

基于STARMA模型的城市暴雨积水点 积水短时预测

郑姗姗^{1,2}, 万 庆¹, 贾明元^{1,2}

(1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 北京 100101;

2. 中国科学院大学, 北京 100049)

摘 要:近年来城市暴雨出现突发和多发态势,导致城市内涝灾害频繁发生,威胁着城市居民的生命和财产安全。随着城市降雨积水监测网的建立,获得分钟尺度的降雨和积水时序监测数据成为可能,实现了城市内涝的实时监控。但目前对监测数据的利用仍显不足,缺乏对其深度分析挖掘,造成监测系统“只监不控”的局面。本文基于城市降雨积水监测网的监测数据,根据积水时间相关性、降雨空间相关性以及降雨积水序列相关性,构建降雨积水的时空自相关移动平均模型(STARMA),对城市暴雨积水点积水过程进行短时预测。STARMA模型已被广泛应用于交通预测、环境变量预测以及社会经济领域,特别是在时空过程机理不清楚、多因素时空变量影响的情况下效果较好。本文首次将该模型应用到降水积水过程拟合和积水短时预测上,同时在方法上改进了传统单变量的STARMA模型,建立降雨和积水双变量的STARMA模型模拟降雨积水过程。并以北京市2012年“7.21”事件降雨积水过程为研究对象,以丰北桥、花乡桥、马家楼桥和六里桥4个积水监测点为例,建立降雨积水的STARMA模型,以5 min为步长作积水5、10、15 min三步预测。验证结果表明,该模型在降雨积水过程中拟合效果较好,模型短时预测精度较高。该项研究能够有效地利用监测数据,提高信息预警和应急指挥能力,为市政防汛或交通等部门提供决策支持。

关 键 词:暴雨积水;短时预测;时空序列;STARMA模型;时间自相关;空间相关

doi: 10.11820/dlkxjz.2014.07.010

中图分类号:P208

文献标识码:A

1 引言

近年来,城市暴雨出现突发和多发态势,导致城市内涝问题突出,威胁城市的正常运行,造成生命和财产的巨大损失。随着城市降雨积水监测网的建立,动态实时了解各地的降雨和积水情况成为可能。但是目前监测网数据仅用于实时信息的发布和展示,缺乏对这些时序数据的深度分析挖掘,造成了监测系统“只监不控”的局面(赵冬泉等, 2012)。本文基于降雨积水的时序监测数据开展积水的短时预测研究,根据积水时间自相关性、降雨空间相关性以及降雨积水的互相关性,建立降雨积水的时空自相关移动平均模型(Spatial-Temporal Auto Regressive and Moving Average, STARMA)(王

佳璆等, 2012),对城市暴雨积水点积水过程进行短时预测。该研究可有效地利用监测数据,提高信息预警和应急指挥能力,为市政防汛或交通等部门提供决策支持。

目前,对城市暴雨积水过程模拟和预测主要集中在城市雨洪模型的应用研究,其中,代表性的城市雨洪模型有SWMM模型、MOUSE模型、STORM模型和Wallingford模型等。城市雨洪模型分为经验模型、概念模型和理论模型(吴险峰等, 2002; 徐宗学, 2009),其中概念模型和理论模型均依据一定的物理基础而建立的。城市雨洪模拟技术已经形成了较为完善的模型框架,但对水文物理过程的认识、数据管理能力和数据资料的缺乏是限制模型模拟精度和应用范围的主要原因(胡伟贤等, 2010)。

收稿日期:2014-02;修订日期:2014-04。

基金项目:国家高科技研究发展计划(863计划)项目(2013AA122302)。

作者简介:郑姗姗(1990-),女,安徽合肥人,硕士研究生,主要研究方向为时空数据分析,E-mail: zhengss@lreis.ac.cn。

通讯作者:万庆(1964-),男,江西南昌人,研究员,主要从事GIS应用与防洪减灾研究,E-mail: wanq@lreis.ac.cn。

949-957 页

随着数据挖掘和大数据技术的发展,对大量数据进行分析处理挖掘已成为常用的方法之一。时间序列分析模型是经验模型,它主要是从数理统计的角度揭示和认识水文过程的复杂特性,是揭示和认识水文过程变化特性的有效手段和重要途径。其中,自回归移动平均模型(Auto Regressive and Moving Average, ARMA)为时序分析模型中序列相关性分析方法的一种。随着对水文现象过程复杂性认识不断提高,探讨多方法耦合或物理成因分析和数理统计分析相结合的方法,是提高水文时间序列分析精度和可靠性的有效手段(桑燕芳等, 2013)。城市降雨积水监测网提供了大量的分钟粒度的时序监测数据,为建立城市雨洪模型提供数据支持。本文基于降雨积水的时空序列数据,将 ARMA 模型和空间分析方法相结合,建立 STARMA 模型来提高积水短时预测的精度和效率。

