

陈燕璇, 刘合香, 谭金凯. 基于等距特征映射降维的台风灾情概率神经网络预评估模型[J]. 灾害学, 2016, 31(3): 20 - 25, 30. [CHEN Yanxuan, LIU Hexiang and TAN Jinkai. Probabilistic Neural Network Pre - Assessment Model Based on Isometric Feature Mapping Dimensional Reduction in Typhoon Disaster[J]. Journal of Catastrophology, 2016, 31(3): 20 - 25, 30.]

## 基于等距特征映射降维的台风灾情概率神经网络预评估模型\*

陈燕璇, 刘合香, 谭金凯

(广西师范学院 数学与统计科学学院, 广西南宁 530023)

**摘 要:** 台风致灾因子、承灾体和灾情之间是一个复杂的非线性动力系统, 准确高效地提取重要指标对台风灾情等级进行预评估, 是防灾救灾工作的重要依据。采用主成分分析、等距特征映射和信息熵特征提取的承灾体关键指标, 和致灾源作为输入神经元, 灾情等级作为输出神经元, 建立台风灾情概率神经网络预评估模型。结果表明, 基于等距映射非线性特征提取的概率神经网络预评估模型的准确率达到 90%。

**关键词:** 概率神经网络; 等距特征映射; 信息熵; 台风; 灾害; 预评估

**中图分类号:** X43; TP3 **文献标志码:** A **文章编号:** 1000 - 811X(2016)03 - 0020 - 07

doi: 10.3969/j.issn.1000-811X.2016.03.004

热带气旋是生成于热带或副热带洋面上, 具有对流和确定气旋性环流的非锋面性涡旋<sup>[1]</sup>。台风作为热带气旋的一种, 登陆我国的台风平均每年有 7 个左右, 影响广西的台风平均每年有 5 个左右。据不完全统计, 1984 - 2012 年, 广西因台风灾害造成的受灾人数为 8 148 万人次, 死亡 801 人, 直接经济损失 1 205 亿元(占气象灾害总经济损失的 12.3%)<sup>[2]</sup>。灾害发生过程根据气象致灾因子预报与承灾体脆弱性快速预测灾情等级是灾害的预评估。人工神经网络是非线性自适应系统, 与耗散、复杂的高阶非线性自然灾害系统相似<sup>[3]</sup>, 灾情等级预评估实际是人工神经网络模式识别问题<sup>[4]</sup>。

近年来, 很多专家学者从不同角度对台风灾害进行预评估研究<sup>[5-8]</sup>。魏章进等<sup>[9]</sup>基于聚类与回归方法建立台风灾情预评估模型; 芮建勋等<sup>[10]</sup>设计了基于元组时间标记法的台风灾害评估的多尺度数据管理模式。在地理信息系统方面, 刘合香等<sup>[11]</sup>利用模糊数学、非线性数据处理方法和 Arc-GIS 空间进行 Kriging 插值, 分析广西洪涝灾害发生的频率, 结果表明洪涝多发生地与实际洪涝灾害拟合较好; 刘少军等<sup>[12]</sup>采用可拓分析方法计算综合关联度判断灾害损失的等级, 建立基于 GIS 的

台风灾害损失评估模型, 以实现台风灾害动态评估。

承灾体的脆弱性是指受到危险因素威胁所有人生命和财产的损害程度, 是灾害风险形成的关键因素<sup>[13]</sup>, 一般当承灾体指标比较多时, 会导致数据的高维特征, 增加数据处理的困难。现在武等<sup>[14]</sup>利用相关分析从致灾因子、承灾体和防灾减灾能力方面选取重要影响因子, 实例证明所选指标的合理性。但相关分析只表示变量间线性关系, 不能反映变量间的非线性关系, 且易受观测值影响。流形学习是根据高维数据空间的内部几何结构, 构造低维流形嵌入, 以实现数据降维, 提高计算效率<sup>[15]</sup>。非线性流形降维方法主要有由局部线性嵌入(LLE)和等距特征映射(ISOMAP)。黄颖等<sup>[16]</sup>利用局部线性嵌入与逐步回归相结合的预报因子挖掘技术, 建立非线性人工智能集合预报模型, 为台风强度客观预报提供了新方法; 等距特征映射算法(ISOMAP)是 Tenenbaum 等<sup>[17]</sup>于 2000 年提出的, 它结合了主成分分析(PCA)和多维标度法(MDS)的算法特征, 在医学方面的肺癌基因数据分析中, 在低维空间揭示出数据集的本质结构, 实现非线性数据降维<sup>[18]</sup>, 而在台风灾害系统指标降维方面的应用还鲜为少见。

\* 收稿日期: 2016 - 01 - 13 修回日期: 2016 - 03 - 11

基金项目: 国家自然科学基金(41465003); 广西研究生教育创新计划项目(YCSZ201585)

