

# 城市蔬菜品种安全监控系统的研究及应用

张国权<sup>1</sup>, 肖莉<sup>1</sup>, 张兵<sup>2</sup>, 岑冠军<sup>1</sup>

(1 华南农业大学理学院, 广东广州 510642; 2 深圳市无公害农产品质量监督检验站, 广东深圳 518040)

**摘要:** 基于深圳市农产品批发市场主要蔬菜品种质量的监测抽样数据, 应用多元统计方法, 根据城市蔬菜品种污染动态响应系数和累积影响系数, 进一步构建蔬菜品种安全指数, 从而建立城市蔬菜品种安全监控系统。

**关键词:** 蔬菜; 安全指数; 多元统计; 监控预警图

中图分类号: O212.4

文献标识码: A

文章编号: 1001-411X(2005)02-0108-04

## Research on urban vegetable security monitoring system and its application

ZHANG Guo-quan<sup>1</sup>, XIAO Li<sup>1</sup>, ZHANG Bing<sup>2</sup>, CEN Guan-jun<sup>1</sup>

(1 College of Sciences, South China Agric. Univ., Guangzhou 510642, China;

2 Safetyfood Quality Supervise Inspection Station, Shenzhen 518040, China)

**Abstract:** By means of multivariate statistical methods, the sampling data from farm produce wholesale market in Shenzhen City were used to analyze in this paper. Based on the construction of vegetable variety pollution dynamic response coefficient and cumulative influence coefficient, the concept of vegetable variety security index is presented, which was used to establish the vegetable variety security control system.

**Key words:** vegetable; security index; multivariate statistic; monitoring and pre-warning chart

当前, 由于环境对农业生产的污染、农业投入品的使用不当以及农产品市场管理的种种弊端, 致使农产品中残留的有毒有害物质超标, 农产品质量安全问题日益凸显, 特别是蔬菜产品农药残留问题, 质量安全事件时有发生, 因此建立蔬菜产品质量安全状况的及时监控系统刻不容缓。本文以深圳市农产品批发市场主要蔬菜产品质量状况的监测抽样数据为考察对象, 根据蔬菜产品农药残留的检测要求, 应用多元统计方法, 在构建城市蔬菜品种污染动态响应指数和累积影响系数的基础上, 进一步提出蔬菜品种安全指数的概念, 建立城市蔬菜产品品种安全监控系统, 以期能科学地、客观地、及时地刻划、测量、监管蔬菜产品品种的质量状况, 同时能对不同蔬菜产品品种质量状况作出分析、评价和走向预测, 及时监控流入市场的蔬菜质量, 并为政府相关部门提

供决策依据。

## 1 城市蔬菜品种安全监控系统的指标体系

### 1.1 城市蔬菜品种安全指数的定义

城市蔬菜品种安全指数 (variety index, 简称  $I_v$ ) 是指将城市市场上某一品种的蔬菜监测检验结果转化成数值形式, 通过分级反映和评价城市市场上某一品种蔬菜的质量安全状况, 并用简单明了的形式向市民提供及时的蔬菜品种质量状况和安全水平的信息。

### 1.2 城市蔬菜品种安全监控系统指标体系的确定

科学指标体系的确定是建立城市蔬菜品种安全监控系统的前提。本文根据深圳市农业标准体系和农产品质量检测网络对城市蔬菜产品监测抽样的技

术指标<sup>1, 2</sup>来建立城市蔬菜品种安全监控系统的指标体系。

定义 1:

$$\begin{cases} x_1 = \frac{\text{该蔬菜品种抽检不合格批次}}{\text{该蔬菜品种抽检总批次}} \times 100\%, \\ x_2 = \frac{\text{该蔬菜品种抽检不合格批量}}{\text{该蔬菜品种抽检总批量}} \times 100\%, \end{cases} \quad (1)$$

