

Urban Waterlogging Risk Assessment Based on Projection Pursuit Dynamic Clustering Model

Jiangtao Wang

Shanghai Maritime University, Shanghai
Email: wangjt12315@163.com

Received: Mar. 20th, 2019; accepted: Apr. 2nd, 2019; published: Apr. 9th, 2019

Abstract

Urban flood risk assessment is critical to the prediction and reduction of flood disaster, and also the mitigation of drainage facilities. However, for flood risk assessment using multiple indicators, the index weight has a great impact on the results of flood risk assessment. In this paper, projection pursuit dynamic cluster method is used in order to reduce the uncertainties of the index weight. Projection pursuit dynamic cluster is a relatively objective method, which can evaluate the sample data comprehensively, and can also consider the preference of decision makers by setting constraint conditions. The SWMM model was established for a drainage system. Then different values for the evaluation index were got through simulation. Then the results of urban waterlogging risk assessment are obtained by the dynamic iteration between clustering center and relative membership degree. Through the analysis of the results, it shows that projection pursuit dynamic cluster method can be used for urban waterlogging risk assessment. Furthermore, it brings important guiding significance for urban disaster prevention and mitigation.

Keywords

Drainage System, Projection Pursuit, Dynamic Cluster, Variable Fuzzy Clustering, Risk Assessment

基于投影寻踪动态聚类模型的城市内涝风险评估

王江涛

上海海事大学, 上海
Email: wangjt12315@163.com

收稿日期: 2019年3月20日; 录用日期: 2019年4月2日; 发布日期: 2019年4月9日

摘要

城市内涝风险评估对于城市防灾减灾以及城市排水设施改造是至关重要的。采用多个指标对城市内涝风险进行评估时, 指标权重对风险评估结果具有重大影响。因此, 针对风险评估中指标权重模糊不确定性问题, 本文基于投影寻踪动态聚类方法对城市内涝风险进行评估。投影寻踪动态聚类是一种对样本数据进行综合评价的客观方法, 并能通过设置约束条件来考虑决策者的偏好。选取某排水区建立SWMM模型进行实例研究, 模拟计算得到评价指标值, 投影寻踪计算得到不同区域的风险等级。通过对结果进行分析可知, 投影寻踪动态聚类可以用于城市内涝风险评估, 而且对城市防灾减灾具有重要指导意义。

关键词

排水系统, 投影寻踪, 动态聚类, 可变模糊聚类, 风险评估

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

城市内涝灾害是与城市化进程一致的, 并且与人类活动和洪水特性有关; 洪涝灾害不仅对城市造成严重影响, 而且还对整个社会发展造成较长时期的不良影响, 甚至影响社会的稳定[1]。我国 2009 年的城市化率已达到 45%, 进入快速发展阶段。随着城市化的快速发展, 近年来, 我国许多城市出现了严重的内涝, 对国家和人民的生命财产造成了重大损失。如 2012 年 7 月, 北京遭遇 61 年来最强特大暴雨, 导致至少 79 位市民失去生命, 160.2 万人受灾, 经济损失达 116.4 亿元[2]。2014 年 3 月底, 深圳“3.30”暴雨造成城市内涝 200 处, 2 人因灾死亡。内涝灾害频繁发生是由许多原因造成的, 一方面, 由于气候变化导致极端降雨频现; 另一方面, 城市化发展进程的不断加快导致下垫面不透水层比例增大, 即径流系数增大, 雨水下渗量减少, 地面截流作用变差, 地表径流的汇流时间变短, 洪峰提前, 并且城市的扩张, 也使得原有管道系统的汇水面积增大, 导致原有管道的负荷增大。

21 世纪, 城市综合防灾减灾能力大小将成为衡量城市整体功能及其安全防卫能力的标志[3]。城市内涝灾害是由极端暴雨在短时间内超过雨水系统设计标准且低于城市防洪标准所导致的, 具有持续时间短等特点。因此, 对城市内涝形成机理和风险评估开展研究, 为城市洪涝灾情预判工作提供及时有效的数据及信息, 根据不同重现期的内涝灾害制定应急预案, 对提高城市防灾减灾能力具有重要意义。