第一作者简介: 陈燕璇(1988 -), 女, 广东汕头人, 硕士研究生, 研究方向为概率统计、数学模型、自然灾害风险分析。

E-mail: colourful48@qq.com

通讯作者: 刘合香(1962 -), 女, 山东茌平人, 教授, 硕士生导师, 研究方向为概率统计、数学模型、自然灾害风险分析。

E-mail: hx\_post@163.com

针对上述问题, 为了比较不同的变量选取方法的预评估效果, 本研究运用主成分分析(PCA)、等距特征映射(ISOMAP)和信息熵特征提取方法试图建立概率神经网络模型, 并将该模型应用于台风灾情预评估工作中。

## 1 模型输入的特征提取方法

### 1.1 ISOMAP 降维方法

ISOMAP 的思想关键在于用测地距离代替欧氏距离, 通过等距映射获得高维数据空间在低维空间的表示, 更好地实现非线性数据降维<sup>[17]</sup>。ISOMAP 算法设定  $K$  个近邻点是相互连接的, 通过欧氏距离构造邻接矩阵, 用 Floyd 算法计算样本点的最短距离, 作为测地距离的逼近<sup>[19]</sup>, 主要包括以下的步骤。

(a) 构造邻域图  $G$ 。在空间  $X$  中的样本点  $x_i$  和  $x_j$ , 其欧氏距离为  $d_x(i, j)$ 。若  $x_i$  是  $x_j$  的  $K$  个近邻点之一, 说明邻域图  $G$  有边, 边长为  $d_x(i, j)$ 。

(b) 计算最短距离。当  $x_i$  和  $x_j$  之间有一条边, 则  $d_c(i, j) = d_x(i, j)$ ; 当  $x_i$  和  $x_j$  之间无边, 则  $d_c(i, j) = \infty$ ; 对  $K=1, 2, \dots, N$ ,  $d_c(x_i, x_j) = \min\{d_c(x_i, x_j), d_c(x_i, x_k) + d_c(x_k, x_j)\}$ ,  $D = \{d_c(x_i, x_j)\}$  是邻域图  $G$  中所有点的最短距离构成。

(c) 构造  $d$  维嵌入。根据多维尺度算法(MDS), 令  $\tau(D_c) = -2^{(-1)} \mathbf{H} \mathbf{D}_c \mathbf{H}$ , 其中,  $\mathbf{H}$  是与  $\mathbf{D}_c$  同阶的单位矩阵, 令  $\lambda_p$  是  $\tau(\mathbf{D}_c)$  的第  $p$  个特征值(降序),  $v_p^i$  是第  $p$  个特征向量的分量, 令  $d$  维嵌入向量  $\mathbf{y}$  的第  $p$  个分量等于  $(\lambda_p)^{1/2} v_p^i$ 。

### 1.2 信息熵方法

信息论中, 若某项指标提供的信息越多, 则对决策的精度和可靠性越大<sup>[20]</sup>。基于熵权理论构建台风灾情指数, 可以避免人为确定指标权重的主观性, 具有更高的客观性和科学性。熵权也可作为选取关键指标的办法, 主要计算步骤如下所示。

(a) 计算熵值  $H_j$

令  $f_{ij} = (1 + x_{ij}^*) / (\sum_i (1 + x_{ij}^*))$ ,  $H_j = -(\sum_i f_{ij} \times \ln f_{ij}) / \ln n$ ,  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, p$ 。其中,  $x_{ij}^* = (x_{ij} - x_{\min}) / (x_{\max} - x_{\min})$ ,  $x_{\min}$  和  $x_{\max}$  表示同一指标下的最小值和最大值。

(b) 计算熵权  $\omega_j$

$\omega_j = (1 - H_j) / (p - \sum_j H_j)$ ,  $i = 1, 2, \dots, n$ ;  
 $j = 1, 2, \dots, p$ 。 (1)

且满足  $0 \leq \omega_j \leq 1$ ,  $\sum_j \omega_j = 1$ 。

## 2 概率神经网络模型

鉴于台风灾情与变化的致灾气象因子和承灾体脆弱性之间是非线性关系, 复杂的相互作用和变化增加了预评估的困难。概率神经网络(PNN)是基于贝叶斯原理, 构造概率密度函数(PDF)分类估计, 在处理非线性问题和模式识别方面比 BP 等更具有显著优势<sup>[21-23]</sup>。PNN 的拓扑结构由输入层、模式层、求和层和输出层组成, 主要计算步骤如下所示。

(a) 输入层

将提取的承灾体关键指标作为输入层神经元, 神经元数目与输入样本维数相等。

(b) 模式层(神经元与给定类别以权值连接)

$$f_{ji}(X) = \frac{1}{(2\pi)^{P/2} \delta^P} \frac{1}{m} \sum \exp - \frac{(X - X_{ji})^T (X - X_{ji})}{2\delta^2}, \quad i = 1, 2, \dots, m。 \quad (2)$$

式中:  $\mathbf{X}$  为降维后提取的关键指标矩阵;  $\mathbf{X}_{ji}$  为类别  $j$  第  $i$  个训练向量;  $m$  为训练样本数目;  $\delta$  为平滑系数;  $P$  为待分类的向量  $X$  及训练向量的维数。

