

# 径流序列分类识别理论模型的建立与应用\*

康 艳, 蔡焕杰

(西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100)

[摘 要] 模糊划分理论具有严谨的理论和物理基础, 考虑了水文现象发生的必然性与随机性以及划分、识别中的模糊性, 是一种无模式可参考的分类理论。文章应用模糊划分理论模型进行了年径流序列的模糊划分, 根据划分结果, 利用LVQ 进行了年径流过程的枯、平、丰识别, 建立了年径流序列的分类识别模型, 并用黄河上游61年径流水文资料进行了验证。结果表明, 建立的年径流系列分类识别理论模型是可行的。

[关键词] 水文年份; 模糊最优划分; LVQ 神经网络; 年径流分类; 年径流识别

[中图分类号] TV 121<sup>+</sup>. 2 [文献标识码] A [文章编号] 1671-9387(2005)10-0129-05

年径流序列的划分与识别, 是水文计算和用水管理中的一个重要步骤。传统年径流序列划分的方法, 是基于确定论与概率论相结合, 根据频率计算结果按一定的频率取值范围进行划分的。但由于年径流现象既受流域水文气象、自然地理、下垫面等因素的影响, 又受到流域系统水量平衡确定性关系的约束, 其变化具有随机性和丰、枯的模糊性<sup>[1]</sup>, 因此, 该法不能反映径流序列的模糊性。

20世纪90年代, 我国大连理工大学的陈守煜教授创建了模糊水文学, 其既考虑了水文现象、事件的必然性与随机性, 又考虑了其在划分、识别与判决过程中的模糊性, 建立了水文系统成因分析、随机分析与模糊集分析相结合的新理论分析体系与方法<sup>[1-5]</sup>, 为水文分析计算和水资源系统模拟提供了一条新途径。模糊水文学的理论基础是模糊划分理论, 根据模糊划分理论模型可进行给定序列的分类, 国内许多学者<sup>[1-21]</sup>进行了这方面的研究, 但该模型仅给出已知序列的划分结果, 无法进行未知序列的分类识别。学习矢量量化(LVQ)神经网络是在监督状态下对竞争层进行训练的一种学习算法, 可实现对输入向量的自动分类<sup>[2]</sup>。因此, 本研究将上述模糊划分理论模型与LVQ神经网络方法相结合, 建立了年径流序列的分类识别模型, 并用黄河上游61年径流水文资料对建立的模型进行了验证, 现将结果报道如下。

## 1 最优模糊划分的理论与模型

设有  $n$  个样本组成样本集合  $X = \{x_1, x_2, \dots, x_n\}$ , 每个样本有  $m$  个指标特征值,  $\vec{x}_j = (x_{1j}, x_{2j}, \dots, x_{mj})^T$ , 则样本集可用矩阵表示为:  $X = (x_{ij})_{m \times n}$ ; 式中,  $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。将样本集经过(1)式规格化处理变换为  $[0, 1]$  区间的模糊指标特征矩阵:

$$R = (r_{ij})_{c \times n_0} \quad (1)$$
$$r_{ij} = (x_{ij} - x_{\min}) / (x_{\max} - x_{\min})$$

设将样本集划分为  $c$  类, 代表每一类  $m$  个指标的特征值称为分类中心。可用模糊分类中心矩阵表示  $c$  个分类的指标特征, 即  $S = (S_{ih})_{m \times c}$ ; 式中,  $i = 1, 2, \dots, m; h = 1, 2, \dots, c$ ;  $S_{ih}$  表示分类中心  $h$  第  $i$  个指标的特征值, 可用向量表示为  $\vec{S}_h = (S_{1h}, S_{2h}, \dots, S_{mh})^T$ 。设将样本集依据  $m$  个指标特征值划分为  $c$  类, 其模糊划分矩阵为  $U = (U_{hj})_{c \times n}$ ; 式中,  $h = 1, 2, \dots, c; j = 1, 2, \dots, n$ ;  $U_{hj}$  为第  $j$  个样本隶属于第  $h$  类的隶属度,  $U_{hj}$  满足下列约束条件:

$$0 \leq U_{hj} \leq 1, \sum_{h=1}^c U_{hj} = 1, U_{hj} > 0 \quad (2)$$

设第  $j$  个样本与第  $h$  类的差异用欧氏距离表示为:  $r_j - \vec{S}_h = \left[ \sum_{i=1}^m (r_{ij} - S_{ih})^2 \right]^{\frac{1}{2}}$ 。式中,  $\vec{r}_j$  为第  $j$  个样本向量;  $h = 1, 2, \dots, c; i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。根据模糊划分矩阵  $U = (U_{hj})_{c \times n}$ , 第  $j$  个样本以

