

# 基于主成分分析法的温室甜瓜生长方程研究

谢恒星<sup>1</sup>, 蔡焕杰<sup>1\*</sup>, 张振华<sup>2</sup>, 王加蓬<sup>1</sup>

(1. 西北农林科技大学教育部旱区农业水土工程重点实验室, 陕西 杨凌 712100;

2. 鲁东大学地理与规划学院, 山东 烟台 264025)

**摘要:** 利用主成分分析法拟合了不同水分下限控制条件下温室甜瓜生长过程, 并对该方法的通用性进行了验证。研究表明, 利用表征作物长势的综合植株性状与有效积温建立 Logistic 经验模型, 模型的决定系数均在 0.9 以上, 模型的拟合精度较高。利用该模型拟合第 3 试验小区的甜瓜生长状况, 观测值与拟合值的相对误差大部分处于 10% 之内。主成分分析既充分保留原始变量的信息量, 又剔除了原变量之间的信息重叠部分, 该方法在作物生长方程模拟中具有很好的实用性。

**关键词:** 甜瓜; 主成分分析; 生长发育方程模型; 综合植株性状; 拟合精度

**中图分类号:** S652 **文献标识码:** A **文章编号:** 1000-7601(2009)06-0134-05

作物的生长发育不仅与遗传生理、生态环境有关, 还受农业气象、土壤肥料及耕作栽培技术等因子的影响和制约<sup>[1~3]</sup>。为了达到对如此复杂系统的研究和把握, 完全依赖物理试验是很难完成的, 要受到人力、财力、天气条件和生长周期等因素的制约。通过作物生长方程或模型的研究, 可以不同程度简化试验过程、加快研究进度、缩短研究周期、节省大量物力财力, 因此模型的研究引起了广泛的关注<sup>[4~6]</sup>。作物生长方程或模型按照其构建原理可以分为机理模型和经验模型, 机理模型具有严密的物理基础, 但由于建模过程中需要输入大量的参数和经过复杂的数学推导而不能被大多数人所理解。经验模型虽然受地域性和时间性的限制而具有一定的局限性, 但由于该模型的建模相对简单和便于理解而得到了广泛的应用<sup>[7,8]</sup>。作物生长经验模型的形式多种多样, 建模过程中一般以作物生长过程的天数或有效积温为自变量, 叶面积指数或干物质重为因变量<sup>[9,10]</sup>。但利用单一的叶面积指数或干物质重来表征作物的生长状况具有一定的片面性, 若利用多个指标如株高、茎粗、叶面积等来表征作物生长状况, 由于指标之间具有一定的相关性而带来信息的重叠, 从而使得作物长势的解释带有不合理的成分<sup>[11,12]</sup>。

主成分分析是把多个指标简化为少数几个综合指标的一种统计分析方法。它采取一种降维处理技

术, 找出几个综合因子来代表原来众多的变量, 这些综合因子尽可能地反映了原始变量的信息量, 而且彼此之间互不相关, 从而达到简化的目的<sup>[13]</sup>。利用主成分来表征作物生长状况, 既可以充分表达作物的长势信息, 又可以消除多个信息相互重叠使得长势的解释变得含混不清的缺点, 因此探讨该方法在作物生长发育方程建立中的应用具有一定的现实意义。

本文以温室甜瓜为例, 研究主成分分析法在温室作物生长发育过程模拟中的应用, 以期作为作物生长模型的建立奠定更为科学、合理的基础。

## 1 材料与方法

### 1.1 试验地概况

试验于 2008 年 7~10 月在西北农林科技大学教育部旱区农业水土工程重点实验室温室大棚内进行。该地地处关中平原中部, 地理坐标为 34°20'N, 108°04'E, 海拔 521 m, 地下水埋深 50 m。属暖温带季风半湿润气候区, 年均日照时数 2 163.8 h, 无霜期 210 d。土壤类型为壤土, 甜瓜定植前测得的土壤有机质含量为 1.95 g/kg, 全氮含量为 0.96 g/kg, 全磷含量为 1.23 g/kg, 全钾含量为 20.21 g/kg。0~100 cm 土层的平均土壤容重 1.42 g/cm<sup>3</sup>, 田间持水量为 24.1%。

