基于动态聚类分析和模糊模式识别法的北京城区汛期降雨时空分布规律研究

刘媛媛[1]，王 毅[2]， 刘洪伟[2]，杜龙刚[3]，刘 舒[1]，柴福鑫[1]

（1.中国水利水电科学研究院，北京 100038；2.北京市人民政府防汛抗旱指挥部办公室，北京 100038；
3.北京市水文总站，北京 100089）

摘 要：以北京城区近十几年汛期50场暴雨过程为样本，利用动态聚类分析和模糊模式识别方法对单站雨型和降雨整体空间分布特征进行了研究，得到北京城区汛期最容易发生的雨型以及时空分布规律。结果表明，北京城区汛期主要发生前单峰型雨，峰值集中且极值雨量大，容易造成内涝积水。发生双峰型雨的次数较少，且极值雨量较小，对内涝积水影响相对较小。

关键词：动态聚类；模糊模式识别；时空分布规律

中图分类号：P333 **文献标识码**：A **文章编号**：1000-0852(2019)01-0074-04

1 引言

随着北京城市的发展，暴雨引发的城市内涝灾害影响越来越大，成为影响城市建设，正常生活的主要气象灾害。越来越多的专家开始关注北京城区极端降雨的特征研究。不同类型的降雨过程，形成的洪峰流量有较大差异。在汇流历时内平均雨强相同的条件下，雨峰在中部或后部的三角形雨型比均匀雨型的洪峰大30%以上[1]，而不同的洪峰特征差异将直接影响城区内涝积水的分布结果。

包高马佐娃等人在20世纪40年代，就将降雨雨型分为7类。包括峰值在前、中、后的单峰型、均匀性，以及峰值分别处于不同位置的双峰型雨。Pilgrim和Cordery[2]提出的雨型与实际降雨过程较为相似。他们将雨峰时段放在出现可能性最大的位置上，雨峰时段在总雨量中的比例取各场降雨雨峰所占比例的平均值。Knifer和Chu[3]根据强度-历时-频率关系得到了一种不均匀的设计雨型—芝加哥雨型，该雨型中任何历时内的雨量等于设计雨量。该方法在国内外应用较为广泛。HUFF[4]通过研究美国伊里诺斯州的暴雨，把雨峰出现位置及降雨历时分为4种类型，根据最大雨强在降雨过程中发生的具体时间段，对每一类雨型设计多种不同频率的时间分配过程。这些方法，主要都是针对某一地区或者某一代表站进行研究的，而对于整体时空分布特征研究相对较少。

目前对城区暴雨的设计计算，主要以该地区某一点的暴雨资料为研究对象，利用该地区的暴雨强度公式或者芝加哥雨型进行计算。这种做法在实际工作中使用方便、快捷。但一场实际降雨过程中，代表站的雨型并不能代表其它雨量站的雨型，采用某一点的资料来代替大的区域对于计算的精度影响很大[5]。各个站的雨型特征虽然有近似之处，但是又各有不同。单一一种设计雨型并不能反映实际降雨特征。

因此，对于雨型特征的分析，必须以城区整体作为研究对象，以各站大量的降雨过程为样本，对每场降雨进行时空分布的特征研究。概括多数降雨过程中，各雨量站雨型及城区整体特征，这样得到的雨型时空规律才会相对客观地反映城区降雨过程的主要特征。

本文基于北京城区14个雨量站，2004~2016年5min间隔的降雨资料，对城区各个雨量站，在降雨过程中的雨型特征进行统计分析计算，得到了北京城区汛期降雨主要几种雨型，这几种雨型可以基本客观地反映出北京城区汛期各雨量站的特征。得到了主要雨型后，将实际发生的降雨过程进行模糊模式识别，进一步分析降雨过程的整体分布规律，对北京城区汛期降雨时空分布规律进行较为完整的研究，为北京城区

收稿日期：2017-12-23
基金项目：国家重点研发计划“一体化综合减灾智能服务系统”项目(2016YFC0803109)
作者简介：刘媛媛(1978-)，女，天津人，高级工程师，主要从事防洪减灾方面的研究工作。E-mail:50237479@qq.com