\* [收稿日期] 2005-01-07  
[基金项目] 国家自然科学基金项目(50179031); 高等学校全国优秀博士学位论文作者专项基金(200052); 西北农林科技大学2004年优秀人才专项基金(04ZR014)  
[作者简介] 康 艳(1977- ), 女, 黑龙江汤原人, 助教, 在读硕士, 主要从事水文水资源研究。

隶属度  $U_{hj}$  隶属于第  $h$  类, 将隶属度定义为权重, 用权距离表示第  $j$  个样本与第  $h$  类的差异程度为  $d(\vec{r}_j, \vec{S}_h) = U_{hj} \left[ \sum_{i=1}^m (r_{ij} - S_{ih})^2 \right]^{\frac{1}{2}}$ 。以全体样本集对全体分类中心之间的权距离平方和最小为目标, 则目标函数为<sup>[1-5]</sup>:

$$\begin{aligned} \min F &= \min_{j=1, h=1}^n \left[ \sum_{i=1}^m U_{hj} (r_{ij} - S_{ih})^2 \right]^{\frac{1}{2}} \\ &= \min_{j=1, h=1}^n U_{hj}^2 \sum_{i=1}^m (r_{ij} - S_{ih})^2 \end{aligned} \quad (3)$$

(3) 式中未知数为模糊分类中心矩阵  $(S_{ih})_{m \times c}$  与模糊划分矩阵  $(U_{hj})_{c \times n}$ 。分别假设  $(S_{ih})_{m \times c}$  和  $(U_{hj})_{c \times n}$  已知, 用极值原理可得

$$S_{ih} = \frac{\sum_{j=1}^n U_{hj}^2 r_{ij}}{\sum_{j=1}^n U_{hj}^2} \quad (4)$$

$$U_{hj} = \frac{1}{\sum_{i=1}^m \left[ \frac{(r_{ij} - S_{ih})^2}{\sum_{i=1}^m (r_{ij} - S_{ih})^2} \right]^{\frac{1}{2}}} \quad (5)$$

模糊划分矩阵和分类中心矩阵求解步骤如下:

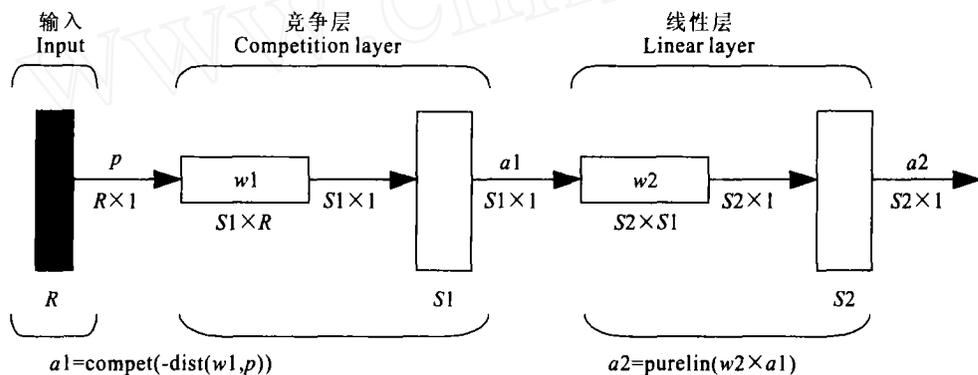


图1 LVQ 网络结构图

Fig 1 LVQ network structure

设输入向量为  $P$ , 目标向量为  $t$ , 且包含  $S_2$  个元素, 即分成  $S_2$  个类, 目标向量必须只有 1 个元素为 1, 其余元素为 0。LVQ 神经网络仿真步骤如下:

**网络初始化。** 利用函数 `initlvq` 可对 LVQ 网络进行初始化, 即  $[w_1, w_2] = \text{initlvq}(P, S_1, t)$ 。其中  $w_1, w_2$  分别为竞争层和线性层的权值。

**网络训练。** LVQ 网络初始化后, 可进行网络的训练。利用函数 `trainlvq` 对 LVQ 网络进行训练, 即:  $tp = [\text{disp- freq max- cycle lr}]$ ;  $[w_1, w_2] = \text{trainlvq}(w_1, w_2, P, t, tp)$ 。其中 `disp- freq` 为显示频率; `max- cycle` 为训练最大步数; `lr` 为学习率。

**网络仿真。** 利用函数 `simulvq` 可完成 LVQ 网

络的仿真, 即  $[a_1, a_2] = \text{simulvq}(P, w_1, w_2)$ , 式中,  $a_1$  为竞争层的输出结果;  $a_2$  为线性层输出结果。

选择划分分类数  $c$  和计算控制精度  $\epsilon$ 。给定一个满足 (2) 式的初始模糊划分矩阵  $(U_{hj}^{(0)})$ ; 把  $U_{hj}^{(0)}$  代入 (4) 式得聚类中心  $S_{ih}^{(1)}$ ; 把  $S_{ih}^{(1)}$  代入 (5) 式得模糊划分矩阵  $U_{hj}^{(1)}$ ; 逐个比较  $U_{hj}^{(1)}$  与  $U_{hj}^{(0)}$  的对应元素, 若对应元素最大差值的绝对值满足计算控制精度, 即  $\max |U_{hj}^{(1)} - U_{hj}^{(0)}| \leq \epsilon$ , 则迭代结束,  $U_{hj}^{(1)}$  可作为满足计算控制精度要求的最优模糊划分矩阵  $U_{hj}^*$ 。否则, 使  $U_{hj}^{(0)} = U_{hj}^{(1)}$ , 重复 ~ 步, 直至满足计算控制精度要求为止; 将所求得的  $U_{hj}^*$  代入 (4) 式得最优模糊划分中心  $S_{ih}^*$ 。

## 2 LVQ 神经网络模型

LVQ 网络由 2 层构成, 第 1 层为竞争层, 第 2 层是将竞争层的分类结果传递到用户定义的目标分类上。通过网络学习, 将输入向量中与目标向量相近的向量分离出来, 其网络结构如图 1 所示。图 1 中,  $R$  为输入数;  $S_1$  为竞争层神经元;  $S_2$  为线性层神经元。文献 [6] 给出了用 Matlab 5.0 进行 LVQ 网络计算的过程。

## 3 模型应用

### 3.1 年径流序列的模糊分类

本研究采用文献 [1] 中 1919~ 1980 年共 61 年的黄河上游月平均径流系列资料对模型进行验证。首先, 应用模糊划分理论模型进行年径流序列的模糊划分。本研究将年径流过程分为枯 (1)、平 (2)、丰 (3) 3 类, 即划分分类数  $c = 3$ , 取计算控制精度  $\epsilon = 0.000001$ , 应用 Fortran 语言编制了相应的计算程序, 经计算可得最优分类中心和最优划分矩阵。最优

分类中心见表1。最优划分矩阵给出了某样本从属于某分类的隶属度, 因此将最优划分矩阵每列元素的最大者取为1, 其他取为0, 表明该列表示的年径流序列从属于该行代表的分类, 经过处理后, 最优划分矩阵见表2。年径流序列的分类结果见表3。

表1 黄河上游径流序列的最优分类中心

Table 1 Optimum classification central of runoff series in upper reach of Yellow River  $m^3/s$

分类 Kinds	月份Months												平均值 Average
	1	2	3	4	5	6	7	8	9	10	11	12	
1	1 690.67	1 570.63	1 377.66	1 158.33	657.61	337.54	284.19	255.80	362.80	472.49	759.33	1 170.63	841.47
2	1 961.87	2 067.35	1 999.12	1 700.51	854.39	431.02	347.13	316.14	406.44	567.15	904.55	1 438.88	1 082.88
3	2 377.99	2 453.18	2 845.75	1 987.91	981.16	488.61	360.39	344.46	449.71	638.41	1 004.14	1 374.16	1 275.49

表2 黄河上游第1~ 61年径流最优划分矩阵

Table 2 Matrix of optimum classification of annual runoff (# 1- # 61) in upper reach of Yellow River