时空序列建模方法(STARMA)假设变量是当前位置 i 及其空间邻接位置过去时期观测值与随机误差项的线性组合,根据历史的样本观测数据建立时空相关—自相关函数得到时空相关性模型(Pfeifer et al, 1980; Kamarianakis et al, 2005)。国内外在 STARMA 模型应用方面已经取得了一些进展, Kamarianakis 等(2005)、韩卫国等(2007)、赵玲等(2012)将该模型应用在交通预测领域;De Luna 等(2005)、王佳璆等(2007)、王佳璆(2008)、Cheng 等(2011)、将该模型应用在环境变量预测上;Martin 等(1975)、Pfeifer 等(1980)、王艳辉等(2005)将 STARMA 模型应用在社会和经济领域等。越来越多的研究表明, STARMA 模型是建模地理时空现象的有效工具,特别是在时空过程机理不清楚、多因素时空变量影响的情况下。本文首次将该方法应用到降水积水序列过程拟合和积水短时预测上,同时在方法上改进了传统单变量的 STARMA 模型,建立双变量的 STARMA 模型来适应降雨积水二元序列过程研究。

2 时空序列模型

2.1 时间自相关移动平均模型 ARMA

时间序列是某一变量按时间先后顺序排列而成的数列,也是所研究系统的历史行为的客观记录,它包含了系统结构特征以及运行规律。时间序列分析可以从数量上揭示某一现象的发展变化规

律,或是从动态的角度刻画某一现象与其他现象之间的内在关系及其变化规律,其最主要的思想是根据预测变量本身或其他相关变量过去的变化规律来预测未来(王振龙等, 2007)。根据研究对象的个数,时间序列可分为一元时间序列和多元时间序列。降雨和积水过程具有相互依存的动态规律性,因此需要建立二元的时间序列模型来描述降雨积水变化过程。

一个系统,如果它在时刻 t 的响应 $z(t)$,不仅与其以前时刻的自身值有关,而且还与其以前时刻进入系统的扰动存在一定的依存关系,那么这个系统就是自回归移动平均模型 ARMA。AR 模型(Auto Regressive)描述的是系统对过去自身状态的记忆, MA 模型(Moving Average)描述的是系统对过去时刻进入系统的噪声的记忆,而 ARMA 模型则是系统对过去自身状态以及各时刻进入系统的噪声的记忆。ARMA(p, q)即 p 阶自回归 q 阶移动平均模型为:

$$z(t) = \varphi_1 z(t-1) + \cdots + \varphi_p z(t-p) + \varepsilon(t) - \theta_1 \varepsilon(t-1) - \cdots - \theta_q \varepsilon(t-q) \quad (1)$$

式中: $z(t-p)$ 是 $t-p$ 时刻对应的系统值, $\varepsilon(t-q)$ 是 $t-q$ 时刻对应的系统扰动, φ_p 和 θ_q 分别对应阶数项的参数;时间自相关性表示变量在时间上的依赖,用时间自相关系数来度量,其取值范围在 $(-1, 1)$, 值越接近于 1, 说明时间序列的自相关程度越高。对于平稳时间序列 $z(t)$, 定义时间延迟期为 k 的自相关函数(Autocorrelation Function, ACF)为:

$$\rho_k = \frac{\text{Cov}(z(t), z(t+k))}{\sigma_{z(t)} \sigma_{z(t+k)}} \quad (2)$$

偏相关函数(Partial Autocorrelation Function, PACF)定义为:

$$\varphi_k = \frac{E(z(t)z(t-k)|z(t-1), \dots, z(t-k+1))}{\sigma(z(t)|z(t-1), \dots, z(t-k+1))} \quad (3)$$

式中: $z(t)$ 为 t 时刻对应的时间变量; k 为时间延迟期; Cov 为协方差; σ 为方差; E 为条件期望。

2.2 时空自相关移动平均模型 STARMA

地理学中,由于连续的时空数据都是经过离散化抽样提取并存储的,因而能够将时空数据看作是空间上有相关关系的时间序列集合,即时空序列。时空序列数据与一般时间序列数据和空间数据相比,时空相关性和时空异质性是其最主要的特征。其中空间相关性描述的是空间区域中某一位置与其邻近位置上变量的相关性(曹志冬等, 2008)。本