防汛工作提供技术支持。

2　资料与方法

2.1　资料及处理

为了较为准确地反映本地区的暴雨雨型规律，本文采用2004~2016年，小时雨量大于30mm、城区各站均有明显降雨的50场典型暴雨过程。以这50场降雨过程为样本，进行统计分析。研究对象主要是北京城区内14个雨量站，如图1所示。

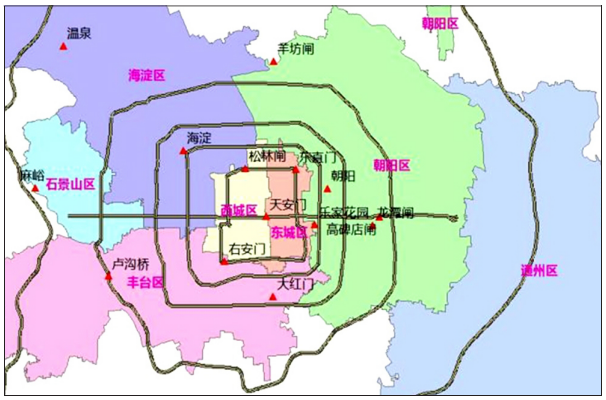

图1　城区14个雨量站分布图

Fig.1 The distribution of the 14 rainfall stations in urban area

2.2　雨型分析方法

2.2.1　雨型特征值向量

本文统计分析的方法，主要采用的是动态聚类分析法。首先根据各站的雨型时空分布的特征，利用聚类分析，将同一种类型的降雨过程归为一类，同一类的内部，特征近似，不同类之间，特征差距较大。提取城区各站每一种类型的聚类中心，并取近似雨型的包络，作为该雨型的典型降雨特征。

将一场24h历时的降雨过程，分为3h间隔的8个时段，每个时段的降雨量占总降雨量的比例为：

$$x_i=\frac{R_i}{R_z}\qquad (i=1,2,\cdots 8)\qquad (1)$$

式中：R_i 为时段雨量；R_z 为24h历时的总雨量，则该24h历时降雨过程的雨型指标，可以用向量 X 表示。

$$X_j=[x_{j1}\ x_{j2}\ \cdots\ x_{jm}]\quad (j=1,2,\cdots 14)\ (m=8)\qquad (2)$$

式中：x_{ji} 为一场实际降雨过程中第 j 雨量站的第 i 个时段的雨型指标。则城区14个站，50场降雨的雨型特征指标向量形成模式样本，进行分析统计。

2.2.2　动态聚类法

动态聚类法是先选择若干样本点作为聚类中心，再按某种聚类准则，使各样本点向各个中心聚集，从而得到初始分类，然后，判断初始分类是否合理，如果不合理，就修改分类，……。以此反复进行修改聚类的迭代运算，直到合理为止。本文就采用动态聚类法中的K-M均值算法，计算各站雨型的聚类中心。

该方法基于使聚类性能指标最小化。这里所用的聚类准则函数是聚类集中每一样本点到该类中心的距离平方之和，并使它最小化。计算步骤为：

(1)选 K 个初始聚类中心：$z_1(1),z_2(1),\cdots\cdots z_K(1)$，括号内的序号为寻找聚类中心的迭代运算次符号。聚类中心的向量值可任意设定。例如可用开头 K 个模式样本的向量值作为初始聚类中心。

(2)逐个将需分类的模式样本{x}按最小距离原则分配给 K 个聚类中心中的某一个 $z_j(1)$。

例如 $i=j$ 时，

$$D_j(k)=\min\{\parallel x-z_i(k)\parallel\}\quad i=1,2,\cdots k\qquad (3)$$

则 $x\in S_j(k)$。

式中：k 为迭代运算次序号；S_j 为第 j 个聚类；其聚类中心为 z_j。

(3)计算各个聚类中心的新的向量值：

$$z_j(k+1)=\frac{1}{N_j}\sum_{x\in S_j(k)}x\quad j=1,2,\cdots K\qquad (4)$$

式中：N_j 为第 j 个聚类域 S_j 中所包含的样本个数。以均值向量作为新的聚类中心，可使聚类准则函数：

$$J_j=[\sum_{x\in S_j(k)}\parallel x-z_j(k+1)\parallel^2]^{1/2}\qquad (5)$$

最小，其中 $j=1,2,\cdots K$。

(4)如果 $z_j(k+1)\neq z_j(k)$，$j=1,2,\cdots K$，则回到(3)，将模式样本逐个重新分类，重复迭代计算。如果 $z_j(k+1)=z_j(k)$，$j=1,2,\cdots K$，算法收敛，计算完毕。

