

DOI:10.13476/j.cnki.nsbdtq.2021.0054

赵菲菲,张青青,张宇,等.基于贝叶斯网络的黄河径流预测[J].南水北调与水利科技(中英文),2021,19(3):511-519. ZHAO F F,ZHANG Q Q,ZHANG Y, et al. The Yellow River runoff forecast based on Bayesian network[J]. South-to-North Water Transfers and Water Science & Technology,2021,19(3):511-519. (in Chinese)

# 基于贝叶斯网络的黄河径流预测

赵菲菲<sup>1,2</sup>,张青青<sup>1,2</sup>,张宇<sup>3</sup>,石旭芳<sup>1,2</sup>,钟德钰<sup>1,2</sup>

(1. 青海大学水利电力学院,西宁 810016;2. 青海大学三江源生态与高原农牧业国家重点实验室,西宁 810016;  
3. 清华大学水沙科学与水利水电工程国家重点实验室,北京 100084)

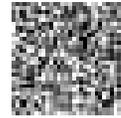
**摘要:**基于黄河流域1979—2018年的ERA-Interim再分析气候与水文数据,以及CMIP5中10个气候模式下3种典型排放情景(RCP2.6、RCP4.5、RCP8.5)的全球气候变化数据,采用离散数据的处理方法,建立黄河流域贝叶斯网络模型,推断黄河流域近40余年来气候要素对径流的影响概率,预测黄河流域未来径流量。结果表明:1979—2018年黄河天然径流量呈减小趋势,基于贝叶斯网络分区间概率预测预报的径流量也呈减小趋势;黄河流域的不同区间(低、中、高)径流量对气候的敏感程度不同,但径流始终与降水相关性最高;在RCP2.6情景下,黄河流域未来20年、60年的径流量为585.50亿、588.57亿 $m^3$ ;在RCP4.5情景下,其值为585.42亿、587.53亿 $m^3$ ;在RCP8.5情景下,其值为593.50亿、585.11亿 $m^3$ 。

**关键词:**贝叶斯网络;气候因子;径流;不确定性问题;黄河流域

中图分类号:TV213;P338

文献标志码:A

开放科学(资源服务)标志码(OSID):



近年来,国际研究表明全球大部分河流受气候变化影响显著,降水、蒸发等气候变化会引起径流等一系列地表水文现象的变化<sup>[1-3]</sup>。例如:姚文艺<sup>[4]</sup>研究表明,1997—2006年黄河上中游径流量减少约23.5%主要是受气候因素的影响;徐文馨等<sup>[5]</sup>基于偏差校正的气候模式集合驱动水量平衡模型,表明降雨是影响长江流域径流变化的重要因素;Ahmadi等<sup>[6]</sup>使用人工神经网络研究气候变化对Kan流域径流的影响,表明2010—2040年在RCP2.6、RCP4.5和RCP8.5排放情景下径流增加分别为4%、26%和2%;Azari等<sup>[7]</sup>研究了伊朗北部戈根罗纳河流域径流对气候变化的响应,表明2040—2069年在A1F1、A2、B1排放情境下,降雨和气温的变化导致年径流增加5.8%、2.8%、9.5%。由此可见,气候变化对径流变化具有十分重要的作用,且作用

机理十分复杂。国内外关于气候要素对水文现象的影响已有许多研究,但对于气候要素与径流的因果关系及相互影响概率的量化仍比较模糊。并且,在水文预报研究中,常用的水量平衡模型、小波分析法、神经网络和模糊推理法等方法一般仅提供水文过程确定性预报,不能对预报结果的不确定性进行定量描述。

针对上述问题并考虑到气候与水文系统之间的关联性和不确定性,利用欧洲中期天气预报中心的再分析气候与径流数据集,采用能够通过概率推理量化不确定性的贝叶斯网络(Bayesian network, BN)<sup>[8]</sup>构建气候要素与黄河流域径流间的因果关系模型,以概率网络的形式明确气候要素与黄河径流间的因果关系及影响概率。BN模型不仅能够评估网络中变量之间的影响概率,还能够进行概率预报。

收稿日期:2020-12-01 修回日期:2021-03-02 网络出版时间:2021-03-12

网络出版地址:https://kns.cnki.net/kcms/detail/13.1430.TV.20210312.1146.002.html

基金项目:国家重点研发计划项目(2017YFC0404303)

作者简介:赵菲菲(1997—),女,湖南邵东人,主要从事水文及水资源研究。E-mail:2282882173@qq.com

通信作者:钟德钰(1970—),男,甘肃金昌人,教授,博士,主要从事泥沙运动基本理论与水沙运动数值模拟。E-mail:zhongdy@tsinghua.edu.cn

例如:张铭等<sup>[9]</sup>对一个水电站水库的中长期径流进行概率水文预报,表明贝叶斯概率预报有利于决策者考虑不确定性并提高预报精度;梁忠民等<sup>[10]</sup>以密赛流域为研究对象,实现了基于高斯混合模型的概率预报,预报结果精度较高。因此,本文进一步利用第五次国际耦合模式比较计划(CMIP5)的 10 种全球气候模式数据,根据所构建的气候要素与径流之间的因果关系模型,对黄河流域未来径流变化趋势进行推理预测。

