ 27.93 \\ 35.95 \\ 537.40 \\ 537.30 \\ -34.46 \end{pmatrix} = \begin{pmatrix} 134.52 \\ -3.17 \\ -1.15 \\ -3.09 \\ -4.68 \\ -4.20 \end{pmatrix} \quad \begin{matrix} b_0 \\ b_1 \\ b_2 \\ b_{11} \\ b_{22} \\ b_{12} \end{matrix} \quad (2-4)$$

④建立数学模型

$$\hat{y} = 134.52 - 3.17x_1 - 1.15x_2 - 3.09x_1^2 - 4.68x_2^2 - 4.2x_1x_2 \quad (2-5)$$

⑤计算理论值 $\hat{y}$

把结构矩阵中的编码值分别代入回归方程,即可求得每处理的理论估计值,结果如下:

$$\begin{aligned} \hat{y}_1 &= 134.52 - 1 \times (-3.17) - 1 \times (-1.15) + 1 \times (-3.09) + 1 \times (-4.68) + 1 \times (-4.2) = 126.87 \end{aligned}$$

$$\text{余类推: } \hat{y}_2 = 128.93 \quad \hat{y}_3 = 132.97$$

$$\hat{y}_4 = 134.88 \quad \hat{y}_5 = 125.42$$

$$\hat{y}_6 = 125.30$$

将上述 y 值填入表 5 中的最后一列。

2.2 模型检验

饱和最优设计的总自由度与回归自由度相等,没有剩余自由度,故不能估计误差。对于试验结果的精度只能用测定值和估计值之间的相关系数来验证。

$R = 0.99159^{**}$ 达到极显著水准,说明 y 与 $\hat{y}$ 非常吻合。亦证明饱和 D-最优设计是回归设计中比较好的一种设计方法。

2.3 模型分析

2.3.1 因子主次分析

偏回归系数绝对值的大小反映了因素的主次,符号的正负反映了因素水平的取值方向。由(2-5)式看出, $|3.17| > |1.15|$ 则 $b_1 > b_2$, 表明在一定的范围内浓度使用量的大小是影响植株高度的主要因素。

2.3.2 主效因子分析

令(2-5)式中 $x_2=0$ 则

$\hat{y}=134.52-3.17x_1-3.09x_1^2$ (2-6)

分别用-1、-0.1315、0.3944、1 四个编码值代入(2-6)式中的 x_1 , 即可求得不同用药量对株高的影响。

同理, 令(2-5)式中的 $x_1=0$ 则

$\hat{y}=134.52-1.15x_2-4.68x_2^2$ (2-7)

用上述 4 个编码值代入(2-6)式中的 x_2 , 即求出不同浸种时间对株高的影响。见表 6。

表 6 两因素取不同编码值时
对株高的影响

因素	编码值(x_j)				极差 (K)
	-1	-0.1315	0.3944	1	
浓度	134.60	134.88	132.79	128.26	6.62
浸种时间	130.99	134.59	133.34	128.69	5.9

2.3.3 交互效应分析

将 4 个编码值分别代入(2-5)式的 x_1 和 x_2 中求出 x_1x_2 的交互效应值, 见表 7。由表 7 可知, 当浓度编码值取-1 水平时, 浸种时间的编码值取 0.3944 时, 交互效应最高, 理论株高达到 135.07cm。

表 7 浓度与浸种时间的交互效应

x_2	x_1			
	-1	-0.1315	0.3944	1
-1	126.87	130.80	130.92	128.93
-0.1315	134.12	134.88	133.08	128.88
0.3944	135.07	133.92	130.95	125.42
1	132.97	129.61	125.30	118.23

2.3.4 边际效应分析

分析边际效应就是分析株高随各因素水平值变化而增减的变化率。

用(2-5)式中的 $\hat{y}$ 分别对 x_1 和 x_2 求偏导数, 求出边际效应方程:

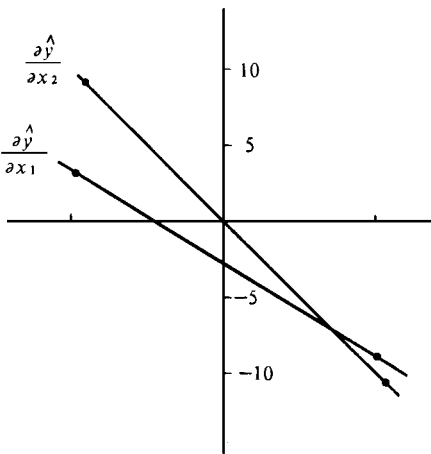
$\frac{\partial \hat{y}}{\partial x_1}=-3.17-6.18x_1$ (2-8)

$\frac{\partial \hat{y}}{\partial x_2}=-1.15-9.36x_1$ (2-9)

对上述边际效应方程分别用-1 和 1(编码值的两个极值)代入, 每式求出两个点, 在直角坐标系中将两个点连成一条直线, 见表 8 和图。

表 8 边际效应值

因素	编码	
	-1	1
$\frac{\partial y}{\partial x_1}$	3.01	-9.35
$\frac{\partial y}{\partial x_2}$	8.21	-10.51



边际效应图

由图可以看出, $\frac{\partial \hat{y}}{\partial x_2}$ 直线的斜率大于 $\frac{\partial \hat{y}}{\partial x_1}$ 直线的斜率, 说明侵种时间对植株高度影响速度快。

2.4 最优农艺措施的选择

2.4.1 模拟试验

用 4 个编码值, 每两两组合, 共组成 $4^2=16$ 个全因子试验, 分别代入(2-5)式中的 x_1 和 x_2 中, 即可模拟出 16 个试验结果(表 9), 其中包括已经实施的 6 个试验在内。如将第一个组合(-1, -1)分别代入 x_1 和 x_2 , 其结果如下:

$\hat{y}=134.52-3.17\times(-1)-1.15\times(-1)-3.09\times(-1)^2-4.68\times(-1)^2-4.2\times(-1)\times(-1)=126.87$
余类推。

2.4.2 定指标约束条件,选择最优组合

根据生产实践定出优选指标的约束条件,本列凡株高在 130cm 以上的组合,均为最优农艺措施组合,从 16 个组合中选出满足约束条件的 9 个组合(见表 9 中画“√”者)。若是多指标试验时,所选的组合必须同时满足各个指标的约束条件。

2.4.3 对入选的 9 个组合进行综合分析,确定最优农艺措施组合

列出最优组合的频数分布表(表 10)进行分析。

对入选的 9 个最优组合进行平均数、标准差、标准误、95%的置信限以及农艺措施的计算,最后得出在马铃薯增产方面采用不同浓度和不同浸种时间应用 ABT 生根粉,对株

高的影响程度表明:浓度在 12.78 ~ 32.41ppm, 浸种时间在 0.86~1.60h 范围

表 9 模拟试验结果

编号	组编码		株高 (cm)	优选 组合
	x1	x2		
1	-1	-1	126.87	
2	-1	-0.1315	134.12	√
3	-1	0.3944	135.07	√
4	-1	1	132.97	√
5	-0.1315	-1	130.80	√
6	-0.1315	-0.1315	134.88	√
7	-0.1315	0.3944	133.92	√
8	-0.1315	1	129.61	
9	0.3944	-1	130.92	√
10	0.3944	-0.1315	133.08	√
11	0.3944	0.3944	130.95	√
12	0.3944	1	125.30	
13	1	-1	128.93	
14	1	-0.1315	128.88	
15	1	0.3944	125.42	
16	1	1	118.23	

内效果为佳。

表 10 最优组合频数分布

编码值及 统计项目	浓 度		浸种时间	
	次数	频率	次数	频率
-1	3	0.333	2	0.222
-0.1315	3	0.333	3	0.333
0.3944	3	0.333	3	0.333
1			1	0.111
合 计	9	1	9	1
平均值 $\bar{x}_j$	-0.2457		-0.0235	
标准差 S_j	0.5916		0.6627	
标准误 $S\bar{x}_j$	0.1972		0.2209	
95%置信限 L	-0.6918~0.2004		-0.5232~0.4762	
农艺措施	12.78~32.41 ppm		0.86~1.60 h	