该算法的结果要受到所选的聚类中心的数目和其初始位置，以及样本模式的几何性质及出入次序等影响。在实际应用中，本文试算了不同的 K 值并选择了不同的聚类中心起始值进行动态聚类计算，以求得该站合理的聚类中心。

得到了聚类中心，也就是北京城区各站经常出现的几种雨型。将各站的聚类中心指标进行对比，并取各站近似雨型向量的包络，进而得到北京城区各站的主要发生的几种雨型。这几种雨型基本反映了城区各站在多数降雨过程中的主要特征。

2.2.3　模糊模式识别

将通过动态聚类分析得到的 k 种雨型特征，写成向量形式为：

$$S_k=(s_{k1},s_{k2},\cdots,s_{km})\ (m=8)\qquad (6)$$

将实际降雨过程各站的降雨过程也写成向量形式，并计算实际雨型与这 k 种雨型的近似程度：

$$\sigma_k=1-\sqrt{\frac{1}{m}\sum_{i=1}^{m}(s_{ki}-x_{ki})^2} \quad (m=8) \qquad (7)$$

利用式(7)计算实际雨型与这 k 种雨型的贴近度，若实际发生的雨型与第 k 种雨型贴近度最小，则该雨型即为第 k 类雨型。

3 结果分析

利用动态聚类分析方法，先对城区各雨量站的单站雨型进行统计分析，再针对每场降雨中，城区各站的雨型分布规律进行研究。经研究发现，北京城区各雨量站的单站雨型和整体分布情况，都有明显的规律性。

3.1 单站雨型分析

经统计分析，城区 14 个站的雨型可聚类为 6 种主要类型。其中单峰型有 3 种，双峰型有 3 种。具体雨型如图 2 所示。

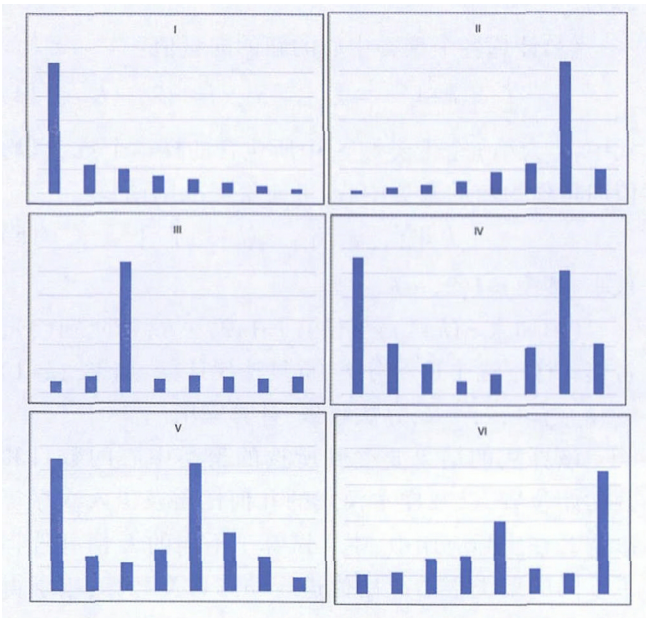

图 2 北京城区汛期常见雨型

Fig.2 The common rain patterns of the urban area during flood season

图 2 中，第 1、2、3 类为单峰型，第 4、5、6 类为双峰型。每种雨型的特征如下：

1 型雨：前单峰型雨。降雨集中在降雨的开始 3~6h 之内，降雨后半程以后降雨就小了，20h 后，降雨过程基本结束。最大峰值时间集中在降雨开始的前 6h 之内。

2 型雨，后单峰型雨。降雨集中在降雨开始后的 18~21h，刚开始降雨时，降雨较少，降雨开始的 18h 以后，降雨量开始增加。而后降雨逐渐增大。

3 型雨，中单峰型雨。降雨集中在降雨开始的 9h 以后，9~12h 之内达到最大。

4 型雨，双峰型雨，有明显的高峰和低谷。第一个集中降雨过程，出现在降雨开始的 3~6h 之内，第一个峰很急，直接达到第一个峰值。第二个集中降雨过程，出现在降雨开始后 18~24h 之内。两个峰值之间有明显的雨量减少过程。