年序 No.	分类 Kinds			年序 No.	分类 Kinds			年序 No.	分类 Kinds		
	枯水年 Dry year	平水年 Normal year	丰水年 Wet year		枯水年 Dry year	平水年 Normal year	丰水年 Wet year		枯水年 Dry year	平水年 Normal year	丰水年 Wet year
1	0	1	0	22	0	0	1	43	0	0	1
2	0	1	0	23	1	0	0	44	1	0	0
3	0	1	0	24	1	0	0	45	0	0	1
4	1	0	0	25	0	1	0	46	0	0	1
5	0	1	0	26	1	0	0	47	1	0	0
6	1	0	0	27	0	1	0	48	0	0	1
7	0	1	0	28	0	0	1	49	0	0	1
8	1	0	0	29	0	1	0	50	0	0	1
9	1	0	0	30	0	1	0	51	1	0	0
10	1	0	0	31	0	0	1	52	1	0	0
11	1	0	0	32	0	1	0	53	0	1	0
12	1	0	0	33	0	0	1	54	1	0	0
13	1	0	0	34	1	0	0	55	1	0	0
14	1	0	0	35	0	1	0	56	1	0	0
15	0	1	0	36	0	1	0	57	0	0	1
16	0	1	0	37	0	0	1	58	0	0	1
17	0	0	1	38	1	0	0	59	1	0	0
18	0	1	0	39	1	0	0	60	0	0	1
19	0	1	0	40	0	0	1	61	0	0	1
20	0	0	1	41	1	0	0				
21	1	0	0	42	0	1	0				

表3 黄河上游第1~ 61年径流序列分类结果

Table 3 Classification results of annual runoff (# 1- # 61) in upper reach of Yellow River

年序 No.	分类 Kinds										
1	2	12	1	23	1	34	1	45	3	56	1
2	2	13	1	24	1	35	2	46	3	57	3
3	2	14	1	25	2	36	2	47	1	58	3
4	1	15	2	26	1	37	3	48	3	59	1
5	2	16	2	27	2	38	1	49	3	60	3
6	1	17	3	28	3	39	1	50	3	61	3
7	2	18	2	29	2	40	3	51	1		
8	1	19	2	30	2	41	1	52	1		
9	1	20	3	31	3	42	2	53	2		
10	1	21	1	32	2	43	3	54	1		
11	1	22	3	33	3	44	1	55	1		

### 3.2 年径流过程的识别

模糊划分理论模型虽然给出了年径流序列枯、平、丰的模糊分类,但是不能进行给定年径流过程的枯、平、丰识别。LVQ神经网络采用LVQ学习规则,通过学习,将输入向量中与目标向量相近的向量分离出来,可为年径流过程的枯、平、丰提供一条非线性的识别途径。本研究通过计算机反复试验,选用

disp. freq= 20, max. cycle= 400, lr= 0.05, 竞争层神经元数 $s_1= 12$ 。应用第1~52年分类结果进行网络训练,其训练结果如表4所示。应用第53~61年分类结果进行网络检验,其检验结果如表5所示。从表5可以看出,LVQ神经网络的识别结果与序列的从属类别吻合,表明用LVQ神经网络可以进行年径流序列枯、平、丰的识别。

表4 黄河上游第1~52年径流资料的LVQ神经网络训练结果

Table 4 Training results of LVQ network of Annual runoff (# 1-# 52) in upper reach of Yellow River

年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification	年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification	年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification
1	2	2	19	2	2	37	3	3
2	2	2	20	3	3	38	1	1
3	2	2	21	1	1	39	1	1
4	1	1	22	3	3	40	3	3
5	2	2	23	1	1	41	1	1
6	1	1	24	1	1	42	2	2
7	2	2	25	2	2	43	3	3
8	1	1	26	1	1	44	1	1
9	1	1	27	2	2	45	3	3
10	1	1	28	3	3	46	3	3
11	1	1	29	2	2	47	1	1
12	1	1	30	2	2	48	3	3
13	1	1	31	3	3	49	3	3
14	1	1	32	2	2	50	3	3
15	2	2	33	3	3	51	1	1
16	2	2	34	1	1	52	1	1
17	3	3	35	2	2			
18	2	2	36	2	2			

表5 黄河上游第53~61年径流资料的LVQ神经网络检验结果

Table 5 Identifying results of LVQ network of Annual runoff (# 53-# 61) in upper reach of Yellow River