文通过建立空间权重矩阵计算其降雨量,权重值为雨量站之间空间相关系数(王劲峰等, 2010)。一般的时空自相关移动平均模型是针对同一个变量的时间相关性和空间相关性建立的, $z_i(t)$ 是由该点过去时刻值和过去扰动值,再加上空间邻近点的过去时刻值和过去扰动值的影响。以空间权重矩阵的形式表示变量空间位置的相关性(王佳璆等, 2012)。STARMA模型如下:

$$z_i(t) = \sum_{k=1}^p \sum_{h=0}^{m_k} \varphi_{kh} W^{(h)} z_i(t-k) - \sum_{l=1}^q \sum_{h=0}^{n_l} \theta_{lh} W^{(h)} \varepsilon_i(t-l) + \varepsilon_i(t) \quad (4)$$

式中: p 是时间自相关阶数; q 是时间移动平均阶数; $W^{(h)}$ 是空间权重矩阵; m_k 是第 k 个时间自相关项的空间阶数; n_l 是第 l 个时间移动平均项的空间阶数; φ_{kh} 和 θ_{lh} 为对应的模型参数; $\varepsilon_i(t)$ 是随机误差。

2.3 改进的降雨积水STARMA模型

根据城市雨洪模型可知,城市水文过程主要是受降雨、汇流、排水过程的影响。各个积水点的汇流面积、地面结构和排水情况各不相同,所以应该针对不同的积水点建立不同的模型,并分别率定参数(丛翔宇等, 2006; 朱晨东, 2011; 潘安君等, 2012)。由于积水的成因机制,积水现状除了受积水自身的时间序列的影响外,还受到降雨的时间序列的影响,所以需建立二元时间序列模型,将降雨作为传递函数考虑进来。积水具有时间上连续、空间不连续的分布特征,所以只考虑积水序列的时间相关性,不考虑积水点之间的空间相关性。降雨具有空间上连续、时间上不连续的分布特征,所以只考虑降雨之间的空间相关性,不考虑降雨之间的时间相关性(王佳璆等, 2012)。通过空间权重矩阵表示雨量站的降雨序列对积水点处降雨的影响权重。由于降雨和积水之间的时间互相关性复杂,当前研究暂不考虑降雨和积水序列之间的时间互相关性。根据以上讨论,为了更好模拟降雨积水过程,需对STARMA模型作出如下修正:

$$z_i(t) = \sum_{k=1}^p \varphi_k z_i(t-k) - \sum_{l=1}^q \theta_l \varepsilon_i(t-l) + \rho_i \sum_{j=1}^n \lambda_{ji} y_j(t-1) + \varepsilon_i(t) \quad (5)$$

式中: $y_j(t)$ 是雨量 j 站点 t 时刻的降雨量, n 为受影响的雨量站点个数。

3 STARMA建模步骤

3.1 模型阶数识别

降雨积水时空序列模型STARMA建模过程如图1所示。为了确定STARMA模型的形式,需要计算时空自相关函数(ACF)和时空偏相关函数(PACF)来识别自相关参数 p 和移动平均参数 q 。通过计算出时间延迟 k 、空间延迟 h 的相邻区域时空自相关函数表和偏相关函数表,观察函数表的空间延迟值和时间延迟值。如果在空间延迟的相邻区域,时间延迟为拖尾,而偏相关函数值在空间延迟为 h 的相邻区域,时间延迟为 p 时截尾,则可以确定模型是STAR(p)即阶数为 p 的自相关过程;如果时间自相关函数值在空间延迟为 h 的相邻区域,时间延迟为 q 时截断,而偏相关函数在所有空间延迟的相邻区域,时间延迟值拖尾,则可以确定模型是STMA(q)即阶数为 q 的移动平均过程;如果时空自相关函数和偏相关函数都拖尾,则表明模型是时空相关移动平均过程STARMA(p, q)(王振龙等, 2007)。STARMA模型的阶数还需要通过一些经验来确定,可以