(c) 求和层(属于某一类别的概率累积)

$$S_j(\mathbf{X}) = \sum_{i=1}^{N_j} W_i f_{ji}(\mathbf{X})。 \quad (3)$$

式中:  $N_j$  表示类别  $j$  的神经元数目;  $W_i$  表示第  $i$  个连接的权重,  $\sum_{i=1}^{N_j} W_i = 1$ 。

(d) 输出层(竞争神经元)

PDF 最大的神经元输出为 1, 表示所对应的那一类为待识别的模式类别, 其他输出神经元输出为 0, 即:

$$y_j = \begin{cases} 1, & S_j = \max\{S_k, k = 1, 2, \dots, m\}; \\ 0, & \text{其他。} \end{cases} \quad (4)$$

## 3 模型应用与分析

### 3.1 数据来源及预处理

本研究的致灾源数据来自《热带气旋年鉴》中 1985-2013 年之间登陆或影响广西的 60 个台风致灾的降水过程数据, 1985-2013 年的社会经济情况数据取自《广西统计年鉴》。灾情数据来自广西气象信息中心、广西农业厅、广西民政厅和广西防汛抗旱指挥部的灾情综述统计。

为了消除指标间的量纲影响, 进行归一化处理:

$$x_{ij}^* = (x_{ij} - x_{\min}) / (x_{\max} - x_{\min})。 \quad (5)$$

式中:  $x_{\min}$  和  $x_{\max}$  表示同一指标下的最小值和最大值, 将指标数据范围压缩到 0~1 之间。

### 3.2 构造灾情指数与划分灾情等级

考虑到台风灾害对生命财产和社会经济的影响,选取受灾人口  $C_1$  (万人)、死亡人口  $C_2$  (人)、农作物受灾面积  $C_3$  (k $km^2$ )、倒塌房屋  $C_4$  (千间)和直接经济损失  $C_5$  (千万元)这5个指标作为构造灾情指数的关键指标。

台风灾情的综合评价方法是对各指标在灾情评估的权重进行确定,利用综合指数来衡量灾情的严重程度<sup>[24]</sup>。本文以信息熵构造综合灾情指数为:

$$I_i = \sum_j \omega_j \times x_{ij}^*, \quad i=1, 2, \dots, n; j=1, 2, \dots, p. \quad (6)$$

式中:  $I_i$ 为综合灾情指数;  $\omega_j$ 是第  $j$  个评价的熵权值;  $x_{ij}^*$ 是归一化的数据,计算结果见表1。

表1 基于熵权的台风灾情评估指标权重

指标	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
权重	0.274	0.157	0.197	0.171	0.200

由表1可以看出,受灾人口  $C_1$ 和直接经济损失  $C_5$ 的权重较大,其次是农作物受灾面积  $C_3$ 和倒塌房屋  $C_4$ 的权重,死亡人口  $C_2$ 的权重最小。受灾人口指的是因灾伤亡人数、因灾失踪人数等,直接经济损失指的是农业、林业、渔业和建筑及室内财产的损失<sup>[25]</sup>,二者是反映灾情的生命财产损失情况的核心指标,因此二者权重较大。而死亡人数由气象因子导致的具有偶然性,受灾区经济、居住环境等条件影响的程度更大<sup>[26]</sup>,因此其权重最小是客观合理。根据灾情指数的大小,我们将灾情划分为几个等级,这是灾后评价、灾情预估

进行救助与管理的重要依据。系统聚类方法基本思想是将个样品分成若干类,距离最小的一对合并成新一类,计算新类与其他类之间的距离,再将距离最近的两类合并,依次聚类直至所有的样品合为一类为止。本文参考马宗晋等<sup>[27]</sup>等级划分的思想,结合系统聚类方法,采用欧氏距离将灾情指数划分为5个等级(I级为微灾,II级为小灾,III级为中灾,IV级为大灾,V级为巨灾)(见表2),避免了主观划分灾情指数等级。

据广西区民政厅的灾情综述,2008年9月24-27日全区受强台风“黑格比”的影响,受灾人口664.99万人,农作物受灾面积656.570 k $km^2$ ,倒塌房屋19358间,直接经济损失高达69.7亿元,由表2可知,灾情指数为0.621,灾情等级为大灾(IV级),是受台风灾害很严重的灾情。2013年8月14-20日全区受热带风暴“尤特”的影响,造成农作物受灾面积59.28 k $km^2$ ,直接经济损失12.82亿元,灾情指数为0.529,灾情等级为大灾(IV级)。同年8月22日20时-25日20时,受台风“潭美”减弱后的环流和西南季风共同影响,造成9市30县(市、区)25.37万人受灾,农作物受灾面积15.18 k $km^2$ ,直接经济损失6600.82万元。“潭美”的移动速度快,在广西的持续降水时间与“尤特”相比要短得多<sup>[28]</sup>,受灾情况较为轻,由表2可知,灾情指数为0.165,灾情等级为微灾(I级)。上述表明该灾情等级划分合理,符合实际情况,可以作为灾后救助与管理的重要依据。