5 型雨，双峰型雨，两个峰值时间间隔较短，而且没有很明显的高峰和低谷，相对比较均匀。

6 型雨，双峰型雨，有两次明显的集中降雨过程。第一个降雨过程，集中在降雨开始的 12~15h 内，第二个降雨过程，集中在降雨开始的 21~24h。后一个峰值较大，主要降雨过程集中在第二个降雨过程中。

由此，得到了北京城区汛期主要发生的 6 种雨型，每种雨型都各有特点和规律。而且该分析结果是以大量实际降雨资料为样本进行统计得到的，基本可以反映出城区各站的主要降雨特征。

3.2 城区整体雨型分布规律分析

得到了城区各站的主要雨型后，对样本中的 50 场实际降雨过程中各个站的雨型，利用模糊模式识别法进行识别，进一步分析城区每场降雨整体分布规律。

经分析，在这 50 场降雨中，城区 14 个站全部集中出现第 1 型和第 3 型的次数最多。具体各站雨型分布，如图 3 所示。图 3 中，城区各站雨型比较类似，各站在降雨开始后 3~9h，集中出现峰值。属于这两种雨型的降雨次数占整个统计样本的 80%以上。该特征与北京城区汛期降雨历史短，降雨集中的特性相吻合，是北京城区汛期出现的主要降雨类型。

该类型降雨，降雨范围大，而且各站出现峰值的时间集中，峰值较大。经统计，样本中属于该类型降雨过程的各站最大 3h 降雨量平均值为 88mm。进一步统计该类型降雨的开始时间可知，该种类型的降雨，经常发生在午后到傍晚，峰值出现在傍晚前后，也就是下班高峰期，晚上 23 点之前，基本完成主要降雨过程。由于该雨型降雨过程中，城区 14 个站都是单峰型雨，强降水出现范围都集中在城区，降雨强度大，对交通影响最大。

城区各站出现双峰型雨的次数相对较少，出现双峰型雨的次数，占整个统计样本的 14%，而且峰值较小。经统计，样本中属于该类型降雨过程的各站最大 3h 降雨量平均值为 44mm，对城区内涝积水影响较小。

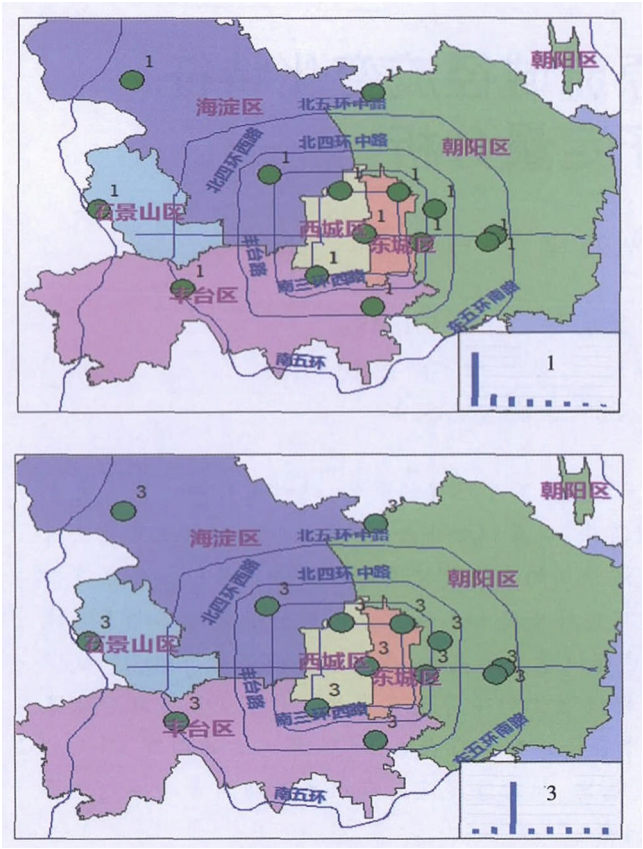

图3　单峰型降雨

Fig.3 The single peak rain

王彬雁[6]等对北京城区汛期气象台站的雨型进行了分析研究，认为北京地区暴雨多单峰，而均匀型和双峰型较少。刘伟东等[7]对2007~2010年北京地区平均逐时降雨量进行研究后认为，北京地区降雨过程存在明显的峰谷区。李建等[8]对北京市1961~2004年降雨资料研究后，认为午后至傍晚北京夏季降雨存在明显的峰值。这些与本文的研究结论相一致。