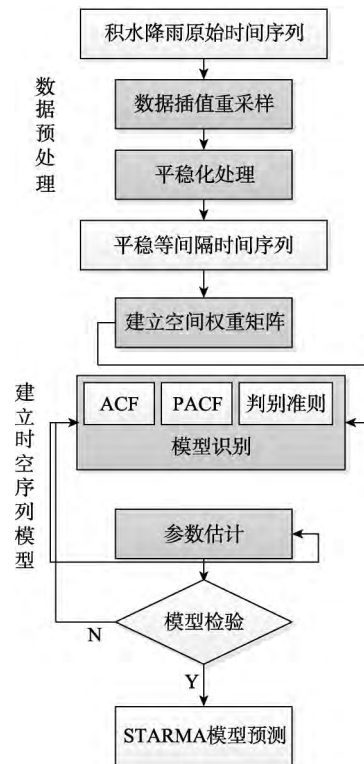


图1 降雨积水时空序列建模流程图

Fig.1 Flowchart of spatial-temporal modeling of rainfall and waterlogging

通过代入不同的阶数进行计算,选择适当的比较准则来确定,如AIC最小信息准则(Akaike Information Criterion, AIC)和BIC贝叶斯信息准则(Bayesian Information Criterion, BIC)(王振龙等, 2007; 王燕, 2008)。BIC准则的定义如下:

$$BIC(n, m) = \ln[\hat{\sigma}_a^2] + r \ln \frac{N}{N} \quad (6)$$

式中: n 和 m 为ARMA对应的模型阶数; $\hat{\sigma}_a^2$ 为模型残差方差的极大似然估计; $r=n+m$ 为模型独立参数的个数; N 为时间序列的长度。

3.2 空间权重矩阵建立

为了将降雨的影响考虑到积水预测中,首先需要确定各个积水点对应的降雨量。因为雨量站和积水站分别分布在城市的不同地方,需要对降雨数据进行空间插值,进而得到积水点对应的降雨量。目前用于降雨资料的空间插值方法有多种,如泰森多边形法、反距离插值法、趋势面法及克里格法(Kriging)等(封志明等, 2004; 朱会义等, 2004; 石朋等, 2005; 王劲峰等, 2010; 周成虎等, 2011)。由于时序数据时间频率较高,需要不断的对数据进行插值计算导致计算量太大,因此,本文选取反距离插值法来估算积水点的降雨量(石朋等, 2005; 王宗敏等, 2012)。根据雨量站和积水站距离的函数来建立积水对应雨量站的空间权重矩阵。 t 时刻积水点 i 处的降雨量可以根据以下公式计算得出:

$$y_i(t) = \sum_{j=0}^n [y_j(t) \times (1/d_{ij}^2)] / \sum_{j=0}^n (1/d_{ij}^2) \quad (7)$$

其各个雨量站对应的权重为:

$$\lambda_{ij} = (1/d_{ij}^2) / \sum_{j=0}^n (1/d_{ij}^2) \quad (8)$$

式中: $y_i(t)$ 是 t 时刻积水点 i 处的降雨量, d_{ij} 是积水点 j 和雨量站 i 之间的距离系数, n 是雨量站的个数, λ_{ij} 是 j 积水点对应的雨量站的空间权重系数, $y_j(t)$ 是 t 时刻第 j 个雨量站对应的降雨量。

3.3 参数估计模型验证

参数估计是估算出STARMA模型的参数值,常用的方法有矩估计、最小二乘法估计和极大似然估计(王振龙等, 2007),本文采用最小二乘法对模型进行参数估计。模型验证需要对模型拟合性能进行评估和对模型残差进行分析,通过计算拟合优度 R^2 、 F 检验和计算误差方差来衡量。拟合性能评估即是对模拟值和实测值之间进行比较,当

两者拟合效果较好或在预定误差范围之内时,模型才可以应用。残差分析是检验残差 μ_i 的特性是否与模型假设中要求的相一致,尤其是对残差序列是否相互独立、是否同方差、是否服从某一假定分布等进行检验。如果残差是随机误差,则说明模型能够完全解释时空数据,否则残差会呈现出一定的模式,即在时空上存在着相关性或变异性(王佳璆等, 2012)。

4 降雨积水STARMA建模与预测

4.1 研究数据及预处理

北京市降雨积水监测网共有84个积水监测点和12个雨量站点,均分布在北京五环内城区内(图2)。其中,积水监测点主要分布在道路上,以高架式立交桥区或下凹式立交桥区为主。本文以北京市2012年“7.21”事件降雨积水过程为研究对象,以丰北桥、花乡桥、马家楼桥和六里桥4个积水监测点(这几个点的积水监测装置均安装在高架桥下)的监测数据为基础进行建模与预测。降雨积水监测数据的基本情况是:每隔5 min记录一次该时刻的积水高度,精度为0.01 m,并记录下每个2 min内的降雨总量,精度为0.5 mm。

利用线性插值方法对数据预处理来插补缺失的数据。时间序列分析要求是等间隔序列,所以需要将积水和降雨序列统一成同一个时间尺度,时间粒度为5 min。由于积水序列是属于非平稳时间序列,所以需要对原数列进行平稳化处理。一般来说,一阶差分可消除线性趋势,二阶差分可消除二次曲线趋势(王燕, 2008)。