式中,  $x_1$  为蔬菜品种批次分量污染频率,  $x_2$  为蔬菜品种批量分量污染频率。

以上定义的指标体系  $X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$  能够反映

出蔬菜品种污染情况的动态过程, 其中每一个指标的变化都会体现出蔬菜品种总体污染情况出现不同程度的变异,  $t=1, 2, \dots, n$ 。

## 2 城市蔬菜品种安全监控系统的建立

### 2.1 城市蔬菜品种污染动态响应系数的建立

由于城市蔬菜品种指标体系  $X(t)$  的变化主要表现为 2 个分量  $x_1(t)$ 、 $x_2(t)$  的同向变化, 但是由于  $x_1(t)$  为计点抽样检验,  $x_2(t)$  为计量抽样检验, 两者抽样检验类型不同, 很有可能会导致  $x_1(t)$ 、 $x_2(t)$  的变化速度不相同, 甚至会出现变化反向的情况, 且  $x_1(t)$ 、 $x_2(t)$  相互影响, 存在着相关性, 所以不能仅依据单分量  $x_1(t)$  或  $x_2(t)$  来计算城市蔬菜品种污染指数, 应该从总体上对 2 个分量  $x_1(t)$ 、 $x_2(t)$  进行综合。

设  $X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$  为因子集序列, 记为  $X(t)$ ,

取  $n$  个样本,  $t=1, 2, \dots, n$ 。作变换  $f[X(t)]$ , 使变换后的变量符合某种概率分布, 以便构造一个监控系统。

考虑到  $x_1(t)$ 、 $x_2(t)$  有很强的相关性, 而主成分分析法是研究如何用多个指标来描述研究单位(个体)的一种统计方法<sup>1, 3</sup>。该方法把原来多个彼此相关的指标(原变量)线性组合为少数几个彼此独立的综合指标(新变量), 线性组合成的新变量是关于原变量的综合指标, 它提取了原来全部指标的主要成分的统计信息。

本文取  $e_i(t)$  为样本矩阵  $X_{n \times 2}$  的主成分变量系数向量, 即为  $X(t)$  构成的相关矩阵的特征值  $\lambda_i$  对应的特征向量。  $Y_1(t)$  称为  $X(t)$  下的蔬菜品种污染第 1 动态响应系数,  $Y_2(t)$  称为  $X(t)$  下的蔬菜品种污染

第 2 动态响应系数。

定义 2:

$$\begin{cases} Y_1(t) = e'_1 X(t) = e_{11}x_1(t) + e_{21}x_2(t), \\ Y_2(t) = e'_2 X(t) = e_{12}x_1(t) + e_{22}x_2(t), \end{cases} \quad (2)$$

式(2)为污染频率的系统控制模型,  $Y_1(t)$ 、 $Y_2(t)$  分别称为第 1、第 2 蔬菜品种污染动态响应系数,  $t=1, 2, \dots, n$ 。

### 2.2 蔬菜品种污染累计影响系数 $T_j^2$

对第  $h$  个主成分  $y_h (h=1, 2)$  有:

$$\text{Var}(y_n) = \frac{1}{n} \sum_{j=1}^n y_{hj}^2 = \lambda_h.$$

定义 3: 蔬菜品种污染累计影响系数

$$T_j^2 = \frac{y_{1j}^2}{\lambda_1} + \frac{y_{2j}^2}{\lambda_2}, \quad (3)$$

其中,  $y_{1j}$  和  $y_{2j}$  分别为本研究中的蔬菜品种污染第 1、第 2 动态响应系数,  $\lambda_1$  和  $\lambda_2$  为特征值, 蔬菜品种污染累计影响系数  $T_j^2$  可用来确定样本集合中的特异点(危险点)。本文中  $T_j^2$  又称为蔬菜品种污染指数。

定理 1: 设  $X_1, X_2, \dots, X_n$  i. i. d  $N_p(u, \Sigma)$ , 其样本主成分为  $Y_i = e'_i (X - \bar{X})$ , 则

$$T_j^2 = \sum_{i=1}^2 \frac{Y_i^2}{\lambda_i} \sim T^2(2, n-1), \quad (4)$$

其中  $e_i$  为特征向量,  $\lambda_i$  为特征值。

### 2.3 利用 $T_j^2$ 分布确定监控预警限

定义 4: 对于给定的检验水平  $\alpha (0 < \alpha < 1)$ , 当样

本点  $j$  满足  $T_j^2 > \frac{2(n-1)}{n-2} F_{2, n-2}(\alpha)$  时, 称样本点  $j$  为特异点(危险点); 而当样本点  $j$  满足  $T_j^2 \leq \frac{2(n-1)}{n-2} F_{2, n-2}(\alpha)$  时, 称样本点  $j$  为非特异点(稳定点)。

