

黄河内蒙古段开河日期预报模型及应用

冀鸿兰¹, 张傲妲^{1, 2}, 高瑞忠¹, 张宝森³, 徐 晶¹

(1. 内蒙古农业大学水利与土木建筑工程学院, 内蒙古 呼和浩特 010018 ;

2. 内蒙古第一水文地质工程地质勘察院, 内蒙古 呼和浩特 010050 ;

3. 黄河水利委员会黄河水利科学研究院, 河南 郑州 450003)

摘要 对黄河内蒙古段的冰情特点进行分析, 选用 1970—1971 年度至 2007—2008 年度共 38 个冬季的冰情资料, 提取合适的预报因子, 分别采用人工神经网络模型和多元线性回归模型, 对黄河内蒙古段的开河日期进行预报。结果表明, 神经网络模型和多元线性回归模型预报样本合格率分别为 86.7% 和 80.0%, 神经网络模型的预报精度高于多元线性回归模型, 更适用于黄河内蒙古段开河日期预报。

关键词 冰凌, 黄河内蒙古段, 开河日期, 多元线性回归, 神经网络

中图分类号 :P338+.4

文献标识码 :A

文章编号 :1006-7647(2012)02-0042-04

Application of the break-up date prediction model in the Inner Mongolia Reach of the Yellow River//JI Hong-lan¹, ZHANG Ao-da^{1, 2}, GAO Rui-zhong¹, ZHANG Bao-sen³, XU Jing¹(1. Water Conservancy and Civil Engineering, College of Inner Mongolia Agricultural University, Hohhot 010018, China ; 2. First Hydrogeology and Engineering Geology Exploration Institute of Inner Mongolia, Hohhot 010050, China ; 3. Yellow River Institute of Hydraulic Research, Yellow River Conservancy Commission, Zhengzhou 450003, China)

Abstract : The ice flood characteristics in Inner Mongolia reach of the Yellow River are analyzed in the papers. Selecting from 1970—1971 year to 2007—2008 year ice conditions data and extracting the appropriate predictors, we use artificial neural network model and multiple linear regression model to forecast the date of break-up in Inner Mongolia reach of the Yellow River. The results show that prediction sample pass rates of the artificial neural network model and multiple linear regression model are respectively 86.7% and 80.0%. This indicates that forecasting accuracy of artificial neural network model is better than the one of multiple linear regression model, and the artificial neural network model is applied to the ice forecast of break-up date.

Key words : ice flood ; Inner Mongolia Reach of the Yellow River ; break-up date ; multiple linear regression ; artificial neural network

黄河内蒙古段处于黄河流域最北端, 总体上是一个“几”字的头部。由于所处的地理位置和特殊的水文气象条件, 黄河内蒙古段几乎每年都产生凌汛灾害, 因此, 除采取一系列工程防凌措施外, 探寻科学可行的冰情预报模型对于防凌实践具有重要的指导意义。

1 黄河内蒙古段冰期气温变化特点

黄河内蒙古段处于黄河流域最北端, 海拔高程在千米以上, 冬半年被蒙古冷高压所控制, 加之高空西北气流的引导, 往往会受大陆冷空气南侵的影响, 天气寒冷。由于河道是自南而北流向, 南北纬差异在 4°以上, 因此气温在地域上存在上游高、下游低的

差别, 而且随时间的变化也呈现出入冬时下游降温比上游降温早的特点, 初春回暖时, 情况则相反。黄河内蒙古段各水文站相对位置如图 1 所示。由于昭君坟水文站 20 世纪 70 年代后的资料系列不完整, 无法收集相关数据, 因此本文选取 1970—1971 年度至 2007—2008 年度共 38 个冬季的实测数据, 分别对巴彦高勒站、三湖河口站和头道拐站 3 个具有代表性的水文站的冰期气温及流量变化特点进行分析, 以此来探寻黄河内蒙古段的冰情变化特点。图 2 和图 3 分别是巴彦高勒站、三湖河口站和头道拐站的气温及流量变化过程曲线(图中年序数是指 1970—1971 年度至 2007—2008 年度共 38 个冬季的顺序号, 其中冬季指从当年 11 月至次年 3 月)。