4.2 模型定阶

在建立STARMA(p, q)模型时,通常利用时间自相关和偏相关函数来确定模型 p, q 阶数。以丰北桥为例,图3显示了时序的自相关函数图和偏相关函数图。自相关系数和偏相关系数均不结尾,可以确定模型是ARMA模型,但是无法确定 p, q 的阶数,还需要通过其他统计量来判断如拟合优度 R^2 、均方根误差RMSE、平均绝对误差MAE和BIC等。BIC准则函数公式共由两部分构成,第一部分反映模型拟合的好坏,第二部分表示模型参数的多少。准则函数的目的是为了拟合残差与参数个数两个目标之间进行不同的权衡。当计算出BIC最小时,

4.4 模型精度评价

根据建立的时空序列模型,利用积水监测点的历史过程数据对积水作三步预测(步长为5 min),预测结果如图4所示。对不同时间预测结果的精度评价指标为均方根误差 *RMSE*、相对平方误差 *RSE* 和相对误差比率 *RE*,评价结果如表4所示。根据预测精度评价表可知模型预测精度较高,平均 *RMSE* 在

0.03 左右,平均 *RSE* 在 0.01 左右,平均误差比率在 5% 左右。随着预测时间从 5 min、10 min、15 min 依次增加,模型预测精度也有所下降。积水的时间跨度范围越长,模型预测精度越精确,如马家楼桥积水时间跨度为 9 个小时,预测精度在 3% 左右;而丰北桥积水时间跨度为 3 个小时左右,预测精度在 6% 左右。

表 2 模型检验统计量

Tab.2 Statistics of model test

统计量	R^2 statistic	F statistic	MAE
丰北桥	0.9718	519.4053	0.0285
花乡桥	0.9901	2.34	0.0113
马家楼桥	0.9863	2.5043	0.0167
六里桥	0.9887	3.28	0.0245

表 3 模型方程的参数估计值

Tab.3 Parameter estimates of models

参数	φ_{i1}	θ_{i1}	ε	ρ_i
丰北桥	0.9559	0.6863	1.5809	0.2863
花乡桥	0.9882	0.6375	0.0037	0.0024
马家楼桥	1.0066	0.5013	-0.0111	0.0029
六里桥	1.0095	0.7798	-0.0257	0.0056

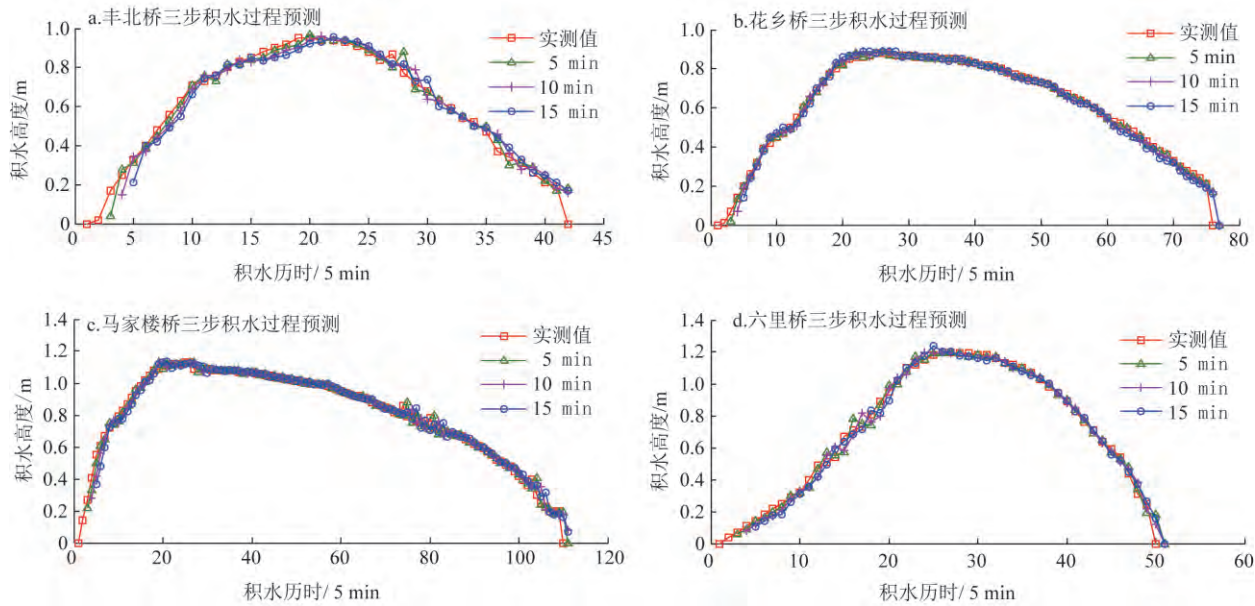


图4 4个积水点的5、10、15min积水预测值

Fig.4 Predicted values of water depth at four waterlogging sites, 5, 10, and 15 minutes