表2 广西1985-2013年台风灾情指数与灾情等级

过程编号	灾情指数/等级	过程编号	灾情指数/等级	过程编号	灾情指数/等级
1985/9/21	0.246/II	2006/5/24	0.105/I	2010/7/22	0.194/II
1986/7/21	0.282/II	2006/6/2	0.467/III	2011/5/7	0.112/I
1987/5/15	0.484/III	2006/6/12	0.359/III	2011/6/14	0.134/I
1988/8/21	0.758/IV	2006/7/15	0.671/IV	2011/6/27	0.224/II
1991/6/14	0.134/I	2006/8/3	0.652/IV	2011/9/29	0.442/III
1993/6/2	0.553/IV	2007/6/6	0.461/III	2012/5/11	0.171/II
1994/6/11	0.892/V	2007/6/13	0.287/II	2012/5/20	0.123/I
1994/7/13	0.850/V	2007/9/8	0.152/I	2012/6/10	0.118/I
1995/8/20	0.596/IV	2008/4/12	0.110/I	2012/6/13	0.124/I
1996/4/18	0.396/III	2008/5/27	0.212/II	2012/6/21	0.309/II
1997/7/2	0.651/IV	2008/6/8	0.773/IV	2012/6/26	0.449/III
1998/6/16	0.882/V	2008/8/6	0.206/II	2012/7/23	0.107/I
2000/6/8	0.430/III	2008/9/24	0.621/IV	2012/8/16	0.387/III
2000/6/19	0.313/II	2008/10/31	0.401/III	2012/10/27	0.167/I
2001/6/7	0.544/IV	2009/7/1	0.587/IV	2013/4/28	0.176/II
2001/6/26	0.298/II	2010/4/19	0.120/I	2013/5/8	0.128/I
2001/7/13	0.718/IV	2010/5/31	0.707/IV	2013/5/14	0.116/I
2003/6/26	0.466/III	2010/6/13	0.585/IV	2013/8/14	0.529/IV
2003/7/18	0.652/IV	2010/6/19	0.253/II	2013/8/22	0.165/I
2005/6/16	0.675/IV	2010/6/23	0.12/I	2013/11/10	0.408/III

### 3.3 基于 ISOMAP 降维、PCA 降维和熵权特征提取的 PNN 预评估模型

承灾体是指一个地区的社会经济和发展状况, 比如人员、农作物、房屋等方面的指标。因此, 本文选取单位面积 GDP  $B_1$  (亿元)、人口密度  $B_2$  (人/ $\text{km}^2$ )、农作物总播种面积  $B_3$  ( $\text{khm}^2$ )、城镇居民人均居住面积  $B_4$  ( $\text{m}^2$ )、农村居民人均生活用房面积  $B_5$  ( $\text{m}^2$ )、人均 GDP  $B_6$  (元/人)、就业人数  $B_7$  (万人)、每万人在校大学生人数  $B_8$  (人)、公路网密度  $B_9$  ( $\text{km}/10^4 \text{ km}^2$ )、每万人拥有床位  $B_{10}$  (床)、每万人拥有医生  $B_{11}$  (人)、电话普及率  $B_{12}$  (部/万人) 共 12 项指标。在致灾源方面, 本研究选取暴雨过程的时间长度  $A_1$  (h)、暴雨过程降水极值  $A_2$  (mm) 和暴雨过程降水均值  $A_3$  (mm)。

通过 ISOMAP 方法进行非线性降维。在构造邻域图  $G$  图时, 采用  $K$ -近邻方法确定样本邻域, 为了保持图连通性, 确定最小值  $K=8$ , 用 Floyd 算法计算最短距离, 作为测地距离的逼近, 降维后的残差曲线图如图 1 所示。

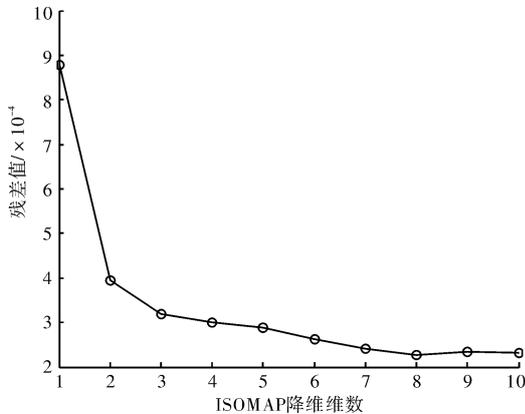


图1 ISOMAP 降维的残差曲线图

ISOMAP 算法降维维数的方法一是当残差曲线出现拐点, 方法二是残差值小于一定的阈值<sup>[12]</sup>。由图 1 可知, 当维数降到 3 维时, 残差曲线出现明显拐点, 且残差值为  $3.199 \times 10^{-4} < 0.05$ , 确定采用 ISOMAP 降维后的 3 维向量代表承灾体原始矩阵的多维向量。