4　结论

本文基于实测降雨过程，利用动态聚类法统计分析出北京城区汛期发生的主要6种雨型。得到雨型后，利用模糊模式识别法，对实际的降雨过程各站的雨型进行识别，从单站雨型和整体分布角度进行统计分析，从而得到北京城区汛期降雨的时空分布规律。结论如下：

（1）基于2004~2016年50场降雨实测资料，利用动态聚类法统计出来的主要6种雨型，基本反映了汛期北京城区降雨的主要特征。且利用模糊模式识别法，将实际发生的降雨过程进行识别，识别出的结论与实际降雨过程特征基本相符。

（2）北京城区汛期主要降雨类型是前单峰型雨，该类型降雨，城区各站集中出现主要降雨过程，峰值集中且极值雨量大。主要降雨过程发生在午后到傍晚，对交通影响最大。

（3）北京城区汛期发生双峰型雨的次数较少，且极值雨量较小，对城区内涝积水影响相对较小。

综上所述，北京城区汛期主要发生的降雨类型是集中型的单峰型降雨，单峰型降雨与双峰型降雨或均匀型降雨相比，降雨更集中，且雨量更大。更容易造成城区内涝积水，对交通影响更大，需要重点防范。由于本次研究采用的降雨资料仅为城区范围，研究范围相对较小。随着降雨资料的逐渐丰富，样本数量逐渐增大，将掌握更为准确的降雨时空分布的特征规律，为防汛工作提供更为可靠的技术支持。

参考文献：

[1] 吴彰春，岑国平，安智敏，等. 坡面汇流的试验研究 [J]. 水利学报，1995,(7):84-89. (WU Zhangchun, CEN Guoping, AN Zhimin. Experimental study on overland flow [J]. Journal of Hydraulic Engineering, 1995,(7):84-89.(in Chinese))

[2] 岑国平，沈晋，范荣生.城市设计暴雨雨型研究[J]. 水科学进展，1998,(3):41-46.(CEN Guoping, SHEN Jin, FAN Rongsheng. Research on rainfall pattern of urban design storm [J]. Advances in Water Science, 1998,(3):41-46.(in Chinese))

[3] Keifer C J, Chu H H. Synthetic storm pattern for drainage design [J]. Journal of Hydraulics Division ASCE, 1957,83(HY4):1-25.

[4] D F Kibler. Urban stormwater hydrology [A]. American Geophysical Union's Water Resources Monograph 7[C], 1982:48-60.

[5] 汪子棚，王亚娟，白国营，等. 北京市暴雨分区研究 [J]. 北京水务，2011,(5):8-11.(WANG Zipeng, WANG Yajuan, BAI Guoying, et al. Study on the zoning of rainstorm in Beijing [J]. Beijing Water, 2011,(5):8-11.(in Chinese))

[6] 王彬雁，赵琳娜，巩远发，等.北京降雨过程分型特征及短历时降雨重现期研究[J].暴雨灾害，2015,(12):302-308. (WANG Binyan, ZHAO Linna, GONG Yuanfa, et al, Characteristics of temporal pattern and return period of short-duration rainfall at Beijing observatory [J]. Torrential Rain and Disasters, 2015,(12):302-308.(in Chinese))

[7] 刘伟东，尤焕苓，任国玉，等.北京地区精细化的降水变化特征[J].气候与环境研究，2014,(1):61-68.(LIU Weidong, YOU Huanling, REN Guoyu, et al. Subtle precipitation characteristics in Beijing area[J]. Climate and Environmental Research, 2014,(1):61-68.(in Chinese))

[8] 李建，宇如聪，王建捷.北京市夏季降水的日变化特征[J].科学通报，2008,(7):829-832. (LI Jian, YU Rucong, WANG Jianjie. Diurnal variation characteristics of summer precipitation in Beijing [J]. Science in China, 2008,(7):829-832. (in Chinese))

（下转第73页）

nese))