表 4 模型预测精度评价表

Tab.4 Accuracy assessment of model prediction

预测 时间 指标	5 min			10 min			15 min		
	$RMSE$	RSE	$MRE/\%$	$RMSE$	RSE	$MRE/\%$	$RMSE$	RSE	$MRE/\%$
丰北桥	0.0369	0.0230	6.0	0.0378	0.0263	5.9	0.0437	0.0374	6.2
花乡桥	0.0226	0.0093	2.1	0.0253	0.0124	2.65	0.0269	0.0148	3.3
马家楼	0.0314	0.0136	2.6	0.0326	0.0151	2.7	0.0364	0.0013	3.4
六里桥	0.0396	0.0111	4.4	0.0410	0.0126	6.0	0.0382	0.0114	5.2

5 结论与讨论

本文基于城市降雨积水监测网的时序数据,将ARMA模型在空间上加以扩展,依据积水时间相关性、降雨空间相关性和降雨积水之间的相关性,构建降雨积水的STARMA模型。本文首次将该方法应用到降水积水序列过程拟合和积水短时预测上,同时,在方法上改进了传统单变量的STARMA模型,建立降雨和积水双变量的STARMA模型来模拟降雨积水过程。以北京市2012年“7.21”事件降雨积水过程为研究对象,以丰北桥、马家楼桥、花乡桥和六里桥4个积水监测点为例,采用改进后的STARMA模型对积水进行5、10、15 min三步预测。实验结果表明STARMA模型在降雨积水过程中拟合效果较好,平均拟合的RMSE在0.03左右;在积水预测方面精度较高,未来5、10、15 min后的积水情况预测精度平均RMSE能达到0.03左右、平均相对误差比率在5%左右。本研究可有效地利用监测数据,提高了信息预警和应急指挥能力,为市政防汛或交通等相关部门提供决策支持。

由于城市积水过程的复杂性,积水过程除了受降雨和积水的影响,还受地面汇流和排水情况等的影响。当前研究仅考虑了降雨和积水的时空相关性。为了提高模型的精度和可靠性,后续研究可以从排水因素和降雨积水序列的时间迟滞性两方面来继续完善模型,并采用将最新监测数据作为验证数据的模型参数动态率定模式,以保证模型的参数能够尽可能的反映实时情况,提高模型预测分析的可靠性和现实性。

参考文献(References)

- 曹志冬,王劲峰,高一鸽,等. 2008. 广州SARS流行的空间风险因子与空间相关性特征. 地理学报, 63(9): 981-993. [Cao Z D, Wang J F, Gao Y G, et al. 2008. Risk factors and autocorrelation characteristics on severe acute respiratory syndrome in Guangzhou. Acta Geographica Sinica, 63(9): 981-993.]
- 丛翔宇,倪广恒,惠士博,等. 2006. 城市立交桥暴雨积水数值模拟. 城市道桥与防洪, (2): 52-55, 152. [Cong X Y, Ni G H, Hui S B, et al. 2006. Numerical simulation of urban stormwater overpass. Urban Roads Bridges & Flood Control, (2): 52-55, 152.]
- 封志明,杨艳昭,丁晓强,等. 2004. 气象要素空间插值方法优化. 地理研究, 23(3): 357-364. [Feng Z M, Yang Y Z,