注:自由度 df=9 t_{0.05}=2.262

优良农艺措施组合选择的详细求算过程:

①计算平均数 $\bar{x}_i$

$$\text{浓度 } \bar{x}_1 = [(-1) \times 3 + (-0.1315) \times 3 + 0.3944 \times 3 + 1 \times 0] \div 9 = -0.2457$$

$$\text{浸种时间 } \bar{x}_2 = [(-1) \times 2 + (-0.1315) \times 3 + 0.3944 \times 3 + 1 \times 1] \div 9 = -0.0235$$

②计算标准差

可用 $f_x - 3600_p$ 直接计算标准差:

$$\text{浓度 } S = 0.5916 \quad \text{浸种时间 } S = 0.6627$$

③计算频率

$$\text{频率} = \frac{\text{各编码值的次数}}{\text{总数 } N}$$

浓度的频率:

$$-1 \text{ 频率} = \frac{3}{9} = -0.333$$

$$-(0.1315) \text{ 频率} = \frac{3}{9} = 0.333$$

$$0.3944 \text{ 频率} = \frac{3}{9} = 0.333$$

$$1 \text{ 的频率} = \frac{0}{9} = 0$$

浸种时间频率计算同上。

④计算标准误 $S\bar{x}_i$

$$\text{标准误} = \frac{\text{标准差 } S}{\sqrt{n}} = \frac{S}{\sqrt{9}} = \frac{S}{3}$$

$$\text{浓度 } S\bar{x}_1 = \frac{0.5916}{3} = 0.1972$$

$$\text{浸种时间 } S\bar{x}_2 = \frac{0.6627}{3} = 0.2209$$

⑤计算 95% 的置信限

95% 的置信限 = 样本平均数 $\pm t_{0.05} \times$ 标准误。当自由度 $df = 9$ 时, $t_{0.05} = 2.262$

浓度 95% 置信限

$$= -0.2457 \pm 2.262 \times 0.1972$$

$$= (-0.6918) - 0.2004$$

浸种时间 95% 置信限

$$= -0.0235 \pm 2.262 \times 0.2209$$

$$= (-0.5232) - 0.4762$$

⑥农艺措施求算

将 95% 的置信限编码值换算成自然变

量:

$$\text{因为 } x_i = \frac{Z_i - Z_{0i}}{\Delta_i};$$

$$\text{所以 } Z_i = Z_{0i} + x_i \cdot \Delta_i$$

即: 自然变量 = 零水平 + 编码值 $\times$ 变化

区间

$$\text{浓度: } Z_{01} = 28 \quad \Delta_1 = 22$$

$$x_{1下} = 28 + (-0.6918) \times 22 =$$

$$12.78 \quad (\text{ppm})$$

$$x_{1上} = 28 + 0.2004 \times 22 =$$

$$32.41 \quad (\text{ppm})$$

$$\bar{x}_1 = \frac{1}{2} (12.78 + 32.41) =$$

$$22.595 \quad (\text{ppm})$$

$$\text{浸种时间: } Z_{02} = 1.25 \quad \Delta_2 = 0.75$$

$$x_{2下} = 1.25 + (-0.5232) \times 0.75 =$$

$$0.8576 \quad (\text{h})$$

$$x_{2上} = 1.25 + 0.4762 \times 0.75 =$$

$$1.6015 \quad (\text{h})$$

$$\bar{x}_2 = \frac{1}{2} (0.8576 + 1.6015) = 1.23 \quad (\text{h})$$

参 考 文 献

- 1 张中发等. 二次回归正交组合设计. 青海农林科技, 1989, 2
- 2 张中发等. 农业现代试验设计. 青海农林科技, 1989, 3
- 3 张永成. ABT 生根粉在马铃薯生产中的应用. 马铃薯杂志, 1991, 5(4): 219~222