采用主成分分析 (PCA) 对承灾体原始矩阵进行线性降维, 见表 3。主成分 1 和主成分 2 的累积贡献率是  $96\% > 80\%$ , 符合主成分降维的要求, 将标准化后的原始数据代入主成分表达式得到承灾体主成分得分矩阵。

表 3 基于 PCA 降维的特征值、贡献率和累积贡献率

	特征值	贡献率	累积贡献率
主成分 1	10.050	0.838	0.838
主成分 2	1.462	0.122	0.960

采用信息熵计算承灾体各个指标的权重, 计算结果见表 4, 选取对承灾体系统起重要影响的指

标 (权重值  $> 0.1$ ) 为单位面积 GDP  $B_1$  (亿元)、每万人拥有床位  $B_{10}$  (床) 这 2 项关键指标。

表 4 基于熵权的承灾体指标提取

指标	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	$B_6$
权重	0.104	0.057	0.055	0.069	0.075	0.097
指标	$B_7$	$B_8$	$B_9$	$B_{10}$	$B_{11}$	$B_{12}$
权重	0.061	0.093	0.084	0.112	0.096	0.097

分别将 ISOMAP 降维后的三维向量、PCA 降维后的主成分得分矩阵和信息熵提取的关键指标, 联合致灾源数据作为神经网络的输入神经元 (矩阵  $p$ ), 将灾情等级作为期望输出 (矩阵  $t$ ), 进行概率神经网络的分类训练和预测。Matlab 创建 PNN 网络的调用函数为  $\text{net} = \text{newpnn}(p, t, \text{spread})$ , 其中,  $\text{spread}$  为网络的扩展速度,  $\text{spread}$  值过大, 需较多的神经元适应网络快速变化, 计算效率差;  $\text{spread}$  值过小, 需较多的神经元适应网络缓慢变化, 网络训练性能差。经过多次实验, 最终确定  $\text{spread}$  取 0.1, 创建的 PNN 是个近邻分类器, 训练准确度高。PNN 具有网络训练性能好, 训练不需大量样本等优点, 故本文选取 10 组登陆或影响广西的重要台风灾害作为测试样本, 其他 50 组样本作为训练样本, PNN 预测结果如图 2~4 所示。

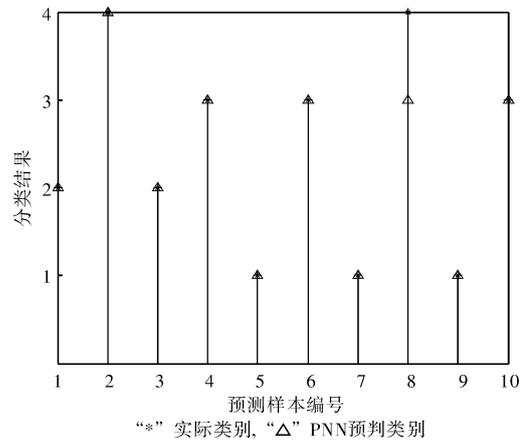


图 2 基于 ISOMAP 降维的 PNN 预评估结果

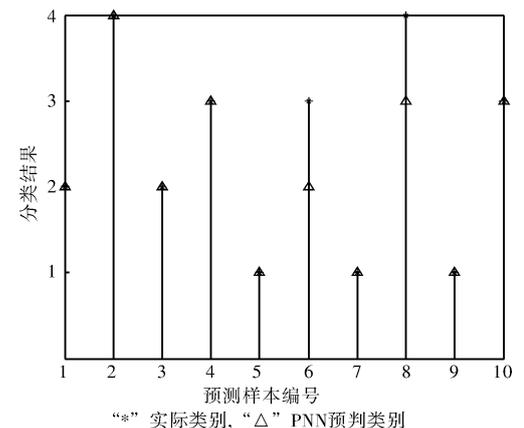


图 3 基于 PCA 降维的 PNN 预评估结果

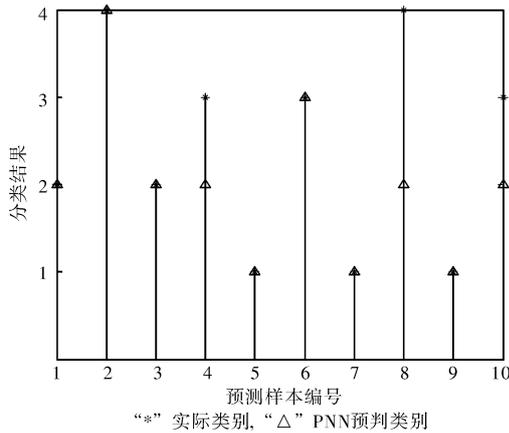


图4 基于信息熵特征提取的 PNN 预评估结果

由图 2 可知，在 10 组预测样本中，基于 ISOMAP 降维的 PNN 预评估只有 8 号样本(1311 号台风)误判，其他样本预判类别和实际类别重合说明预判准确，分类效果最好。由图 3 可知，基于主成分分析降维的 PNN 预评估中，6 号样本(1213 号台风)和 8 号样本(1311 号台风)误判，其他样本预判准确。由图 4 可知，基于信息熵特征提取的重要指标的 PNN 预评估模型，4 号样本(1117 号台风)、8 号样本(1311 号台风)和 10 号样本(1330 号台风)误判，其他样本预判准确。基于三种变量提取方法的 PNN 预评估结果对比如表 5 所示。

