

# 一种基于模糊数学的生物灾害预测模型

胡建根

(江西农业大学 理学院, 江西 南昌 330045)

**摘要:** 生物灾害的产生与多种因素有关, 灾害预测是减灾的重要措施。本文运用模糊数学理论建立了一个预测生物灾害的数学模型, 通过输入历史和现时的影响某种生物灾害的数据, 可预测这种生物灾害发生的强度。

**关键词:** 生物灾害; 模糊数学; 预测; 模型

中图分类号: S431.3

文献标志码: A

文章编号: 2095-3704 (2012) 02-0219-03

## A Based on Fuzzy Mathematics Biological Disaster Prediction Model

HU JIAN-GEN

(College of Sciences, Jiangxi Agricultural University, Nanchang)

**Abstract:** Biological disasters are related to many factors, disaster forecast is the important measure of disaster reduction. In this paper, using the theory of fuzzy sets to establish a mathematical model of prediction of biological disasters, inputting the history and current impact of some biological disaster data, can predict the biological disaster intensity.

**Key Words:** Biological disasters; Fuzzy sets; Forecast; Model

生物灾害是一种常见的自然灾害, 对农业生产影响很大, 它与气候、环境的变化及人类活动等因素有关, 其作用机理比较复杂。生物灾害产生的原因具有模糊性, 它对农业生产的破坏程度也具有模糊性。本文试图运用模糊数学理论, 根据历史上有关某种生物灾害的数据信息, 建立一种预测生物灾害的数学模型, 研究某种生物灾害发生的程度与各影响因素的变化规律。

## 1 历史资料收集整理

### 1.1 变量假设

假设要预测某种生物病害或虫害, 根据历史资料信息, 找出产生此种生物灾害的相关因素, 作为样本指标, 每一年作为一个样本。构造论域  $U = \{x_1, x_2, \dots, x_n\}$ ,  $x_i$  表示第  $i$  年 ( $i=1, 2 \dots n$ ),  $(x_{i1}, x_{i2}, \dots,$

$x_{im})$ , 表示第  $i$  年  $m$  个相关因素的数据向量,  $x_{ij}$  表示第  $i$  年第  $j$  个相关因素的数值 ( $i=1, 2, \dots, n; j=1, 2, \dots, m$ )。

### 1.2 建立数据矩阵

根据收集到的数据, 建立数据矩阵:

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

### 1.3 数据标准化

由于不同的相关因素所对应的数据有不同的量纲, 为了使不同量纲的量能进行比较, 需对数据作适当的变换, 使所有数据都在区间  $[0, 1]$  上, 这就是数据标准化。标准化步骤如下:

收稿日期: 2012-06-22

基金项目: 2010 年江西农业大学校基金项目

作者简介: 胡建根, 男, 江西新余人, 副教授, 主要研究方向为不确定数学、数据分析及其应用。

(1) 标准差变换

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_k} \quad (i=1,2,\dots,n; j=1,2,\dots,m)$$

其中

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}, \quad s_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$$

经过上述变换后, 每个变量的均值为 0, 标准差为 1, 且消除了量纲的影响。但是, 这样得到的  $x'_i$  还不一定在区间 [0, 1] 上, 因此还要作下列变换:

(2) 极差变换

$$y_{ij} = \frac{x'_{ij} - \min_{1 \leq i \leq n} \{x'_{ij}\}}{\max_{1 \leq i \leq n} \{x'_{ij}\} - \min_{1 \leq i \leq n} \{x'_{ij}\}} \quad (j=1,2,\dots,m)$$

经变换得标准矩阵:

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{bmatrix}$$

## 2 建立模糊相似矩阵

模糊矩阵定义: 如果对于任意  $i=1, 2, \dots, n; j=1, 2, \dots, m$ , 都有  $r_{ij} \in [0, 1]$ , 则称矩阵

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}$$

为模糊矩阵。

模糊相似矩阵定义: 设论域  $U=\{x_1, x_2, \dots, x_n\}$ ,  $R$  对应的为相似矩阵,  $I$  为单位矩阵, 若满足: (1) 自反性  $I \leq R$ , (2) 对称性  $R^T=R$ , 则称  $R$  为模糊相似矩阵。

对论域  $U=\{x_1, x_2, \dots, x_n\}$  对应的标准矩阵  $Y$ , 应用模糊聚类方法, 确定  $x_i$  与  $x_j$  之间的相似程度, 用  $r_{ij}=R(x_i, x_j)$  表示。下列两种方法可求  $x_i$  与  $x_j$  之间的相似程度  $r_{ij}$ :

(1) 数量积相似法

$$r_{ij} = \begin{cases} 1 & i = j \\ \frac{1}{M} \sum_{k=1}^m y_{ik} \cdot y_{jk} & i \neq j \end{cases} \quad \text{其中}$$

$$M = \max_{i \neq j} \left( \sum_{k=1}^m y_{ik} \cdot y_{jk} \right),$$

显然  $|r_{ij}| \in [0, 1]$ , 若  $r_{ij}$  中出现负值, 则可采用以下方法将  $r_{ij}$  压缩到 [0, 1] 上:

$$r'_{ij} = \frac{r_{ij} + 1}{2}, \quad \text{则 } r'_{ij} \in [0, 1]$$

(2) 指数相似法

$$r_{ij} = \frac{1}{m} \sum_{k=1}^m e^{\left[ -\frac{3(x_{ik} - x_{jk})^2}{4s_k^2} \right]}, \quad \text{其中}$$

$$s_k = \frac{1}{n} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2$$

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik}, \quad (k=1,2,\dots,m)$$