目前, 有很多学者针对城市内涝风险评估进行了研究和应用, 常用的研究方法主要包括定性分析方法和定量分析方法, 如基于历史灾情数理统计方法、基于情景进行模拟分析方法、基于遥感数据分析方法以及多指标体系评价方法等[4]。黄维选取海口市主城区为研究区域, 借助 ArcGIS 和 InfoWorks ICM 进行模拟计算, 综合考虑淹没深度和淹没时间两个因素对研究区的内涝风险进行评估, 得到了不同频率设计暴雨重现期情景下的内涝风险图[5]。戴晶晶等以苏州市城市中心区为例, 采用数值模拟的手段评估识别内涝风险, 为城市内涝治理及科学决策提供依据[6]。黄铁兰等从致灾因子危险性、孕灾环境敏感性、承载体易损性及防灾减灾能力 4 个方面选取评价指标, 构建了城市内涝风险评价模型进行城市内涝风险

评估, 为城市内涝防治提供了科学参考[7]。王磊等采用投影寻踪方法对城市内涝风险进行了评估, 并对影响因素的影响程度进行了分级比较[8]。

然而, 城市内涝灾害受多种因素影响, 例如气象、城市下垫面等。受主观性影响, 不同影响因素的权重对城市内涝风险评估结果影响较大。为此, 基于改进的投影寻踪方法, 本文提出将投影寻踪动态聚类模型用于城市排水系统内涝风险评估。在构建投影寻踪动态聚类评估模型的基础上, 根据最优投影方向向量确定评估体系各指标权重。

2. 投影寻踪动态聚类原理

投影寻踪法是基于降维思想的一种数据分析和处理方法, 最早由 Friedman 于 1974 年提出, 把高维数据向低维空间进行投影, 方便处理多因素的复杂问题[9] [10]。针对投影寻踪中密度宽度参数的不确定问题, 倪长健等提出了投影寻踪动态聚类方法, 有效避免了投影寻踪模型的不足[11] [12]。投影寻踪动态聚类模型和建立步骤如下:

将第 i 个样本的第 j 个指标表示为 $x_{ij}^0 (i=1, \dots, n; j=1, \dots, m; n$ 为样本数, m 为指标数)。

1) 对各指数据进行无量纲化处理, 即归一化, 如公式(1)~(2)所示。

对于越大越优的指标

$$x_{ij} = \frac{x_{ij}^0 - x_{j\min}^0}{x_{j\max}^0 - x_{j\min}^0} \quad (1)$$

对于越小越优的指标

$$x_{ij} = \frac{x_{j\max}^0 - x_{ij}^0}{x_{j\max}^0 - x_{j\min}^0} \quad (2)$$

其中, $x_{j\max}^0$ 和 $x_{j\min}^0$ 分别为第 j 个指标的样本最大值和最小值。

2) 投影可以从不同角度观察数据, 寻找能够最大程度地反映高维数据特征和最能充分挖掘数据信息的最佳观察角度即最优投影方向。本文采用线性投影, 即将高维数据投影到线性空间。设 a 为 m 维单位投影方向向量, 其分量为 a_1, a_2, \dots, a_m , 则 x_{ij} 的投影特征值 z_i 可表示为

$$z_i = \sum a_j x_{ij} \quad (i=1, \dots, n) \quad (3)$$

3) 本文应用动态聚类方法来构造投影寻踪动态聚类模型的投影指标。假设整个样本的投影特征值序列组成的集合为 $\Omega = \{z_1, z_2, \dots, z_n\}$, 依据实际情况或要求, 采用动态聚类法将其聚为 p ($p \leq n$) 类[13], 具体实现步骤如下:

①随机选取 p 个点作为 p 个聚核, 记为

$$L^0 = (A_1^0, A_2^0, \dots, A_p^0) \quad (4)$$

②根据 L^0 , 把 Ω 中的点分为 p 类, 记为

$$\Theta^0 = (\Theta_1^0, \Theta_2^0, \dots, \Theta_p^0) \quad (5)$$

其中 $\Theta_i^0 = \{z \in \Omega | d(A_i^0 - z) \leq d(A_j^0 - z), j=1, 2, \dots, p, j \neq i\}$, $d(A_i^0 - z)$ 为点 A_i^0 和集合 Ω 中任一点的绝对值距离。

③从 L_0 出发, 得到新的聚核 L^1

$$L^1 = (A_1^1, A_2^1, \dots, A_p^1) \quad (6)$$

④重复以上步骤, 由此得到一个分类结果序列 $V^k = (L^k, \Theta^k), k=1, 2, \dots$, 记

$$D(A_i^k, \Theta_i^k) = \sum_{z_i \in \Theta_i^k} |z_i - A_i^k| \quad (7)$$

$$u_k = \sum_{i=1}^p D(A_i^k, \Theta_i^k) \quad (8)$$

当 $\frac{|u_{k+1} - u_k|}{u_{k+1}} \leq \varepsilon$ 时终止计算, 其中 ε 是一个充分小的误差允许值。

$\Theta_h (h=1, 2, \dots, p)$ 为由属于第 h 类的所有样本投影特征值构成的集合, 定义 $d(z_i, z_j)$ 为任意两投影特征值间的绝对值距离。为了表示类内样本的邻近程度, 则用如下公式表示

$$dd(a) = \sum_{h=1}^p D_h(a) \quad (9)$$

$$D_h(a) = \sum_{z_i, z_j \in \Theta_h} d(z_i, z_j) \quad (10)$$

其中 $dd(a)$ 愈小, 则类内样本的聚集程度越高。为了表征有样本序列投影特征值的离散程度, 则记投影分散度为

$$ss(a) = \sum_{z_i, z_j \in Q} d(z_i, z_j) \quad (11)$$

根据 $ss(a)$ 和 $dd(a)$ 的定义, 提出投影寻踪动态聚类模型的投影指标 $QQ(a)$ 表示为

$$QQ(a) = ss(a) - dd(a) \quad (12)$$

投影寻踪动态聚类模型的投影指标定义为投影分散度与类内聚集度之差, 表示类间所有样本投影特征值的距离之和。显然, 投影分散度越大或类内聚集度越小, 则投影指标越大。当 $QQ(a)$ 取得最大值时, 就实现了类间样本尽量散开、类内样本尽量集中的聚类目的。

4) 模型求解。

本文采用遗传算法求解如下优化问题, 即寻找最能反映数据特征的最优投影方向向量 a^* , 满足 $QQ(a)$ 取得最大值。

$$\begin{cases} \max QQ(a) \\ \|a\| = 1 \end{cases} \quad (13)$$

3. 实例分析

借助美国 EPA 的 SWMM 模型进行排水管网系统水动力模拟计算。SWMM 模型是美国环保局为了设计和管理城市暴雨而研制的综合性数学模型, 它可以模拟完整的城市降雨径流和污染物运动过程, 被很多国家和地区的研究者用于排水系统设计规划和管理中[14] [15] [16]。SWMM 模型包括径流模块、输送模块、扩展的输送模块、调蓄/处理模块和接纳水体模块等主要模块。选取某排水系统为例具体说明城市洪水风险评估过程。该研究区域面积约为 34.11 hm², 区域内排水管网系统分布如图 1 所示。该排水系统由 69 个排水管井、72 根排水管道、1 个排水口组成。采用 2014 年 8 月 6 日 6 h 暴雨过程(图 3)作为模型的降雨输入进行水动力学模拟计算, 获得排水系统的水力状况, 某时刻部分管线的水位图如图 2 所示。选取评价指标, 并对各种指标进行处理分析。本文选用管井最大积水深度、最大洪水流量、洪水总水量、洪水持续时间、集水区不透水面积、集水区总面积以及所在道路等级等 7 个评价指标[8], 指标样本集共有 69 个样本, 部分样本数据如表 1 所示。