年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification	年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification	年序 No.	模糊划分 Fuzzy classification	网络识别 Network identification
53	2	2	56	1	1	59	1	1
54	1	1	57	3	3	60	3	3
55	1	1	58	3	3	61	3	3

## 4 结论

年径流过程不仅受流域水文气象、自然地理、下垫面等诸多确定性因素的影响,而且其变化具有随机性和丰枯的模糊性。本研究将模糊水文学原理和学习向量量化神经网络结合起来,用于年径流过程枯、平、丰的分类与识别。在分类中,模糊划分理论模型因为综合考虑了水文现象的必然性与随机性,以及在划分、识别过程中的模糊性,且有严谨的理论和物理基础,所以可应用于年径流枯、平、丰的划分。但模糊划分理论模型只能进行模糊分类,并不能进行模式识别,而LVQ神经网络采用LVQ学习规则,

能够将输入向量中与目标向量相近的向量分离出来,可为年径流过程的枯、平、丰提供一条非线性的识别途径。因此,采用LVQ神经网络进行枯、平、丰的识别具有充分的理论依据。实例验证表明,LVQ神经网络的识别结果与序列的从属类别吻合,表明LVQ神经网络可以进行年径流序列枯、平、丰的识别。本研究的模型虽然需要经过若干次复杂的迭代计算,但其计算过程在计算机上非常容易实现。因此,本研究所建立的模型可用于实测年径流序列枯、平、丰的分类与识别,同时也为实际年径流过程枯、平、丰水文年的分类与识别提供了一条途径。

## [参考文献]

- [1] 陈守煜 模糊水文学与水资源系统模糊优化原理[M]. 大连: 大连理工大学出版社, 1990
- [2] 陈守煜 区域水资源可持续利用评价理论模型与方法[J]. 中国工程科学, 2001, 3(2): 33- 38
- [3] 陈守煜 模糊划分理论分析模型及其在水文上的应用[J]. 水利学报, 1991, (12): 25- 30
- [4] 陈守煜 中长期水文预报综合分析理论模式与方法[J]. 水利学报, 1997, (8): 15- 21.
- [5] 陈守煜 区域可持续发展模糊模式识别评价量化理论[J]. 大连理工大学学报, 1999, 39(2): 5- 8
- [6] 楼顺天, 施 阳 基于MATLAB的系统分析与设计——人工神经网络[M]. 西安: 西安电子科技大学出版社, 1998

## The establishment and application of theory model for the classification and identification of runoff series

KANG Yan, CAI Huan-jie

(College of Water Conservancy and Architectural Engineering, North China University of Technology, Beijing 100048, China)

**Abstract:** Fuzzy classified theory has strict theory and physics base. It integrates inevitability and randomness and fuzzy classification and identification of hydrological phenomenon. The author applied fuzzy classified theory model to fuzzily classify annual runoff series. According to classified results, using LVQ to identify dry year, normal year or wet year of annual runoff series, identification and classification model of annual runoff series was established, and 61 years' water-data in upper reach of Yellow River were confirmed. The results show that the model can be used to classify and identify runoff.

**Key words:** water year; fuzzy optimum classification; LVQ Networks; runoff classification; runoff identification

(上接第128页)

**Abstract ID:** 1671-9387(2005)10-0125-EA

## The application research of DPSA and large-scale system decomposition-coordination for cascaded hydroelectric short-term optimal scheduling

LILiang<sup>1</sup>, HUANG Qiang<sup>1</sup>, XIAO Yan<sup>2</sup>, XIAO Zhi-juan<sup>1</sup>

(1 Key Laboratory of Northern Water Resources, Environment and Ecology, Ministry of Education, Xi'an University of Technology, Xi'an, Shaanxi 710048, China;  
2 Wujiang Hydropower Development Corporation Ltd., Guiyang, Guizhou 550002, China)

**Abstract:** The mathematical model in cascaded hydroelectric short-term optimal scheduling is introduced. Then two methods, DPSA and Large-Scale System decomposition-coordination, are contrasted and analyzed in detail from calculating idea and process. Both methods will arrive convergence because of the identical idea. Finally, an example of two methods is given, whose results can be referred in real project. The two methods can be chosen on the need of precision and time etc.

**Key words:** short-term optimal scheduling of plant; POA arithmetic; DPSA arithmetic; large-scale system decomposition-coordination