因为  $\frac{y_{1j}^2}{\lambda_1} + \frac{y_{2j}^2}{\lambda_2} \geq \frac{2(n-1)}{n-2} F_{2, n-2}(\alpha) \triangleq C_\alpha$

则有  $\frac{y_{1j}^2}{\lambda_1} + \frac{y_{2j}^2}{\lambda_2} \leq C_\alpha$  (5)

是一个椭圆, 所以在  $y_1, y_2$  平面上可以做出椭圆图。这里  $C_\alpha$  称为阈值, 本研究中  $C_\alpha$  分为警戒区阈值  $W_\alpha$  和危险区阈值  $A_\alpha$ , 同时  $C_\alpha$  也是  $T^2(2, n-1)$  分布的上  $100\alpha$  百分位数。

当样本点  $j$  落在椭圆外时, 即  $\frac{y_{1j}^2}{\lambda_1} + \frac{y_{2j}^2}{\lambda_2} > C_\alpha$ , 认为该样本点是在检验水平  $\alpha$  下的特异点; 当样本点  $j$

落在椭圆内时, 即  $\frac{Y_1^2}{\lambda_1} + \frac{Y_2^2}{\lambda_2} \leq C_\alpha$ , 认为该样本点是在检验水平  $\alpha$  下的非特异点, 见图 1.

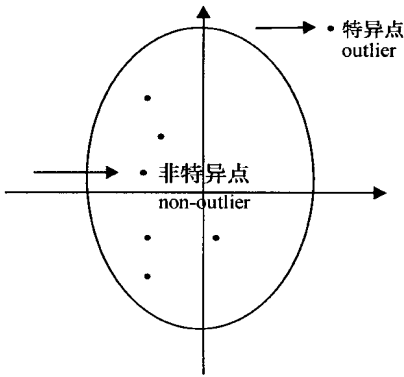


图 1 椭圆图

Fig. 1 Elliptic chart

根据图 1, 本文通过取不同的检验水平  $\alpha$  来获得不同的监控预警限. 通过对历史样本进行聚类分析<sup>[3]</sup>发现: 检验水平  $\alpha = 0.15$  时, 得到阈值  $W_{0.15} = L_S$  为警戒限; 检验水平  $\alpha = 0.05$  时, 得到阈值  $A_{0.05} = L_W$  为行动限. 样本点  $j$  如果满足  $T_j^2 > \frac{2(n-1)}{(n-2)} F_{2, n-2}(0.05)$ , 认为该样本点是特异点(危险点).

2.4 城市蔬菜品种安全指数  $I_V$  的计算公式:

$$\begin{cases} I_V = 100 - 10 \times T^2(t), & T^2(t) < 10 \\ I_V = 0, & T^2(t) \geq 10 \end{cases} \quad (6)$$

其中,  $T^2(t) = \frac{Y_1^2}{\lambda_1} + \frac{Y_2^2}{\lambda_2}$ ,  $Y_1$  为第 1 主成分得分;  $Y_2$  为第 2 主成分得分,  $t = 1, 2, \dots, n$ .

2.5 城市蔬菜品种安全指数监控预警图

为了便于广大市民识别蔬菜品种安全水平, 利用蔬菜品种安全指数和带警戒限的均值控制图原理<sup>[4]</sup>建立城市蔬菜品种安全指数监控预警图(见图 2).