基金项目 国家自然科学基金(51060001); 内蒙古自治区自然科学基金(2010MS0618); 黄河水利科学研究院中央级公益性科研院所专项(HKY-JBYW-2010-04); 内蒙古水利厅科技计划(201007)

作者简介 冀鸿兰(1970—), 女, 内蒙古呼和浩特人, 教授, 博士, 主要从事水文水资源与工程模糊集研究。E-mail :honglanji@163.com

通讯作者 张傲妲(1985—), 女, 内蒙古赤峰人, 硕士研究生, 主要从事水文水资源与工程模糊集研究。E-mail :zad2010@yeah.net

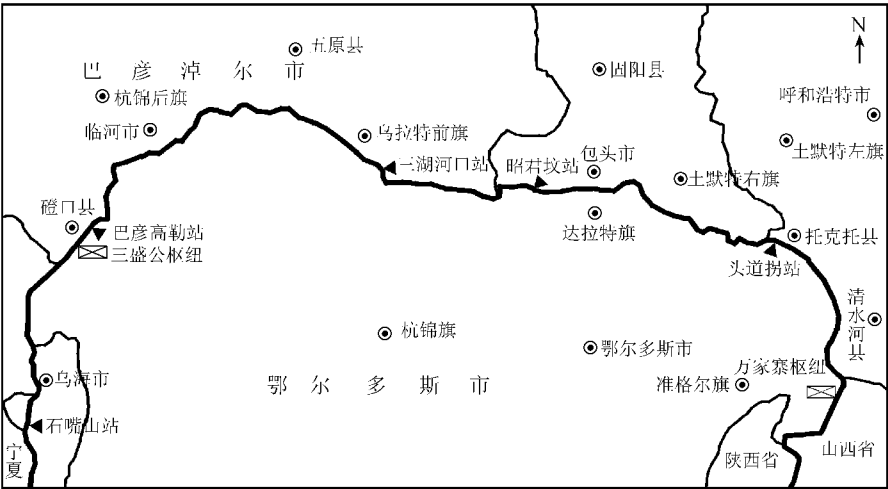


图1 黄河内蒙古段各水文站相对位置示意图

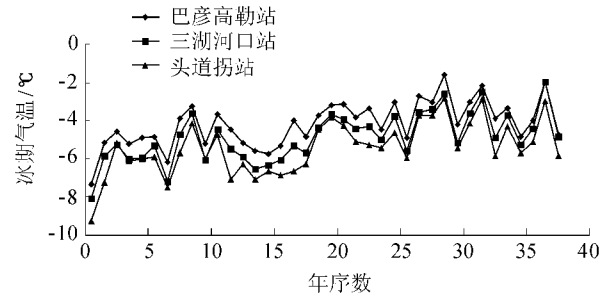


图2 巴彦高勒站、三湖河口站和头道拐站的冰期气温变化曲线

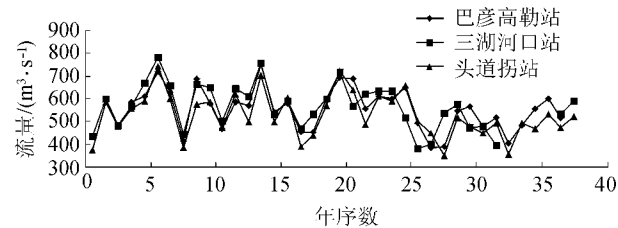


图3 巴彦高勒站、三湖河口站和头道拐站的冰期流量变化曲线

从图2、图3中的冰期气温是从当年11月到次年3月的平均气温可知,巴彦高勒站、三湖河口站、头道拐站的气温变化趋势基本一致,但巴彦高勒站多年平均气温较三湖河口站和头道拐站高,故其流凌、封冻时间较晚,开河时间较早,其他两个站则相反。从图3可看出,流量变化趋势同气温变化趋势基本一致。