根据表 5 可知，基于 ISOMAP 降维的 PNN 预测灾情等级的准确率最高(90%)，运行时间最短(0.227 1 s)，预评估效果最好。造成不同预测效果的原因是，PNN 模型输入指标提取的方法将直接或间接影响预评估结果的准确度，因此要深入分析致灾源、承灾体与灾情之间的信息传递，有效的特征选取方法能提高 PNN 预评估的精准度。

在三种特征选取方法中，基于信息熵特征提取模型，因只选取对承灾体系统的 2 项关键指标对

预评估会造成一定的信息损失，虽然计算效率有所提高，但对预评估的精确度有一定影响。基于主成分特征提取的模型，提取了承灾体系统线性相关度高的特征分量，比信息熵方法减少信息损失，但忽略了高维数据内部的非线性结构。实验结果表明，ISOMAP 比 PCA 更好的地将高维数据的非线性流形结构挖掘，保留更多原始信息，计算效率最高，采用 ISOMAP 算法对承灾体指标进行非线性降维过程中，邻域  $K$  的取值和降维维数  $n$  是决定其降维效果优劣的重要参数<sup>[17]</sup>， $K$  取值过大会导致数据集变成局部邻域， $K$  取值过小会导致邻域图不连通，本实验  $K$  取最优值 8。在  $K$  取定值情况下，由残差曲线图的明显拐点确定降维维数  $n$ 。若  $n$  值过大，会增加数据冗余度，若  $n$  值过小，会导致数据集分离部分被映射到一起。本实验由残差曲线图确定  $n$  为 3，降维效果最好。

此外，我们注意到 1311 号台风“尤特”在三种特征选取的 PNN 预评估模型均预测错误，可视作异常点分析。1311 号台风“尤特”以热带风暴强度进入广西，先后造成 15 日(15 日 08 时—16 日 08 时)和 18 日(18 日 08 时—19 日 08 时) 2 个特大暴雨，部分暴雨落区重叠，造成当地重大生命财产损失<sup>[29]</sup>。根据广西民政厅统计，台风“尤特”造成全区直接经济损失 12.82 亿元，153.17 万人受灾，灾情等级划分为大灾(IV 级)较合理。

### 4 结论与展望

(1) 基于概率神经网络的灾情等级识别是利用其强大的非线性处理能力，将致灾因子和承灾体的特征空间映射到灾情等级类型空间中，形成了一个较强容错能力和自适应能力的识别网络系统，网络训练不需大量样本，总收敛于 Bayes 优化解，

表 5 基于三种方法选指标的 PNN 预评估结果

预测样号	过程编号	台风编号	ISOMAP 降维 PNN 分类结果	PCA 降维 PNN 分类结果	基于信息熵选指标 PNN 分类结果
1	2008/8/6	0809	II(正)	II(正)	II(正)
2	2008/9/24	0814	IV(正)	IV(正)	IV(正)
3	2010/7/22	1003	II(正)	II(正)	II(正)
4	2011/9/29	1117	III(正)	III(正)	II(错)
5	2012/7/23	1208	I(正)	I(正)	I(正)
6	2012/8/16	1213	III(正)	II(错)	III(正)
7	2012/10/27	1233	I(正)	I(正)	I(正)
8	2013/8/14	1311	III(错)	III(错)	II(错)
9	2013/8/22	1312	I(正)	I(正)	I(正)
10	2013/11/10	1330	III(正)	III(正)	II(错)
预测准确率/%			90	80	70
运行时间/s			0.227 1	0.250 1	0.234 4

稳定性高。

(2)在采用概率神经网络进行台风灾情等级预评估模型时,模型输入的影响因子选择是一个重要问题,选择起关键作用的因子,提高计算效率和模型预评估的精确度。主成分分析方法本质是个二阶统计特征的线性映射方法,而很多自然灾害系统的数据信息都包含高阶非线性关系中。基于等距映射的非线性特征提取方法可以充分有效地挖掘预报因子的信息,算法效率高,具有全局优化性和渐近收敛性的优点,且在台风灾情评估指标选取方面鲜少研究。

(3)ISOMAP 算法适用于自然灾害系统的指标降维。本研究基于等距映射的非线性特征提取建立概率神经网络预评估模型,该模型预测精度高,对台风可能造成的灾害进行有效的预评估,有助于相关部门更好的对灾情的预防与救助做出有效的决策。