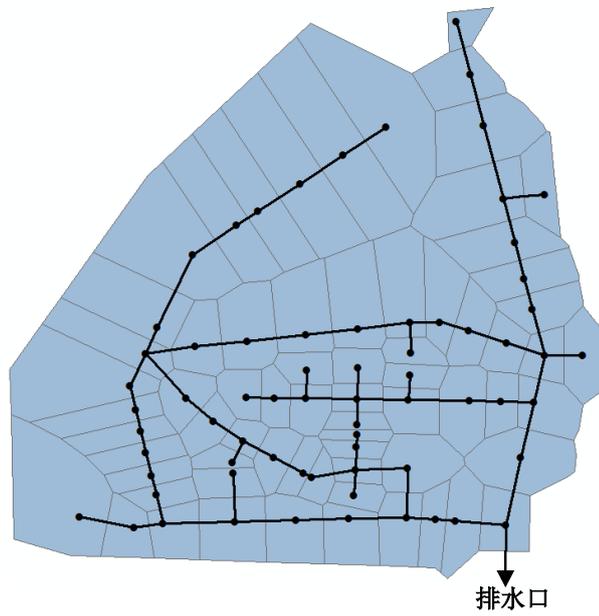


Figure 1. Drainage system
图 1. 某排水系统

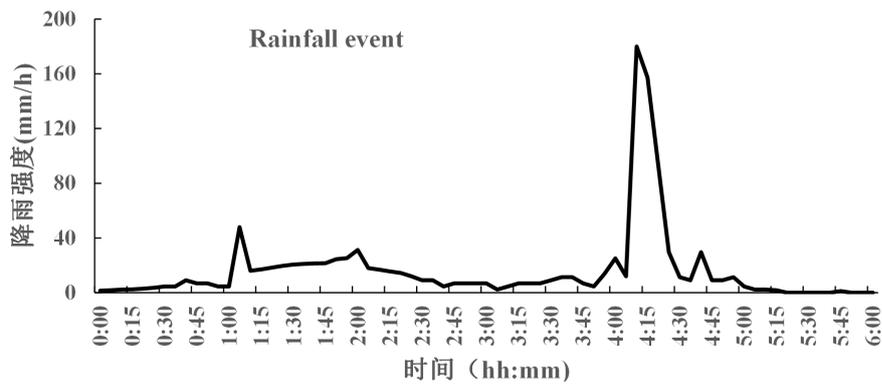


Figure 2. 6 h rainfall hyetographs of the 6 August 2014 event
图 2. 2014年8月6日6h暴雨过程线

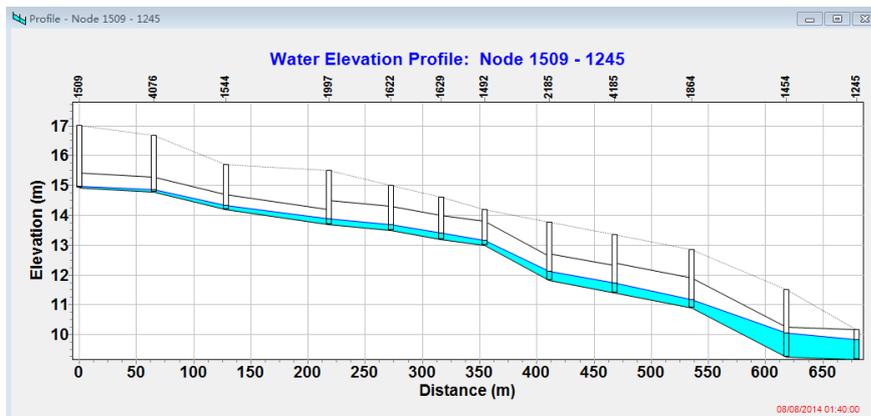


Figure 3. Water level diagram of some pipes at 01:40:00
图 3. 管线在 01:40:00 时刻的水位图

Table 1. Flood risk evaluation factor sample set and values of projection and risk degree
表 1. 城市洪水风险评估指标集及投影值、风险等级值