### 1.2 试验设计

试验采用单因素完全随机区组设计, 4 个滴灌

收稿日期: 2009-05-20

基金项目: 国家自然科学基金项目 (50779059, 50609022)

作者简介: 谢恒星 (1981-), 男, 山东兖州人, 在读博士, 从事节水灌溉新技术研究。E-mail: xiehengxing@nwsuaf.edu.cn。

\* 通讯作者: 蔡焕杰 (1962-), 男, 河北藁城人, 教授, 博士生导师, 从事农业节水和水资源高效利用研究。

灌水下限处理(灌水下限分别为 T1:55% $\theta_f$ ,T2:66% $\theta_f$ ,T3:75% $\theta_f$ ,T4:85% $\theta_f$ 。灌水上限统一为 95% $\theta_f$ ,其中  $\theta_f$  为田间持水量),1 个沟灌对照 CK,各 3 次重复,其中的 2 个重复试验小区用来建模,另外的一个重复试验小区用来进行方程模型验证。沟灌按照常规经验灌水,即土壤含水率降低到田间持水量的 60%左右开始灌水,灌水上限为田间持水量。宽窄行种植,宽、窄行距分别为 0.9 m 和 0.6 m,株距 0.4 m。南北向做垄,垄高、宽分别为 0.2 m 和 0.5 m,每小区长 3.5 m,宽 3 m,为了防止水分的侧渗,小区之间用深 0.7 m 的塑料薄膜间隔。供试品种为“一品天下 208”,3 叶 1 心时定植在大棚内,单蔓整枝,20 节位打顶。甜瓜的其它管理措施均与大田甜瓜相同。观测内容包括株高( $x_1$ )、茎粗( $x_2$ )、叶片数( $x_3$ )、叶面积( $x_4$ )、叶绿素含量( $x_5$ )和叶面积指数( $x_6$ )等共 6 项指标,每指标选取长势均一的 4 株测定,平均值作为该项指标的观测值。株高采用米尺从甜瓜基部量取,茎粗采用游标卡尺量取基部茎秆直径;叶面积采用十字交叉法(由于摘除侧枝的原因,叶面积和叶片数均为主枝叶片的参数值),叶面积指数由叶面积仪测定<sup>[14]</sup>,气温由放置在温室内的蒸发蒸腾农田小气候温度传感器采集。试验数据利用 Micro-Excel 2003 和 DPS 6.55 处理分析,分析所用数据为甜瓜从苗期到打顶之间的植株特性参数(打顶后叶片数不再增加)。

## 2 结果与分析

### 2.1 主成分分析法建模步骤

主成分分析法的建模步骤大体上可以分为原始数据标准化,计算相关系数,计算特征值、特征向量,根据累计贡献率选取主成分,计算主成分载荷,构建主成分表达式等过程,每一过程的具体实施步骤如下<sup>[13]</sup>:

设观测样本原始数据矩阵为

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

其中,  $n$  为样本数,  $p$  为变量数。

(1) 原始数据标准化。由于各变量量纲往往不同,而不同量纲的数据不能放在一起进行比较,因此需要对原始数据进行标准化处理,以消除量纲的影响,使其具有可比性。数据标准化公式为

$$x_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k} \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, p)$$

(2)

$$\text{其中, } \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik}, S_k^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2.$$

(2) 计算样本矩阵的相关系数矩阵。计算公式为

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \quad (3)$$

(3) 计算特征值和特征向量。对应于相关系数矩阵  $R$ ,用雅可比方法求特征方程  $|R - \lambda I| = 0$  的  $p$  个非负的特征值,即

$$r_n \lambda^p + r_{n-1} \lambda^{p-1} + \cdots + r_1 \lambda + r_0 = 0 \quad (4)$$

的特征多项式,求  $\lambda_1, \lambda_2, \dots, \lambda_p$  并使  $\lambda_i$  按大小排列,即

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$$

对应于特征值  $\lambda_i$  的相应特征向量为

$$C^{(i)} = (C_1^{(i)}, C_2^{(i)}, \dots, C_p^{(i)}), i = 1, 2, \dots, p \quad (5)$$