气温和流量与流凌时间及封、开河时间也存在很大关系。河段封河天数的多少,主要取决于当年11月、12月及次年3月的气温及流量的变化,即气

温低,流量小,流凌、封冻早,开河晚;气温高,流量大,流凌、封冻晚,开河早。表1是巴彦高勒站、三湖河口站和头道拐站的流凌、封河和开河日期统计表。

从图2、图3和表1可看出,气温的多年变化趋势与巴彦高勒站、三湖河口站和头道拐站的流凌、封河、开河日期的变化趋势相吻合,即巴彦高勒站的平均气温高于其他两个站,故其流凌、封河日期较晚,而开河日期早于其他两站。

2 开河日期预报模型

2.1 BP神经网络模型

神经网络对于解决非线性复杂问题具有很大的优势。近年来,神经网络理论模型在水资源领域中的应用也越来越多,如年径流预报、洪水预报、水质评价、河流污染物浓度预报、冰塞水位预测、地下水资源评价等^[1-7]。国内对于冰凌预报的研究并不多,但由于冰凌变化受多种因素的影响,至今没有一种广泛认可的预报方法。陈守煜等^[8-12]提出模糊优选神经网络预报模型,预报精度较高。赵晓慎等^[13]在分析物理成因的基础上,采用模糊识别、聚类及优选理论综合模糊成因的方法建立了冰凌预报模型。本文将BP神经网络模型用于黄河内蒙古段开河日期预报,并与多元线性回归模型进行对比。

2.2 多元线性回归模型

多元线性回归分析是通过建立数学模型描述一个变量如何随其他变量的变化而变化。

表1 巴彦高勒站、三湖河口站和头道拐站凌期特征日期统计

水文站	流凌日期			封河日期			开河日期		
	最早	最晚	平均	最早	最晚	平均	最早	最晚	平均
巴彦高勒	11月14日	12月28日	11月31日	11月23日	1月13日	12月18日	2月24日	3月27日	3月14日
三湖河口	11月6日	12月2日	11月19日	11月21日	1月1日	12月8日	3月6日	3月28日	3月21日
头道拐	11月6日	12月2日	11月19日	1月21日	1月11日	12月12日	3月1日	4月1日	3月22日

2.2.1 模型方程

假设因变量 y 与多个自变量 $x_1, x_2, \dots, x_k$ 之间有如下关系：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + e \quad (1)$$

式中： $\beta_j (j=1, 2, \dots, k)$ 为 $k+1$ 个回归系数； e 为随机误差项。式(1)即为多元线性回归模型方程。

2.2.2 模型计算步骤

- a. 根据研究目的确定因变量和自变量。
- b. 收集样本数据。
- c. 从收集到的样本数据中应用最小二乘法得到样本回归方程：

$$\hat{y}(x_1, x_2, \dots, x_k) = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k \quad (2)$$

- d. 对总体回归方程进行总检验和回归系数检验。

总检验的原假设为 $\beta_1 = \beta_2 = \dots \beta_k = 0$ 检验的统计量为

$$F = \frac{S_{SR}/k}{S_{SE}/(n - k - 1)} \quad (3)$$

式中： S_{SR} 残差平方和； S_{SE} 为回归方差平方和。

回归系数检验的原假设为 $\beta_j = 0$,检验的统计量为 $t_j = \frac{b_j - \beta_j}{\sqrt{\sum (x - \bar{x})^2}}$ 。

3 预报模型的应用

3.1 预报因子的筛选

根据预报因子与预报对象相关系数的大小选取预报因子,各预报因子与预报对象的相关系数见表2。

由表2的计算结果得到各水文站选取的预报因子如下：

- a. 巴彦高勒站预报因子： x_1 为2月16—25日逐日累计气温， x_2 为2月16—25日流量均值， x_3 为封冻期最大冰厚， x_4 为2月1日至气温稳定转正日期的天数。开河历时 y 为计算起点2月1日至开河日期的天数。