## 1 研究区域概况

黄河流域陆域范围位于北纬  $32^{\circ}20' \sim 42^{\circ}62'$ , 东经  $76^{\circ}14' \sim 119^{\circ}70'$ , 面积为 79.5 万  $\text{km}^2$ , 约占国土总面积的 8.3%。黄河流域地域辽阔,气候与水文要素时空分布不均,差异明显,20 世纪 90 年代以来,实测径流量与天然径流量均呈减少趋势,年际变化和年内分配差异大,导致黄河流域水资源供需矛盾尤为紧张。

## 2 数据与方法

### 2.1 数据收集与整理

用于概率推理的气候径流数据采用欧洲中期天气预报中心的 ERA-Interim 再分析资料,空间分辨率为  $0.75^{\circ} \times 0.75^{\circ}$ , 时间跨度为 1979—2018 年,在概率推理与径流预测时将 ERA 再分析日数据集处理为月平均值,包括气温、比湿、蒸发、垂向风速、气压、降水和径流等 7 个变量。另外,根据已有研究在刻画未来黄河流域气候时选用的 CMIP5 全球气候模式<sup>[11-12]</sup>,选取如表 1 所示的 10 个 CMIP5 气候模式中排放情景 RCP2.6、RCP4.5、RCP8.5 下相应的气候要素,利用集合平均的方式对黄河未来径流进行预报。不同气候模式的结果可能存在偏差:Li 等<sup>[13]</sup>提出了等距累积分布函数匹配方法(EDCDFm)来校正气候模式中的数据偏差;高谦<sup>[14]</sup>利用 EDCDFm 和累积分布函数变换(CDF-t)对 CMIP5 中的 5 个全球模式进行偏差订正比较,表明了 EDCDFm 方法订正效果较好,因此,本文运用 EDCDFm 方法,以 ERA 再分析数据作为基准,对所选取的 10 个气候模式的气候要素数据进行偏差校正。需指出的是,由于 ERA 再分析数据及 10 个全球气候模式的分辨率各不相同,故在进行数据校正前将所有格点数据统一插值为水平分辨率为  $2.81^{\circ} \times 2.81^{\circ}$  的月均值数据。

表 1 CMIP5 全球气候模式信息

Tab. 1 CMIP5 global climate model information table

CMIP5 模式名称	所属国家	分辨率	所选气候要素
MPI-ESM-LR	德国	$1.87^{\circ} \times 1.87^{\circ}$	降水、 压强、 温度、 蒸发、 比湿、 风速
IPSL-CM5A-LR	法国	$3.75^{\circ} \times 1.87^{\circ}$	
GISS-E2-R	美国	$2.50^{\circ} \times 2.00^{\circ}$	
CanESM2	加拿大	$2.81^{\circ} \times 2.81^{\circ}$	
HadGEM2-ES	英国	$2.48^{\circ} \times 0.937^{\circ}$	
MIROC-ESM	日本	$2.81^{\circ} \times 2.81^{\circ}$	
MIROC-ESM-CHEM	日本	$2.81^{\circ} \times 2.81^{\circ}$	
MIROC5	日本	$1.40^{\circ} \times 1.40^{\circ}$	
NORESM1-M	挪威	$2.50^{\circ} \times 1.87^{\circ}$	
NORESM1-ME	挪威	$2.50^{\circ} \times 1.87^{\circ}$	

### 2.2 研究方法

#### 2.2.1 数据离散方法

为简化模型,需将连续数据离散化。ChiMerge 是一种基于统计学思想的自底向上合并的离散算法。刘晓明等<sup>[15]</sup>用等宽法和 ChiMerge 两种方法对比离散数据,并构建 BN 进行简单的预测分析,表明 ChiMerge 算法在离散过程中考虑了类别信息,是一种较为科学的有监督的离散方法。ChiMerge 算法具体过程为:首先根据变量数值大小,对连续变量进行排序,每个变量的初始取值即为初始的区间状态;然后计算相邻分箱卡方值  $\chi$ ,合并卡方最小的两项,循环直至遇到以下停止条件<sup>[16]</sup>。(1)分箱个数等于指定的分箱数目(max\_interval)。(2)最小卡方值大于卡方阈值(chi2\_threshold)。其卡方计算公式为

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \quad (1)$$

式中: $m$  表示比较的区间数目,此处为 2; $k$  为类别数; $A_{ij}$  表示  $i$  区间  $j$  类别的样本个数; $E_{ij}$  表示期望的  $A_{ij}$  的频数,  $E_{ij} = \frac{R_i \times C_j}{C}$  ( $R_i$  表示  $i$  区间,  $R_i = \sum_{j=1}^k A_{ij}$ ;  $C_j$  表示  $j$  类,  $C_j = \sum_{i=1}^m A_{ij}$ ;  $C$  表示总样本数)。