[15] 杨晓华,杨志峰,郦建强,等.水文模型参数识别算法研究及展望[J]. 自然科学进展, 2006,16(6):657-66. (YANG Xiaohua, YANG Zhifeng, LI Jianqiang, et al. Research and prospects of identificable algorithm for hydrological model parameters [J]. Progress in Natural Science, 2006,16(6):657-66. (in Chinese))

[16] 郝振纯,谷军方,杜富慧. 随机 GP 法在概念性水文模型参数优选中的应用[J]. 中山大学学报, 2009,48(6):18-23. (HAO Zhenchun, GU Junfang, DU Fuhui. Application of stochastic GP algorithms optimization to conceptual hydrologic model parameters [J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2009,48(6):18-23. (in Chinese))

[17] 周振,吴志超. 基于呼吸计量法的异养菌动力参数优化估计[J].中国环境科学, 2010,35(6):53-54+125. (ZHOU Zhen, WU Zhichao. Optimization estimation of kinetic parameters for heterotrophic biomass based on batch respirometric method [J]. China Environmental Science, 2010.35(6):53-54+125. (in Chinese))

[18] 白杰,刘和,殷波,等. C/N 比跳控污泥厌氧发酵产酸的数学模型研究[J]. 中国环境科学, 2015.35(1):3303-3309. (BAI Jie, LIU He, YIN Bo, et al. Mathematical modeling of the anaerobic fermentation for VFAs production from sludge by C/N ratio regulation [J]. China Environmental Science, 2015,35(1):3303-3309. (in Chinese))

[19] 邹宗森,施汉昌,李鑫玮. 反硝化池滤池的数值模拟与模型校正[J]. 中国环境科学, 2014,34(10):2505-2512. (ZOU Zongsen, SHI Hanchang, LI Xinwei. Simulation and model calibration of denitrification biological filter [J]. China Environmental Science, 2014,34 (10):2505-2512. (in Chinese))

[20] 王景瑞,胡立堂,尹文杰. iTOUGH2 反演模型在地下水模拟中的应用[J]. 水文地质工程地质, 2015,42(1):35-41. (WANG Jingrui, HU Litang, YIN Wenjie. Application of iTOUGH2 to groundwater modeling [J]. Hydrogeology & Enginering Geology, 2015,42(1):35-41. (in Chinese))

Application of GMS Model in Hydrological and Water Quality Simulation

ZHU Junyan, LI Cuimei, HE Jingxiong, XIANG Song

(*School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou* 215009, *China*)

Abstract:GMS model is a widely-used system in groundwater simulation. This paper took a town in Kunshan as a study area to simulate the hydrological conditions and water quality of the groundwater using the GMS model. The conceptual and mathematical model was established with generalized method, and the main parameters of the model was calibrated with iterative method, and then MODFLOW and MT3DMS were used to simulate and forecast the hydrological condition and water quality and migration trends of pollutants in the study area. The results show that the simulated values are consistent with measured values, which means that the calibrated GMS model can better reproduce the characteristics of the underground water with strong reliability, and can be applied to simulate and forecast migration trends of pollutants in watersheds.

Key words:GMS model; hydrology and water quality; simulation; forecasting; groundwater; pollutant

(上接第 77 页)

Study on Temporal Distribution of Precipitation in Beijing City during Flood Period Based on Dynamic Cluster Analysis and Fuzzy Pattern Recognition

LIU Yuanyuan[1], WANG Yi[2], LIU Hongwei[2], DU Longgang[3], LIU Shu[1], CHAI Fuxin[1]

(1. *China Institude of Water Resources and Hydropower Research, Beijing* 100038, *China*; 2.*Office of Flood Control and Drought Relief Headquarters of Beijing Municipal People's Government, Beijing* 100038, *China*; 3.*Beijing Hydrological Center, Beijing* 100089, *China*)

Abstract:Dynamic cluster analysis and fuzzy pattern recognition method were used to study the spatial distribution characteristics of rain pattern and rainfall at single station, based on the 50 typical rainstorm processes in Beijing City during the flood periods in recent years. The results show that the unimodal type is mainly temporal pattern. The maximum precipitation is concentrated and the extreme value rainfall is large, which is easy to cause waterlogging. The bimodal type is seldom.

Key words: dynamic cluster analysis; fuzzy pattern recognition; characteristics of temporal pattern