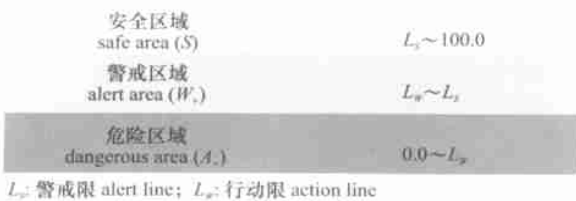


图 2 蔬菜品种安全指数监控预警图

Fig. 2 The supervising and pre-warning chart of vegetable security index

图 2 中,  $L_S, L_W$  是 2 条水平线, 分别称为警戒限和行动限, 这 2 条水平线将监控预警图分为 3 个区

域: 安全区域( $S$ )、警戒区域( $W_+$ )、危险区域( $A_+$ ), 该监控预警图中还有一条中心限( $U$ )<sup>[5]</sup>. 实际运行时, 若某蔬菜品种安全指数落在  $A_+$ , 表示过程失控, 即该蔬菜品种污染严重, 质量差; 若某蔬菜品种安全指数落在  $W_+$ , 表示该蔬菜品种污染比较严重, 质量较差, 需向市民发出警告, 如果某蔬菜品种连续一段时间都落在  $W_+$ , 表示过程呈现失控趋势, 即该蔬菜品种在这段时间会产生污染严重后果, 菜农应该及时地减少对该蔬菜品种农药的使用量; 若某蔬菜品种安全指数落在  $S$ , 表示该蔬菜品种污染水平低、质量好; 若某蔬菜品种安全指数落在  $L_S \sim U$ , 表示该蔬菜品种质量良好; 若某蔬菜品种安全指数落在  $U \sim 100.0$ , 则表示该蔬菜品种质量优良.

监控预警限  $L_S, L_W, U$  的计算公式分别为:

$$L_S = 100 - 10 \times T_{p, n-1}^2(\alpha) = 100 - 10 \times \frac{p(n-1)}{n-p} F_{p, n-p}(\alpha), \quad (\alpha = 0.25) \quad (7)$$

$$L_W = 100 - 10 \times T_{p, n-1}^2(\alpha) = 100 - 10 \times \frac{p(n-1)}{n-p} F_{p, n-p}(\alpha), \quad (\alpha = 0.05) \quad (8)$$

$$U = 100 - 10 \times \frac{p(n-1)}{n-p-2}, \quad (9)$$

其中,  $p = 2$ .

经过对深圳市 2 大农贸市场的蔬菜品种监测抽样历史数据的研究和计算, 同时结合实际情况, 可以将  $L_S$  确定在 50.6 左右,  $L_W$  确定在 13.1 左右,  $U$  确定在 73.3 左右, 那么蔬菜品种安全水平的分级见图 3.

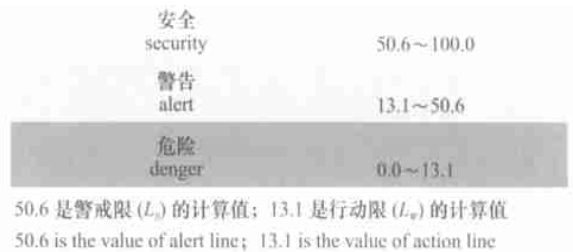


图 3 蔬菜品种安全水平分级标准

Fig. 3 Vegetable variety security index and the grade standard of security level

2.6 对未来蔬菜品种的监控

根据上述理论, 如有某蔬菜品种安全水平落在危险区域, 说明该蔬菜品种是特异点(危险点), 从而判断该蔬菜品种安全监控系统失控, 一个失控的蔬菜品种安全监控系统是不稳定的, 无法对未来蔬菜品种进行监控, 因此, 为保证蔬菜品种安全监控系统

的稳定性并对未来蔬菜品种能进行监控, 就应剔除特异点(危险点), 同时进行补充采样, 然后利用剩下的蔬菜品种按照上述理论重新计算, 如此反复进行下去, 直到各蔬菜品种安全水平都是非特异点为止<sup>[6]</sup>。

### 3 实例分析

以 2003 年 1~2 月份菜心的监测抽样数据为例, 首先利用幂变换法<sup>[7]</sup>对原始数据进行预处理, 对经过预处理的原始数据进行正态分布拟合检验, 结果表明: 经过预处理的原始数据服从正态分布, 同时满足  $(x-u)' \sum^{-1} (x-u) \leq \chi_p^2(0.5)$ <sup>[6]</sup>, 再运用上述原理和方法计算得到 2003 年 1~2 月份菜心品种安全指数  $I_V$ <sup>[8]</sup>, 结果见表 1 和图 4。

表 1 2003 年 1~2 月份菜心品种安全指数

Tab. 1 *Brassica parachinensis* variety security index from January to February, 2003