- Ding X Q, et al. 2004. Optimization of spatial interpolation methods for climate resources. Geographical Research, 23(3): 357-364.]
- 韩卫国,王劲峰,高一鸽,等. 2007. 区域交通流的时空预测与分析. 公路交通科技, 24(6): 92-96. [Han W G, Wang J F, Gao Y G, et al. 2007. Forecasting and analysis of regional traffic flow in space and time. Journal of Highway and Transportation Research and Development, 24(6): 92-96.]
- 胡伟贤,何文华,黄国如,等. 2010. 城市雨洪模拟技术研究进展. 水科学进展, 21(1): 137-144. [Hu W X, He W H, Huang G R, et al. 2010. Review of urban storm water simulation techniques. Advances in Water Science, 21(1): 137-144.]
- 潘安君,侯爱中,田富强,等. 2012. 基于分布式洪水模型的北京城区道路积水数值模拟: 以万泉河桥为例. 水力发电学报, 31(5): 19-22. [Pan A J, Hou A Z, Tian F Q, et al. 2012. Numerical simulation of Beijing urban roads flood water distributed model: a case study of Wanquan River Bridge. Journal of Hydroelectric Engineering, 31(5): 19-22.]
- 桑燕芳,王中根,刘昌明. 2013. 水文时间序列分析方法研究进展. 地理科学进展, 32(1): 20-30. [Sang Y F, Wang Z G, Liu C M. 2013. Research progress on the time series analysis methods in hydrology. Progress in Geography, 32(1): 20-30.]
- 石朋,芮孝芳. 2005. 降雨空间插值方法的比较与改进. 河海大学学报: 自然科学版, 33(4): 361-365. [Shi P, Rui X F. 2005. Comparison and improvement of spatial rainfall interpolation methods. Journal of Hohai University: Natural Sciences, 33(4): 361-365.]
- 王佳璆. 2008. 时空序列数据分析和建模[D]. 广州: 中山大学. [Wang J Q. 2008. Analysis and modeling of spatial-temporal series data[D]. Guangzhou, China: Sun Yat-sen University.]
- 王佳璆,程涛. 2007. 时空预测技术在森林防火中的应用研究. 中山大学学报: 自然科学版, 46(2): 110-113, 116. [Wang J Q, Cheng T. 2007. Applications of spatio-temporal forecasting to forest fire prevention. Acta Scientiarum Naturalium Universitatis Sunyatseni: Natural Sciences, 46(2): 110-113, 116.]
- 王佳璆,邓敏,程涛,等. 2012. 时空序列数据分析和建模. 北京市: 科学出版社. [Wang J Q, Deng M, Cheng T, et al. 2012. Analysis and modeling of spatial-temporal series data. Beijing, China: Science Press.]
- 王劲峰,廖一兰,刘鑫. 2010. 空间数据分析教程. 北京: 科学出版社. [Wang J F, Liao Y L, Liu X. 2010. Spatial data

- analysis tutorial. Beijing, China: Science Press.]
- 王艳辉, 贾利民, 王卓, 等. 2005. 基于时空序列的铁路客运量数据挖掘预测. 中国铁道科学, 26(4): 130-135. [Wang Y H, Jia L M, Wang Z, et al. 2005. Study on prediction of railway passenger traffic volume based on time-space serial. China Railway Science, 26(4): 130-135.]
- 王燕. 2008. 应用时间序列分析. 北京: 中国人民大学出版社. [Wang Y. 2008. Application time series analysis. Beijing, China: China Renmin University Press.]
- 王振龙, 胡永宏. 2007. 应用时间序列分析. 北京: 科学出版社. [Wang Z L, Hu Y H. 2007. Application time series analysis. Beijing, China: Science Press.]
- 王宗敏, 张杰, 赵红领, 等. 2012. 城市暴雨内涝计算中降雨插值算法的选取. 人民黄河, 34(8): 24-26. [Wang Z M, Zhang J, Zhao H L, et al. 2012. Selection of the interpolations of rainfall for calculating the water logging disasters caused by the urban storms. Yellow River, 34(8): 24-26.]
- 吴险峰, 刘昌明. 2002. 流域水文模型研究的若干进展. 地理科学进展, 21(4): 341-348. [Wu X F, Liu C M. 2002. Progress in watershed hydrological models. Progress in Geography, 21(4): 341-348.]
- 徐宗学. 2009. 水文模型. 北京: 科学出版社. [Xu Z X. 2009. Hydrological Model. Beijing, China: Science Press.]
- 赵冬泉, 王浩正, 陈吉宁, 等. 2012. 监测技术在排水管网运行管理中的应用及分析. 中国给水排水, 28(8): 11-14. [Zhao D Q, Wang H Z, Cheng J N, et al. 2012. Application and analysis of monitoring technology in the operation and management of the drainage network. China Water & Wastewater, 28(8): 11-14.]
- 赵玲, 王佳璐, 邓敏, 等. 2012. 基于行车时间的道路网络时空自相关模型研究. 中南大学学报: 自然科学版, 43(10): 4114-4122. [Zhao L, Wang J Q, Deng M, et al. 2012. Spatial-temporal autocorrelation research about road network based on travel time. Journal of Central South University: Science and Technology, 43(10): 4114-4122.]
- 周成虎. 2011. 地理信息系统空间分析原理. 北京: 科学出版社. [Zhou C H. 2011. GIS spatial analysis theory. Beijing, China: Science Press.]
- 朱晨东. 2011. 关于北京下凹式立交桥积水的思考. 北京水务, 160(5): 4-5. [Zhu C D. 2011. Reflect on the Beijing recessed overpass waterlogging. Beijing Water, 160(5): 4-5.]
- 朱会义, 贾绍凤. 2004. 降雨信息空间插值的不确定性分析. 地理科学进展, 23(2): 34-42. [Zhu H Y, Jia S F. 2004. Uncertainty in the spatial interpolation of rainfall data. Progress in Geography, 23(2): 34-42.]
- Cheng T, Wang J Q, Li X. 2011. A hybrid framework for space-time modeling of environmental data. Geographical Analysis, 43(2): 188-210.
- De Luna X, Genton M G. 2005. Predictive spatio-temporal models for spatially sparse environmental data. Statistica Sinica, 15(2): 547-568.
- Kamarianakis Y, Prastacos P. 2005. Space-time modeling of traffic flow. Computers & Geosciences, 31(2): 119-133.
- Martin R L, Oeppen J E. 1975. The identification of regional forecasting models using space-time correlation functions. Transactions of the Institute of British Geographers, 66: 95-118.
- Pfeifer P E, Jay Deutsch S. 1980. Stationarity and invertibility regions for low order starma models: stationarity and invertibility regions. Communications in Statistics-Simulation and Computation, 9(5): 551-562.