并且满足

$$C^{(i)} C^{(j)} = \sum_{k=1}^p C_k^{(i)} C_k^{(j)} = \begin{cases} 1 (i = j) \\ 0 (i \neq j) \end{cases} \quad (6)$$

(4) 计算贡献率和累计贡献率。贡献率的计算公式为

$$CR = \lambda_i / \sum_{i=1}^p \lambda_i \quad (7)$$

累计贡献率的计算公式为

$$CRA = \sum_{k=1}^p \left[ \lambda_k / \sum_{i=1}^p \lambda_i \right] \quad (8)$$

主成分分析的基本思想就是选取尽量少的  $m$  个主成分来进行综合评价,同时还要使损失的信息量尽可能少。一般取累计贡献率达 85% 的特征值  $\lambda_1, \lambda_2, \dots, \lambda_m (m \leq p)$  对应的主成分作为主成分分析的个数。

(5) 计算主成分载荷。主成分载荷为主成分  $Z_k$  和变量  $x_i$  之间的相关系数,计算公式为

$$p(z_k \cdot x_i) = \sqrt{\lambda_k} l_{ki} \quad (i = 1, 2, \dots, p, k = 1, 2, \dots, m) \quad (9)$$

(6) 计算主成分得分,构造各主成分表达式。

### 2.2 甜瓜植株性状的主成分分析

以 T3 处理为例,说明主成分分析法在甜瓜生长模型建立中的应用。因各植株性状参数的单位不同,为了使其具有可比性,首先对原始数据进行标准化处理,以消除量纲的影响。设经标准化处理的数据记为  $X_n$ ,对各植株性状间进行相关性分析,相关系数矩阵如表 1。

表 1 经标准化处理后植株性状间相关系数矩阵

Table 1 Correlation coefficients matrix between standardized data of plant characteristics

植株性状 Plant characteristic	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
$X_1$	1					
$X_2$	0.989**	1				
$X_3$	1.000**	0.989**	1			
$X_4$	0.962**	0.979**	0.960**	1		
$X_5$	0.892*	0.933**	0.894*	0.868*	1	
$X_6$	0.989**	0.998**	0.988**	0.977**	0.936**	1

注: \*\*表示相关性在 0.01 水平达到显著, \*表示相关性在 0.05 水平达显著。

Note: \*\* means significant difference at 0.01 level, while \* means significant difference at 0.05 level.

由表 1 可知, 甜瓜植株性状间均具有很强的相关性,  $X_6$  与其它 5 个性状参数的相关性均达到 0.01 极显著水平,  $X_5$  与  $X_1$ 、 $X_3$  和  $X_4$  间的相关性均达到 0.05 显著水平, 因此利用单一的植株性状参数来表示植株生长状况不能完全代表植株长势信息而具有一定的片面性, 而利用多个原始指标来表示又会存在信息的重叠。

利用 DPS 数据处理软件中的主成分分析功能模块, 计算植株性状参数数据矩阵的特征值、特征向量、贡献率和累计贡献率等特征参数, 结果如表 2、表 3。

表 2 植株性状各主成分的特征值和百分率

Table 2 Eigenvalue and percentage of plant characteristics principal components

主成分 Principal component	特征值 Eigenvalue	百分率(%) Percentage	累计百分率(%) Cumulative percentage
1	5.788	96.460	96.460
2	0.159	2.655	99.114
3	0.051	0.854	99.968
4	0.002	0.031	99.999
5	0.000	0.001	100.000
6	0.000	0.000	100.000

由表 2 可知, 主成分 1 的累计贡献率达到 96.460%, 基本上保留了原来 6 个因子的全部信息, 根据 85% 累计贡献率原则, 选取第 1 主成分作为主成分分析的依据。