编号 number	日期(1月) date(January)	安全指数 security index	编号 number	日期(2月) date(February)	安全指数 security index
1	01~02	95.59	14	03~04	99.08
2	03~05	79.69	15	05~06	85.81
3	06~07	85.81	16	07~09	98.23
4	08~09	54.38	17	10~11	85.81
5	10~12	91.57	18	12~13	59.45
6	13~14	85.81	19	14~16	93.76
7	15~16	91.25	20	17~18	94.11
8	17~19	30.75	21	19~20	76.63
9	20~21	85.81	22	21~23	70.24
10	22~23	92.73	23	24~25	76.38
11	24~26	88.96	24	26~27	46.51
12	27~28	85.81			
13	29~30	85.81			

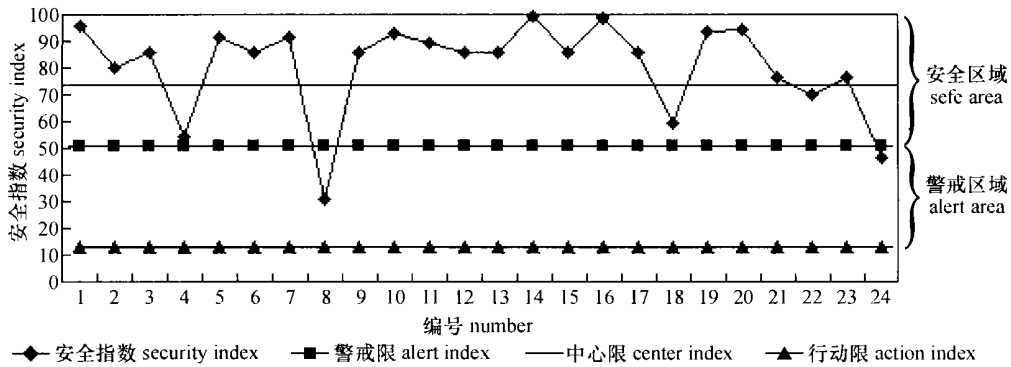


图 4 2003 年 1~2 月份菜心安全指数监控预警图

Fig. 4 The supervising and pre-waming chart of *Brassica parachinensis* security index from January to February, 2003

从表 1 和图 4 可以清楚地了解 1~2 月份菜心的质量安全状况: 1 月 17~19 日、2 月 26~27 日这 2 个时间段中菜心品种安全指数分别为 30.75、46.51, 落在  $W_+$  内, 处于警告水平, 说明菜心在上述 2 个时间段里有污染, 质量状况不是很好, 应向市民和菜农分别发出警告, 市民购买这类蔬菜后应多用水清洗, 同时应警告菜农减少对菜心的农药施用量; 除这 2 个时间段之外, 菜心品种安全指数均落在  $S$  内, 处于安全水平, 并且在 1~2 月份中大部分菜心品种安全指数都落在  $U \sim 100.0$  这个区域, 说明菜心在这 2 个月里的质量状况大部分时间都是优良的, 市民可放心购买。

参考文献:

[1] DB 440300/T 20-2001, 无公害蔬菜检测技术规程[S].

[2] DB 440300/T 20-2001, 无公害蔬菜[S].  
 [3] 张尧庭, 方开泰. 多元统计分析引论[M]. 北京: 科学出版社, 1982. 79-289.  
 [4] 汪仁官, 王振羽. 带警戒限的均值控制图中平均链长的计算公式[J]. 数理统计与管理, 2002, 21(2): 48-52.  
 [5] JACKSON J E. Principal Components and Factor Analysis, Part I—Principal Components[J]. Journal of Quality Technology, 1980, 12: 201-213.  
 [6] JOHNSON RA, WICHERN D W. Applied multivariate statistical analysis[M]. [s. l.]: International Edition, 1998. 490-497.  
 [7] BOX G E P, COX D R. An analysis of transformation (with discussion)[J]. Journal of the Royal Statistical Society (B), 1964, 26(2): 211-252.  
 [8] 高惠璇, 耿直, 李贵斌, 等. SAS 系统 SAS/STAT 软件使用手册[M]. 北京: 中国统计出版社, 1997. 115-400.

【责任编辑 李晓卉】