(4)采用信息熵构造灾情指数,对灾情指数采用系统聚类方法进行灾情划分等级,克服了人为划分灾情等级的主观随意性。

(5)本研究对异常极值点的预测仍存在误差,还需对台风灾情预评估模型进一步优化。台风灾害系统是由致灾源的强度、承灾体的脆弱性及防灾减灾应急能力等因素共同作用的结果,关于台风灾情的预评估工作还需考虑更多实际的影响因素,例如台风登陆路径、最大风速和最低气压等预报因素,还可以将城市护林、排水和城市建设实施的损害情况等一并考虑到灾情预评估的研究中。

## 参考文献:

- [1] 温克刚, 杨年珠. 中国气象灾害大典 - 广西卷[M]. 北京: 气象出版社, 2007.
- [2] 黄雪松, 廖雪萍, 覃卫坚. 广西热带气旋特征变化与灾损变化态势[J]. 气象研究与应用, 2014, 35(1): 2-6.
- [3] 彭昱忠, 王谦, 元昌安, 等. 数据挖掘技术在气象预报研究中的应用[J]. 干旱气象, 2015, 33(1): 19-27.
- [4] 叶雯, 刘美南, 陈晓宏, 等. 基于模式识别的台风风暴潮灾情[J]. 海洋通报, 2004, 23(4): 65-70.
- [5] 刘合香, 简茂球. 基于粒子群 - 投影寻踪和遗传 - 神经网络集成的预测模型[J]. 中山大学学报: 自然科学版, 2012, 51(5): 113-119.
- [6] 陈佩燕, 杨玉华, 雷小途. 我国台风灾害成因分析及灾情预评估[J]. 自然灾害学报, 2009, 18(1): 64-73.
- [7] 徐明, 雷小途, 杨秋珍. 应用联合极值分布评估热带气旋影响风险——以“海葵”对上海地区影响为例[J]. 灾害学, 2014, 29(3): 124-130.
- [8] 吴先华, 徐中兵, 袁迎蕾, 等. 台风灾害的关联经济损失评估——以江苏省为例[J]. 灾害学, 2014, 29(2): 77-83.
- [9] 魏章进, 隋广军, 唐丹玲. 基于聚类与回归方法的台风灾情统计评估[J]. 数理统计与管理, 2014, 33(3): 400-407.
- [10] 芮建勋, 张发勇, 鲍曙明, 等. 面向台风事件与灾害影响评估的时空数据管理模式[J]. 灾害学, 2015, 30(3): 43-46.
- [11] 刘合香, 秦川, 倪增华. 组合权重和 ArcGIS 相结合的广西洪涝灾害风险分析[J]. 灾害学, 2015, 30(3): 76-79.
- [12] 刘少军, 张京红, 何政伟, 等. 基于 GIS 的台风灾害损失评估模型研究[J]. 灾害学, 2010, 25(2): 64-67.
- [13] Blaikie P, Cannon T, Davis I, et al. At Risk: Natural Hazard, People's Vulnerability and Disasters[M]. London: Routledge, 1994: 210.
- [14] 现在武, 胡丽. 台风灾害评估中的影响因子分析[J]. 自然灾害学报, 2015, 24(1): 203-213.
- [15] 倪艳. Isomap 算法在地震属性参数降维中的应用[J]. 西南民族大学学报: 自然科学版, 2008, 34(2): 397-400.
- [16] 黄颖, 金龙, 黄小燕, 等. 基于局部线性嵌入的人工智能台风强度集合预报模型[J]. 气象, 2014, 40(7): 806-815.
- [17] Tenenbaum J B, Silva V, Langford J C. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290(5500): 2319-2323.
- [18] 翁时锋, 张长水, 张学工. 非线性降维在高维医学数据处理中的应用[J]. 清华大学学报: 自然科学版, 2004, 44(4): 485-488.
- [19] 尹焕. 基于 ISOMAP 的机械故障诊断方法研究与应用[D]. 广州: 华南理工大学, 2012: 15-22.
- [20] 张星. 自然灾害灾情的熵权综合评价模型[J]. 自然灾害学报, 2009, 18(6): 189-191.
- [21] Specht D F. Probabilistic neural networks[J]. Neural Networks, 1990, 3(1): 109-118.
- [22] 郭联金, 罗炳军. PNN 与 BP 神经网络在钢板表明缺陷分类中的应用研究[J]. 机电工程, 2015, 32(3): 352-357.
- [23] 王雨虹, 付华, 张洋, 等. 基于 KPCA 和 CIPSO-PNN 的煤与瓦斯突出强度辨识模型[J]. 传感技术学报, 2015, 28(2): 271-277.
- [24] 胡永宏, 贺思辉. 综合评价方法[M]. 北京: 科学出版社, 2000.
- [25] 王志强, 陈思宇, 吕雪峰, 等. 风暴潮灾害受灾人口与直接经济损失评估方法研究进展[J]. 北京师范大学学报, 2015, 51(3): 274-279.
- [26] 陈舜华, 吕纯濂, 李吉顺. 福建省台风灾害评估试验[J]. 中国减灾, 1994, 4(3): 31-34.
- [27] 赵阿兴, 马宗晋. 自然灾害损失评估指标体系的研究[J]. 自然灾害学报, 1993, 2(3): 1-7.
- [28] 张芳琳. “潭美”入桂待一天暴雨集中桂中北[N]. 南宁晚报, 2013-8-24(5).
- [29] 陈见, 赖珍权, 罗小莉, 等. “尤特”超强台风残留低涡引发的广西特大暴雨成因分析[J]. 暴雨灾害, 2014, 33(1): 20-25.