根据特征值和特征向量计算第 1 主成分的主成分载荷, 结果如表 4。

表 3 植株性状各主成分的特征向量

Table 3 Eigenvector of plant characteristics principal components

植株性状 Plant characteristic	主成分 1 Principal component 1	主成分 2 Principal component 2	主成分 3 Principal component 3	主成分 4 Principal component 4	主成分 5 Principal component 5	主成分 6 Principal component 6
$x_1$	0.412	-0.244	-0.441	-0.088	0.431	-0.619
$x_2$	0.415	-0.023	0.059	0.632	-0.596	-0.262
$x_3$	0.412	-0.231	-0.463	0.145	0.083	0.731
$x_4$	0.406	-0.337	0.764	0.028	0.357	0.103
$x_5$	0.389	0.879	0.051	0.041	0.262	0.053
$x_6$	0.415	-0.001	0.043	-0.754	-0.506	0.001

表 4 植株性状主成分载荷

Table 4 Principal component loading of plant characteristics

主成分 Principal component	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
主成分 1 Principal component 1	0.990	0.999	0.990	0.976	0.936	0.999

由主成分载荷可以构建主成分的表达式:

$$Z_{T3} = 0.990x_1 + 0.999x_2 + 0.990x_3 + 0.976x_4 + 0.936x_5 + 0.999x_6 \quad (10)$$

式中,  $Z_{T3}$  为综合植株性状, 其含义为原植株性状的线性加权组合, 其作用为既充分保留了原植株性状特征的信息量, 又去除了原多个性状特征的信息重复部分, 使得植株性状的表达更为合理和科学。

利用同样的方法, 可以得到另外 3 个滴灌处理和 1 个沟灌对照的甜瓜植株性状的主成分表达式, 其表达式分别为:

$$Z_{T1} = 0.983x_1 + 0.990x_2 + 0.998x_3 + 0.985x_4 + 0.979x_5 + 0.989x_6 \quad (11)$$

$$Z_{T2} = 0.990x_1 + 0.989x_2 + 0.990x_3 + 0.985x_4 + 0.982x_5 + 0.989x_6 \quad (12)$$

$$Z_{T4} = 0.990x_1 + 0.989x_2 + 0.994x_3 + 0.969x_4 + 0.984x_5 + 0.995x_6 \quad (13)$$

$$Z_{CK} = 0.990x_1 + 0.955x_2 + 0.993x_3 + 0.975x_4 + 0.975x_5 + 0.992x_6 \quad (14)$$

## 2.3 甜瓜植株生长模型

Logistic 植物生长模型是应用广泛的一种经验模型, 该模型的表达式为:

$$y = \frac{c_1}{1 + e^{c_2 + c_3x}} \quad (15)$$

式中,  $x$  为环境因子(生长天数或有效积温),  $y$  为植物生长性状(叶面积指数或干物质重),  $c_1$ 、 $c_2$  和  $c_3$  为待求参数。本例中,  $x$  取有效积温,  $y$  为甜瓜综合植株性状。甜瓜生育阶段的有效积温定义为<sup>[15]</sup>:

$$T_e = \sum(T_d - T_b) \quad (16)$$

式中,  $T_d$  为日平均气温(每日 2:00、8:00、14:00 和 20:00 的平均气温),  $T_b$  为甜瓜发育基点温度(苗期和开花坐果期均为 15°C, 伸蔓期为 13°C<sup>[16]</sup>)。以累计

有效积温为自变量, 分别以综合植株性状  $Z_n$  和叶面积指数  $x_6$  为因变量, 5 个处理的 logistic 植物生长模型参数值如表 5。

表 5 不同水分处理条件下植株生长的 Logistic 模型参数(A/B)

Table 5 The Logistic model parameters of plant development in different watering treatments

水分处理 Treatment	T1	T2	T3	T4	CK
C <sub>1</sub>	14.946/2.884	30.463/6.438	18.955/2.580	18.925/2.092	27.394/2.899
C <sub>2</sub>	6.548/3.920	4.008/4.356	4.553/4.302	4.646/4.250	4.088/4.301
C <sub>3</sub>	-0.029/-0.014	-0.013/-0.012	-0.018/-0.017	-0.019/-0.019	-0.014/-0.017
R <sub>2</sub>	0.999/0.993	0.995/0.968	0.992/0.998	0.988/0.987	0.996/0.989

注: A/B: A 代表由综合植株性状与有效积温建立的 Logistic 方程参数, B 代表由叶面积指数与有效积温建立的 Logistic 方程参数。

Note: A/B: A stands for Logistic formula parameters constructed with plant characteristics and effective accumulative temperature, while B stands for that with leaf area index and effective accumulative temperature.