表2 巴彦高勒站、三湖河口站和头道拐站各预报

因子与预报对象的相关系数

水文站	因子	相关系数
巴彦高勒	2月16—25日石嘴山流量均值	0.0252
	1月9—26日逐日累计气温	-0.0393
	1月9—26日流量均值	0.1298
	1月9—26日水位均值	0.0442
	2月1日至气温转正日期天数	0.6131
	2月16—25日逐日累计气温	-0.6844
	2月16—25日流量均值	-0.5021
	最大冰厚	0.7364
	2月21日—3月2日巴彦高勒流量均值	0.0317
	1月9—26日逐日累计气温	-0.6283
三湖河口	1月9—26日流量均值	0.6052
	1月9—26日水位均值	-0.0158
	2月1日至气温转正日期天数	0.7903
	2月21日—3月2日逐日累计气温	-0.3597
	2月21日—3月2日流量均值	0.0505
	最大冰厚	0.5210
	2月21日—3月2日三湖河口流量均值	-0.1055
头道拐	1月9—26日逐日累计气温	-0.2943
	1月9—26日流量均值	0.3320
	1月9—26日水位均值	0.1478
	2月1日至气温转正日期天数	0.6252
	2月21日—3月2日逐日累计气温	-0.5222
	2月21日—3月2日流量均值	0.7430
	最大冰厚	0.5408

- b. 三湖河口站预报因子： x_1 为1月9—26日逐日累计气温， x_2 为1月9—26日流量均值， x_3 为封冻期最大冰厚， x_4 为2月1日至气温稳定转正日期天数。开河历时 y 为计算起点11月1日至封河日期的天数。

- c. 头道拐站预报因子： x_1 为2月21日—3月2日逐日累计气温， x_2 为2月21日—3月2日流量均值， x_3 为封冻期最大冰厚， x_4 为2月1日至气温稳定转正日期天数。开河历时 y 为计算起点2月1日至开河日期的天数。

3.2 模型应用及比较

图4为应用BP神经网络模型对巴彦高勒站、三

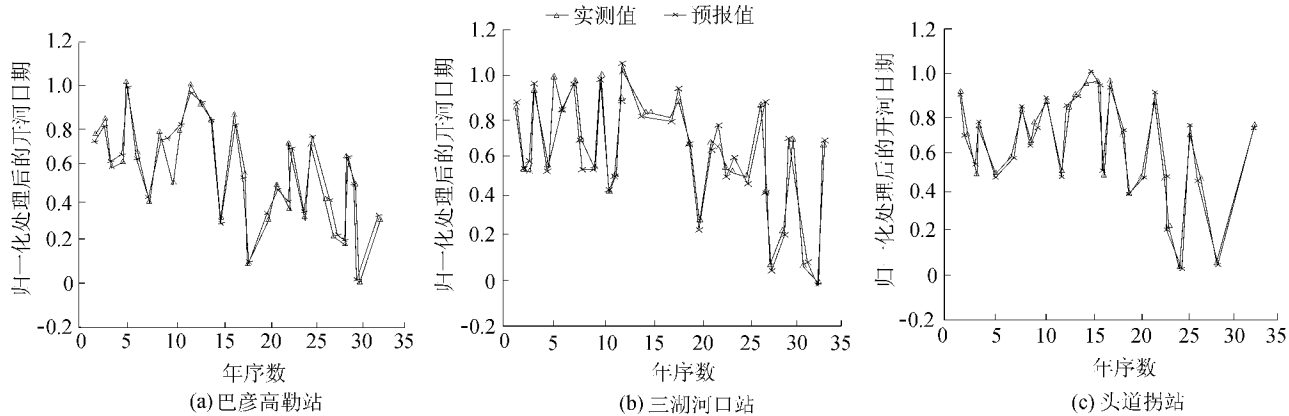


图4 巴彦高勒站、三湖河口站和头道拐站开河日期预报神经网络模型训练拟合结果

湖河口站和头道拐站开河日期进行的网络训练拟合图,可以看出,巴彦高勒站、三湖河口站和头道拐站开河日期预报值与实测值均拟合得比较好,说明神经网络模型可用于黄河内蒙古段开河日期预报,且适用性较强。