管井编号	最大积水深度	最大洪水流量	洪水总水量	洪水持续时间	集水区不透水面积	集水区总面积	所在道路等级	投影值	风险等级
J1	0.282	0.426	0.262	0.31	0.251	0.184	1	1.11	2
J2	0.203	0.436	0.217	0.283	0.431	0.322	1	1.1	2
J3	0.134	0.347	0.144	0.186	0.282	0.208	1	0.99	2
J4	0.215	0.362	0.219	0.266	0.305	0.225	1	1.06	2
J5	0.314	0.327	0.294	0.343	0.11	0.076	1	1.09	2
J6	0.169	0.299	0.173	0.321	0.095	0.064	0.5	0.61	2
J7	0.079	0.177	0.081	0.163	0.388	0.289	0.5	0.57	2
J8	0.331	0.319	0.306	0.357	0.138	0.097	1	1.11	2
J9	0.151	0.198	0.137	0.255	0.048	0.028	0	0.18	1
J10	0.302	0.321	0.281	0.335	0.174	0.124	1	1.09	2
J11	0.059	0.154	0.057	0.23	0.149	0.105	0.5	0.5	2
J12	0.29	0.352	0.268	0.324	0.187	0.135	1	1.09	2
J13	0.228	0.448	0.214	0.258	0.961	0.729	1	1.24	2
J14	0.168	0.306	0.168	0.271	0.182	0.16	0.5	0.62	2
J15	0.282	0.426	0.262	0.31	0.251	0.184	1	0.48	1
...

对上述 69 个排水管井的各项指标进行归一化处理, 并采用投影寻踪动态聚类模型进行计算分析, 即将上述指标分别代入式(1)~(12), 建立投影寻踪动态聚类模型, 然后采用遗传算法对式(13)进行优化求解。经过迭代计算, 目标函数值收敛, 且最佳投影方向向量 $a^* = (0.596, 0.185, 0.104, 0.130, 0.150, 0.115, 0.739)$ 。最佳投影方向 a^* 反映了各指标重要程度的大小, 即为客观权重, 由此表明道路等级和最大积水深度是影响此区域内涝风险等级的主要因素。将 a^* 代入式(3)得到各个排水管井的样本投影值, 如图 4 为不同管井编号的投影值, 投影值越大, 则内涝风险程度越高。由图 4 可以得知投影值聚类结果较好, 得到了明确的聚类结果, 内涝高、中、低风险区很容易辨识, 即当投影值 $1.5 \leq z_i \leq 2.0$ 时, 内涝风险分级为高风险区(3 级); 当 $0.5 \leq z_i < 1.5$ 时, 内涝风险分级为中风险区(2 级); 当 $0 \leq z_i < 0.5$ 时, 内涝风险分级为低风险区(1 级)。通过 GIS 分析得到了 2014 年 8 月 6 日 6 h 暴雨条件下城市内涝风险分布图, 如图 5 所示。内涝风险高风险区如图中深色区域所示, 主要分布在道路等级为一级道路区, 对积水较为敏感, 且此区域集水区域面积较大, 当发生上述暴雨时, 管井易溢水。因此, 内涝风险区划结果与实际情况相符。