由表 5 可知, 由综合植株性状和叶面积指数建立的 Logistic 经验模型均具有很高的拟合精度, 模型的决定系数均在 0.9 以上, 且与叶面积指数模型相比, 综合植株性状模型决定系数稍有提高, 但由单一的叶面积指数建立的生长模型具有一定的片面性, 不能很好地代表植株的全部生长信息, 所以由综合

植株性状建立的增长模型在拟合精度较高的条件下更具有一定的优越性。

为了检验该模型的通用性, 分别把第 3 个重复试验小区的甜瓜植株性状特征进行主成分分析, 将所得的综合植株性状带入式(11)~(14), 计算得到综合植株性状的拟合值, 观测值与拟合值的关系如表 6。

表 6 甜瓜植株综合性状预测误差

Table 6 The forecast error of muskmelon comprehensive plant characteristics

植株性状特征值 Plant characteristic value	日期 Data(m-d)					
	08-04	08-10	08-14	08-20	08-26	
T1	观测值 Measured value	0.124	2.832	6.658	11.695	14.206
	拟合值 Fitted value	0.338	2.974	6.421	11.928	14.082
	相对误差 Relative error	-172.581	-5.014	3.56	-1.992	0.873
T2	观测值 Measured value	1.448	4.623	7.473	11.178	16.481
	拟合值 Fitted value	1.796	4.649	6.939	11.622	16.347
	相对误差 Relative error	-24.033	-0.562	7.146	-3.972	0.813
T3	观测值 Measured value	0.713	3.767	7.374	10.902	15.17
	拟合值 Fitted value	1.083	4.022	6.655	11.476	14.954
	相对误差 Relative error	-51.893	-6.769	9.75	-5.265	1.424
T4	观测值 Measured value	0.877	3.475	7.676	11.09	15.412
	拟合值 Fitted value	1.04	4.029	6.753	11.711	15.186
	相对误差 Relative error	-18.586	-15.942	12.024	-5.6	1.466
CK	观测值 Measured value	1.586	4.088	7.249	11.124	16.031
	拟合值 Fitted value	1.614	4.423	6.727	11.407	15.954
	相对误差 Relative error	-1.778	-8.198	7.201	-2.539	0.484

由表 5 可知, 除了 8 月 4 日的甜瓜植株综合性状的相对误差较大外, 其它日期的相对误差均较小, 大部分的相对误差绝对值在 10% 之内, 说明该模型具有一定的通用性, 可以用来进行甜瓜生长方程的模拟。8 月 4 日甜瓜正处于苗期, 土壤比较松动, 株高的量取精度受到一定的影响, 且甜瓜的茎秆较为

脆弱, 出于保护植株的目的, 游标卡尺的卡位较为松动, 这些因素可能是造成模型预测精度降低的原因。

### 3 结论与讨论

主成分分析采用降维处理方法, 将原来多个相互联系的因素综合成几个少数代表性的变量, 该变

量既充分保留了原来因素的信息量,又避免了新变量间信息的重复,从而达到了科学简化的目的。代表作物长势的因素很多,可以归纳为株高、茎粗、叶片数、叶面积和叶绿素含量等一系列的指标,这些指标之间存在着较强的相关性,如果利用这些全部原始指标建模,势必造成模型庞大且不能做出合理的解释。本文利用主成分分析法,以温室甜瓜为例,把代表作物长势的一系列指标综合为综合植株性状,该指标可以根据原变量信息的贡献值确定权重系数,既充分表达了植株长势信息,又避免了原信息的重叠,利用综合植株性状与有效积温建立 Logistic 经验生长模型,经检验,模型拟合精度较高,可以用来进行作物生长方程的模拟。

代表植株长势的因素还包括地下部分的根长、根粗、须根数等指标,本文仅利用地上易测量部分的常规指标建模,还具有一定的片面性,且影响作物长势的环境因素还包括土壤肥力、太阳辐射等因素,本文仅利用有效积温建模,也是一种理想化的方法。甜瓜打顶之后如何利用果实膨大速度、果实重量等指标来建立整个生育期的生长模型,这些问题需要进行进一步深入研究。