- [13] 《铁路安全管理条例》(国务院令第 639 号)[Z]. 北京: 国务院, 2013.
- [14] 《铁路主要技术政策》(铁道部令第 34 号)[Z]. 北京: 铁道部, 2013.
- [15] 《关于印发高速铁路防灾安全监控系统管理办法(暂行)的通知》(铁运[2010]28 号)[Z]. 北京: 铁道部, 2010.
- [16] 《中国铁路总公司关于印发〈铁路技术管理规程〉(高速铁路部分)的通知》(铁总科技[2014]172 号)[Z]. 北京: 中国铁路总公司, 2014.
- [17] 《长春工务段 2012 高速铁路降雪应急处置预案》(长工高发[2012]296 号)[Z], 长春: 长春工务段, 2012.
- [18] 《关于印发铁路暴风雨雪雾等恶劣天气应急预案(暂行)的通知》(铁运[2008]236 号)[Z]. 北京: 铁道部, 2008.
- [19] 《关于印发冰雪天气动车组列车限速暂行规定的通知》(铁运[2011]17 号)[Z]. 北京: 铁道部, 2011.
- [20] 日本法国德国高速铁路运营安全专题资料汇编[R]. 北京: 铁道部, 2010: 249-252.
- [21] 胡立庆. 气象与铁路运输安全[J]. 铁道运输与经济, 1984, 4(1): 35-38.
- [22] 姜强, 薛洁. 铁路大风监测与行车安全试验[J]. 新疆气象, 2006, 29(3): 25-27.
- [23] 刘丽霞. 高速铁路防灾气象监测系统设计[J]. 计算机测量与控制, 2010, 18(9): 36-38.
- [24] 杨靖波, 李正, 杨风利, 等. 2008 年电网冰灾覆冰及倒塔特征分析[J]. 电网与水利发电进展, 2008, 24(4): 4-8.
- [25] 拉有玉, 李永乐, 何向东, 等. 兰新铁路第二双线防风技术及工程设计[J]. 石家庄铁道大学学报(自然科学版), 2010, 23(4): 104-108.
- [26] 杨加伦, 朱宽军, 刘彬, 等. 输电线路冰区分布图绘制关键技术[J]. 电力建设, 2013, 34(9): 31-36.

## Research on Level Standards of High-impact Weather over High-speed Railway Safety Operation

CUI Xinqiang, ZHOU Xiaolan, FU Jia, DAI Juan and LIU Jin  
(Hubei Meteorological Service Center, Wuhan 430205, China)

**Abstract:** In order to ensure high-speed railway in safe operation of the bad weather conditions. Based on the survey of specification of high-speed railway safety operation & management and related standards, the papers have built the Level Standards of Weather Conditions Impacting by lightning, wind, precipitation, snowfall, haze, dust storms temperature, wire icing and hail and so forth over High-speed railway after integrating the achievements of meteorological research. Moreover, the application test results show that our works can provide the certain reference to the operational safe management and meteorological service Assurance for High-speed railway.

**Key words:** high-speed railway; safety operation; standards of weather; level

.....

(上接第 25 页)

## Probabilistic Neural Network Pre-Assessment Model Based on Isometric Feature Mapping Dimensional Reduction in Typhoon Disaster

CHEN Yanxuan, LIU Hexiang and TAN Jinkai  
(College of Mathematics and Statistics Sciences, Guangxi Teachers Education University,  
Nanning 530023, China)

**Abstract:** Typhoon hazard, between hazard bearing body and the disaster is a complex nonlinear dynamical system; accurately and efficiently extract the important indicators for the pre-assessment of typhoon disaster grade is an important basis for disaster prevention and relief work. In this paper, we apply principal component analysis, isometric feature mapping and entropy to extract key indicators of hazard bearing body, with hazard source as the input neurons, and disaster grade as output neurons, establishing probabilistic neural network pre-assessment model in typhoon disaster. The results show that the accuracy of probabilistic neural network pre-assessment model based on the non-linear feature extraction isometric feature mapping reaches 90%, the model has a satisfactory level of accuracy and generalization ability, provide a new way for natural disaster risk assessment, having certain reference value.

**Key words:** probabilistic neural network; Isometric Feature Mapping; entropy; typhoon; disaster; pre-assessment