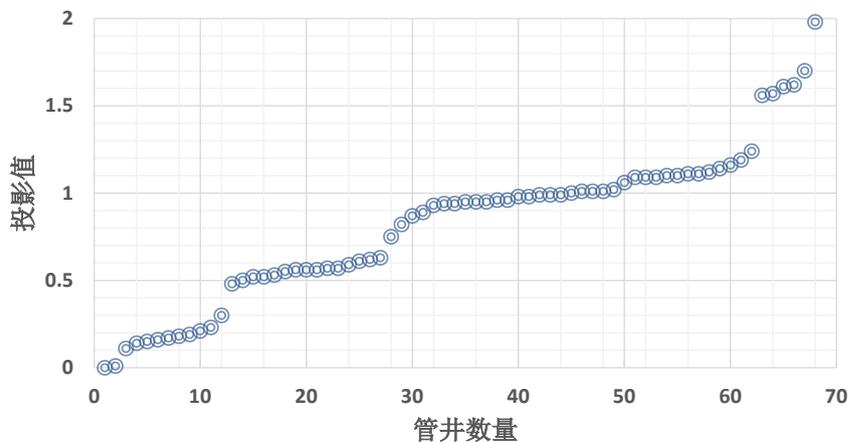


Figure 4. Projection values of different wells
图 4. 不同管井编号的投影值



Figure 5. Flood risk map of the study area
图 5. 研究区城市内涝风险图

4. 结语

本文通过多指标分析对城市排水系统进行内涝风险评估，可以有效应用于城市防灾减灾。通过实例研究发现投影寻踪动态聚类模型可以用来处理多因素复杂问题的聚类分析。投影寻踪动态聚类模型通过遗传算法求解非线性优化问题可以有效避免投影寻踪密度宽度不确定的问题，这样求得的各指标权重较为客观，在一定程度上减轻了赋权对风险评估的影响。所以，投影寻踪动态聚类模型为城市内涝风险多指标评估提供了一种有效方法。然而，不同城市区域风险等级数目的确定有待进一步研究。

参考文献

- [1] 周振民. 城市水务学[M]. 北京: 科学出版社, 2013.
- [2] 谢映霞. 从城市内涝灾害频发看排水规划的发展趋势[J]. 城市规划, 2013(2): 45-50.
- [3] 金磊. 城市灾害学原理[M]. 北京: 气象出版社, 1997.
- [4] 孙章丽, 朱秀芳, 潘耀忠, 等. 洪水灾害风险分析进展与展望[J]. 灾害学, 2017(3): 125-130+136.
- [5] 黄维. 城市排水管网水力模拟及内涝风险评估[D]. [硕士学位论文]. 广州: 华南理工大学, 2016.
- [6] 戴晶晶, 刘增贤, 陆沈钧. 基于数值模拟的城市内涝风险评估研究——以苏州市城市中心区为例[J]. 中国水利, 2015(13): 20-23.
- [7] 黄铁兰, 陈小芬, 苏章灿. 基于遥感和GIS的城市内涝风险评价——以广州市天河区为例[J]. 人民珠江, 2017(2): 89-95.
- [8] 王磊, 周玉文. 基于投影寻踪的城市排水系统洪涝风险评估模型[J]. 中国给水排水, 2011, 27(23): 78-82.
- [9] Friedman, J.H. and Tukey, J.W. (1974) A Projection Pursuit Algorithm for Exploratory Data Analysis. *IEEE Transactions on Computers*, **23**, 881-890. <https://doi.org/10.1109/T-C.1974.224051>
- [10] Glover, D.M. and Hopke, P.K. (1994) Exploration of Multivariate Atmospheric Particulate Compositional Data by Projection Pursuit. *Atmospheric Environment*, **28**, 1411-1424. [https://doi.org/10.1016/1352-2310\(94\)90204-6](https://doi.org/10.1016/1352-2310(94)90204-6)

-
- [11] 倪长健, 王顺久, 崔鹏. 投影寻踪动态聚类模型及其在地下水分类中的应用[J]. 四川大学学报(工程科学版), 2006, 38(6): 29-33.
- [12] 王卓, 倪长健. 投影寻踪动态聚类模型研究及其在洪灾评定中的应用[J]. 四川师范大学学报(自然科学版), 2008, 31(5): 635-638.
- [13] 任若恩, 王惠文. 多元统计数据分析: 理论、方法、实例[M]. 北京: 国防工业出版社, 1997.
- [14] 丛翔宇, 倪广恒, 惠士博, 等. 基于 SWMM 的北京市典型城区暴雨洪水模拟分析[J]. 水利水电技术, 2006(4): 64-67.
- [15] 符锐, 罗龙洪, 刘俊, 等. SWMM 模型中的低影响开发模块在排水防涝系统中的应用[J]. 水电能源科学, 2014(9): 71-74.
- [16] 王慧亮, 吴泽宁, 胡彩虹, 等. 基于 GIS 与 SWMM 耦合的城市暴雨洪水淹没分析[J]. 人民黄河, 2017(8): 31-35+43.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org