#### 参考文献:

- [1] 赵伟霞,蔡焕杰,单志杰,等.无压灌溉日光温室番茄高产指标[J].农业工程学报,2009,25(3):16-21.
- [2] 常水晶,左兵,王晓辉,等.光照、温度及盐分对钠猪毛菜种子萌发的影响[J].干旱区地理,2008,31(6):897-903.
- [3] Shafi M, Bakht J, Hassan M J, et al. Effect of cadmium and salinity stresses on growth and antioxidant enzyme activities of wheat (*Triticum aestivum* L.) [J]. Bull Environ Contam Toxicol, 2009, 82: 772-776.

- [4] Timmermans B G H, Vos J, Stomph T J. The development, validation and application of a crop growth model to assess the potential of *Solanum sisymbriifolium* as a trap crop for potato cyst nematodes in Europe [J]. Field Crops Research, 2009, 111: 22-31.
- [5] 殷红.作物生产系统模拟模型研究进展[J].杂粮作物,2000,20(3):30-33.
- [6] Bojacá C R, Gil R, Cooman A. Use of geostatistical and crop growth modelling to assess the variability of greenhouse tomato yield caused by spatial temperature variations [J]. Computers and Electronics in Agriculture, 2009, 65: 219-227.
- [7] 孙忠富,陈人杰.温室园艺作物生长发育模型研究现状与发展趋势[J].园艺学报,2001,28(增刊):700-704.
- [8] 李建明.温室厚皮甜瓜幼苗生长模型与生态生理变化规律研究[D].杨凌:西北农林科技大学,2006.
- [9] Adams S R, Pearson S, Hadley P. Modeling growth and development of pansy cv. Universal Violet in response to photo-thermal environment: application for decision support and scheduling [J]. Acta Horticulturae, 1996, 417: 23-32.
- [10] 林忠辉,项月琴,莫兴国,等.夏玉米叶面积指数增长模型的研究[J].中国生态农业学报,2003,11(4):69-72.
- [11] 张旭东,蔡焕杰,付玉娟,等.黄土区夏玉米叶面积指数变化规律的研究[J].干旱地区农业研究,2006,24(2):25-29.
- [12] 李建明,邹志荣.温度、光辐射及水分对温室甜瓜幼苗干物质积累与分配的影响及其模拟模型[J].应用生态学报,2007,18(12):2715-2721.
- [13] 唐启义,冯明光.实用统计分析及其 DPS 数据处理系统 [M].北京:科学出版社,2002.
- [14] 王燕,蔡焕杰,陈新明,等.根区局部控水无压地下灌溉对番茄生理特性及产量、品质的影响 [J].中国农业科学,2007,40(2):322-329.
- [15] 曹卫星,罗卫红.作物系统模拟及智能管理 [M].北京:华文出版社,2001.
- [16] 中国农业科学院郑州果树研究所.中国西瓜甜瓜 [M].北京:中国农业出版社,2000.

## Research on muskmelon growth formula in greenhouse based on principal component analysis

XIE Heng-xing<sup>1</sup>, CAI Huan-jie<sup>1\*</sup>, ZHANG Zhen-hua<sup>2</sup>, WANG Jia-peng<sup>1</sup>

(1. State Key Laboratory for Agricultural Soil and Water Engineering of

Ministry of Education, Northwest A&F University, Yangling, Shaanxi 712100, China;

2. College of Geography and Planning, Ludong University, Yantai, Shandong 264025, China)

**Abstract:** The development process of muskmelon in greenhouse with different water supply was simulated by principal component analysis, and the universal application of the model was tested. The results showed that the logistic model constructed with the comprehensive plant characteristics and effective accumulative temperature had very high precision with the determining coefficient above 0.9. The plant development process in the 3<sup>rd</sup> trial district was simulated with the model, and the relative error between measured value and fitted value was mostly among 10%. The principal component analysis model had pretty practicality in plant development process simulation with initial information existed fully, and repetition information deleted.

**Keywords:** muskmelon; principal component analysis; plant development formula model; comprehensive plant characteristic; fitted precision