从而得到模糊相似矩阵:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}$$

## 3 聚类判断, 预测

3.1 确定模糊等价矩阵

将模糊矩阵  $R$  改造成模糊等价矩阵  $R^*$ , 其方法是用二次方法:

$R \rightarrow R^2 \rightarrow R^4 \rightarrow \cdots$ , 直到  $R^k \cdot R^k = R^k$  为止, 则得到传递闭包  $t(R) = R^* = R^k$

$$R^* = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ s_{n1} & s_{n2} & \cdots & s_{nm} \end{bmatrix}$$

3.2 最佳阈值  $\lambda$  的确定

模糊相似矩阵的  $\lambda$ -截矩阵定义: 设  $A=(a_{ij})_{n \times m}$

为模糊相似矩阵,对于任意  $r_{ij} \in [0, 1]$ , 称

$$A_\lambda = (a_{ij}^{(\lambda)})_{n \times m} \text{ 为 } A = (a_{ij})_{n \times m} \text{ 的 } \lambda\text{-截矩阵, } \lambda$$

称为阈值,其中

$$a_{ij}^{(\lambda)} = \begin{cases} 1 & a_{ij} \geq \lambda \\ 0 & a_{ij} < \lambda \end{cases}$$

例如,设论域  $U = (x_1, x_2, x_3, x_4)$ , 有4个相关因子,通过上述运算,得到模糊等价矩阵:

$$R^* = \begin{bmatrix} 1 & 0.5 & 0.2 & 0 \\ 0.5 & 1 & 0.1 & 0.3 \\ 0.2 & 0.1 & 1 & 0.8 \\ 0 & 0.3 & 0.8 & 1 \end{bmatrix}$$

若取  $\lambda = 1$ , 则得到1-截矩阵

$$R_1^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

则论域  $U$  可分四类:  $\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}$ ;

若取  $\lambda = 0.8$ , 则得到0.8-截矩阵

$$R_{0.8}^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

则论域  $U$  可分三类:  $\{x_1\}, \{x_2\}, \{x_3, x_4\}$ ;

取  $\lambda = 0.5$ , 则得到0.5-截矩阵

$$R_{0.5}^* = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

则论域  $U$  可分二类:  $\{x_1, x_2\}, \{x_3, x_4\}$ 。

在模糊聚类分析中,根据阈值  $\lambda$  对论域  $U = \{x_1, x_2, \dots, x_n\}$  进行分类。不同的  $\lambda$  可得到不同的分类,要得到合理的分类就需要确定最佳阈值  $\lambda$ 。一般有两种方法确定最佳阈值  $\lambda$ : (1) 根据实际需要,调整  $\lambda$  的值,从1慢慢减少,结合专家意见,确定  $\lambda$

的值,从而得到在  $\lambda$  水平上的合理的等价分类;(2) 用F统计量确定  $\lambda$  的最佳值,此种方法不需要专家意见,但计算较复杂。

### 3.3 分类定级,预测

根据最佳阈值  $\lambda$ , 在  $R^*$  上,结合该种生物的危害或虫害历史上的历史危害程度,将论域  $U = \{x_1, x_2, \dots, x_n\}$  分成若干等级类,例如,危害程度可分:一类(危害轻年份);二类(危害较重年份);三类(危害重年份);四类(危害严重年份)。

将要预测年份的有关该生物的  $m$  个变量的数据输入上述模型,经过运算,取最佳阈值  $\lambda$ , 可判断该年份归并为哪一类,从而可预测该年该生物灾害危害的轻重程度。

## 4 讨论

(1) 由于该模型中,论域  $U$  中的样本个数  $n$ , 影响因子数  $m$ , 阈值  $\lambda$  都是可变的,因此可动态地进行分类,根据实际需要,对生物灾害危害程度进行等级分类,且是动态分类。该模型可进行计算机编程或运用数学软件来计算,只要输入某年影响该生物灾害  $m$  个因子的数值,即可预测该年发生该生物灾害危害的轻重程度,从而制订各种抗灾措施。

(2) 该模型不具备分析某种生物灾害产生的作用机理,但可通过增加、减少、变动相关影响因子来运行该模型,从而了解各影响因子对灾害产生的影响程度。

### 参考文献:

- [1] Zadeh L A. Fuzzy sets. Information and Control, 1965, 8: 338-353
- [2] 汪培庄. 模糊集合理论及其应用[M]. 上海: 上海科学技术出版社, 1983.
- [3] 夏敬源. 我国重大农业生物灾害暴发现状与防控成效[J]. 中国植保导刊, 2009, 1: 6-9.
- [4] 曹志勇, 邱靖, 曹志娟, 等. 基于改进型神经网络的植物病虫害预警模型的构建[J]. 安徽农学通报, 2010, 38(1): 538-540.
- [5] 谢季坚, 刘承平. 模糊数学方法及其应用(第三版)[M]. 华中科技大学出版社, 2006.