Short-term forecasting of waterlogging at urban storm-waterlogging monitoring sites based on STARMA model

ZHENG Shanshan^{1,2}, WAN Qing¹, JIA Mingyuan^{1,2}

(1. State Key Laboratory of Resource and Environmental Information System, Institute of Geographic Science and Natural Resources Research, CAS, Beijing 100101, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Storm struck cities frequently and often suddenly in recent years, leading to urban inundation and threatening life and property in these cities. With the establishment of urban storm-waterlogging monitoring network, real-time and time series rainfall and waterlogging data at the temporal resolution of a few minutes can be easily acquired. Real-time monitoring of inundation thus can be achieved and this provides new ways for research of inundation in cities. At present, however, there is a general lack of in-depth data mining and analysis of the observed data, which leads to the fact that urban inundation monitoring systems are used only for monitoring purposes and waterlogging control is not an integrated part of the system. Based on the monitoring data of urban inundation monitoring system, according to the temporal autocorrelation of waterlogging, spatial correlation of rainfall and the correlation of rainfall and waterlogging, a spatial and temporal auto regressive and moving average model (STARMA) has been built for short-term forecasting of waterlogging in this study. Auto regressive and moving average model (ARMA) is one of the correlation analysis methods of time series data. ARMA combined with spatial analysis methods generates the STARMA model. STARMA model is an effective means for modeling the spatiotemporal processes of geographic phenomena, especially when the mechanism of the spatiotemporal processes is unclear or multiple spatial and temporal variables are involved. STARMA model has been applied in traffic prediction, environment variable prediction and the social and economic fields. In this study, the model is applied in rainfall and waterlogging process simulation and short-term forecasting for the first time. In order to simulate rainfall and waterlogging processes, the traditional STARMA univariate model is modified to create a bivariate model of rainfall and waterlogging. Based on urban inundation monitoring data on 12 July 2012 in Beijing, using Fengbei Bridge, Huaxiang Bridge, Majialou Bridge and Liuli bridge as examples, the STARMA models were built respectively to predict water depth with a 5 minute step at 5, 10, and 15 minutes. The modeling process included creating rainfall stations' spatial weight matrix, model identification, parameter estimation and model verification. The STARMA model form was determined by autocorrelation function and partial autocorrelation, in addition to A-information criterion or Bayesian information criterion. Spatial weight values were calculated by inverse distance weighting (IDW). Parameter estimates were derived by the least square method. The simulation results show that the STARMA model predictions fit well with observed data and accuracy of short-term forecasting is high. The root mean square error (RMSE) is about 0.03, the relative square error (RSE) is about 0.01 and the average error rate is about 5%. When the prediction time increased from 5 to 15 minutes, prediction accuracy slightly decreased. This method improves prediction accuracy and reliability as compared to traditional hydro model simulation and prediction. The research uses urban inundation monitoring network data to predict short-term waterlogging. On the one hand it takes full advantage of the monitoring data, and on the other hand it improves the ability of disaster early warning and emergency command, thus provide decision support for related government departments.

Key words: storm waterlogging; short-term forecasting; space-time sequence; STARMA model; temporal autocorrelation; spatial correlation