运用多元线性回归模型进行开河日期预报,得到的多元线性回归方程通过了统计检验,结果如表3所示。

表3 巴彦高勒站、三湖河口站和头道拐站多元线性回归模型统计检验结果

水文站	回归方差平方和	残差平方和	<i>F</i>	<i>P</i>
巴彦高勒	1 818.490	506.260	38.315	0.000
三湖河口	965.875	381.704	44.283	0.000
头道拐	919.930	543.585	16.359	0.000

从表3可知,巴彦高勒站、三湖河口站和头道拐站的多元线性回归方程通过了总检验,因为统计量*F*对应的显著性水平*P*值均等于0,小于显著性水平5%,所以拒绝所有回归系数等于零的假设,说明开河历时与所选取的预报因子之间存在线性关系。

将1970—1971年度至2002—2008年度共38个冬季的前33个冬季的实测数据作为训练样本,后5个冬季的实测数据作为预报检验样本,训练结果及其与多元线性回归模型预报检验结果对比分别见表4和表5。

根据水文情报预报规范^[14],巴彦高勒站、三湖河口站和头道拐站的预报精度均在80%以上,预报方案均为甲等方案,两种模型的精度均较高,说明两种模型对于黄河内蒙古河段的开河日期预报适用性较强。

从表5对比结果可知,神经网络模型的预报检

验合格率均值为86.7%,而多元线性回归模型的预报检验合格率均值为80.0%,说明神经网络模型的预报精度较多元线性回归模型高,更适于黄河内蒙古河段的开河日期预报。

表5 巴彦高勒站、三湖河口站和头道拐站开河日期预报模型预报样本检验合格率对比

水文站	神经网络模型 预报样本合格率/%	多元线性回归模型 预报样本合格率/%
巴彦高勒	80	80
三湖河口	100	80
头道拐	80	80

4 结 论

a. 巴彦高勒站、三湖河口站和头道拐站的冰期气温变化总体趋势是巴彦高勒站多年平均气温较其他两个站高,所以其流凌、封河日期较晚,而开河日期较早,其他两个站变化趋势则相反。

b. 从所选取的预报因子来看,气温和流量是影响冰情变化的主控因素。

c. 从两种模型的预报检验结果来看,神经网络模型的预报精度高于多元线性回归模型。由于黄河封开河受多种因素的影响,所以神经网络模型对于黄河内蒙古段开河日期预报的适用性较强,更适合解决上述受复杂因素影响的问题。

d. 考虑到黄河内蒙古段的自然地理条件及受各种因素的影响,应采取必要的人类工程活动(诸如通过调整水库的调度方式)来改变其封、开河的情势,以避免或减少冰凌灾害。

表4 巴彦高勒站、三湖河口站、头道拐站开河日期预报模型预报检验结果对比

水文站	年度	实测值	预报值		误差/d		合格否	
			神经网络模型	多元线性回归模型	神经网络模型	多元线性回归模型	神经网络模型	多元线性回归模型
巴彦高勒	2003—2004	3月8日	3月1日	3月8日	2	0	合格	合格
	2004—2005	3月18日	3月23日	3月2日	5	2	合格	合格
	2005—2006	3月9日	3月4日	3月9日	5	0	合格	合格
	2006—2007	2月24日	3月4日	3月4日	8	8	不合格	不合格
	2007—2008	3月16日	3月12日	3月14日	4	2	合格	合格
三湖河口	2003—2004	3月14日	3月14日	3月16日	0	3	合格	合格
	2004—2005	3月21日	3月16日	3月23日	5	2	合格	合格
	2005—2006	3月19日	3月17日	3月16日	2	3	合格	合格
	2006—2007	3月17日	3月15日	3月17日	2	0	合格	合格
	2007—2008	3月21日	3月16日	3月15日	5	6	合格	不合格
头道拐	2003—2004	3月14日	3月21日	3月17日	7	3	不合格	不合格
	2004—2005	3月19日	3月18日	3月24日	1	5	合格	合格
	2005—2006	3月15日	3月2日	3月21日	5	6	合格	合格
	2006—2007	3月16日	3月15日	3月15日	1	1	合格	合格
	2007—2008	3月14日	3月16日	3月18日	2	4	合格	合格

(下转第88页)