#### 2.2.2 静态贝叶斯网络

贝叶斯网络由网络结构  $G$  和参数  $\theta$  构成,网络结构  $G=(I, B)$ 。其中: $I$  为节点集合,每个节点对应一个随机变量; $B$  为有向边的集合;参数  $\theta$  表示每个节点的条件概率分布,能反映节点与其父节点的概率依赖关系的强弱。贝叶斯网络规定,当一个节点的父节点概率分布确定之后,该节点条件独立于其非后代节点构成的任何节点子集。这一性质便于简化变量间的联合概率分布计算,提高推理效率。

令  $X_i$  为网络中节点  $i$  代表的随机变量,利用链规则,则包含  $n$  个变量的联合概率分布计算公式<sup>[17-18]</sup>为

$$P(X_1, X_2, \dots, X_n) = P(X_1) P(X_2 | X_1) \cdot P(X_3 | X_1, X_2), \dots, P(X_n | X_1, X_2, \dots, X_{n-1}) = \prod_{i=1}^n P(X_i | X_1, X_2, \dots, X_{i-1}) \quad (2)$$

利用上述性质,即  $P(X_n | X_1, X_2, \dots, X_{i-1}) = P(X_i | \pi(X_i))$ ,公式(2)可简化<sup>[19]</sup>为

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi(X_i)) \quad (3)$$

式中: $\pi(X_i)$ 表示节点  $X_i$  的父节点。

### 3 基于贝叶斯网络的气候-径流系统模型

基于贝叶斯网络模型构建气候-径流系统模型的流程见图 1,首先确定影响黄河径流的气候变量,通过专家知识等建立 BN 结构,然后利用极大似然估计法进行参数学习,用变量消元法进行贝叶斯推理。

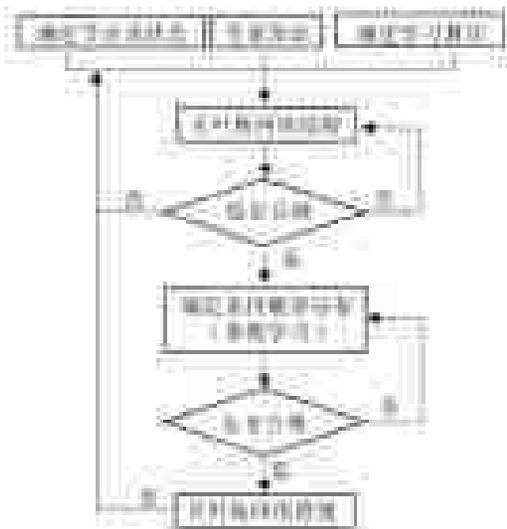


图 1 气候-径流系统模型构建流程

Fig. 1 Construction diagram of the climate-runoff Bayesian model

#### 3.1 基于贝叶斯网络模型的建立

##### 3.1.1 贝叶斯网络变量的确定

结合专家知识库和其他学者对黄河流域气候与径流相关关系的研究成果,最终确定了 6 个气候因子,与径流构成变量集  $\{E, T_p, T, P, V, H_u, R\}$ ,其中: $E$ 代表蒸发量; $T_p$ 代表总降雨量; $T$ 代表温度; $P$ 代表压强; $V$ 代表垂向风速; $H_u$ 代表比湿; $R$ 代表径流。最大程度保证原始信息不丢失地离散数据,是得到训练数据集的关键。采用 ChiMerge 算法将连续变量分为 3 个离散区间(表 2),分别代表低、中、高 3 种状态,用 L、M、H 表示。

表 2 各变量离散区间(月值)

Tab. 2 Discrete interval of variables (month value)

变量	离散区间		
	L	M	H
$E/\text{mm}$	$<8.99$	$8.99 \sim 41.31$	$>41.31$
$T_p/\text{mm}$	$<7.63$	$7.63 \sim 45.79$	$>45.79$
$T/\text{K}$	$<267.86$	$267.86 \sim 284.18$	$>284.18$
$P/\text{Pa}$	$<79\,500.19$	$79\,500.19 \sim 79\,775.45$	$>79\,775.45$
$V/(\text{Pa} \cdot \text{s}^{-1})$	$<0.077$	$0.077 \sim 0.091$	$>0.091$
$H_u/(\text{kg} \cdot \text{kg}^{-1})$	$<1.21 \times 10^{-3}$	$1.21 \times 10^{-3} \sim 3.45 \times 10^{-3}$	$>3.45 \times 10^{-3}$

##### 3.1.2 贝叶斯网络结构确定及参数学习

郭晨露等<sup>[19]</sup>研究青藏高原珠峰地区蒸发皿蒸发量与其他气候因子的关系,发现蒸发皿蒸发量在干季主要受风速和温度影响,受相对湿度影响比较小。柳春等<sup>[20]</sup>利用多元回归方法分析黄河流域蒸发量与气候因子的关系,发现蒸发量的上升主要与气温上升有关,而蒸发量的下降主要与风速减小有关,蒸发对温度和风速有直接依赖关系。姜师等<sup>[21]</sup>通过集合预报发现风速、压强、850 hPa 温度以及 850 hPa 与 925 hPa 比湿是大范围暴雨的敏感因子。韦晓伟等<sup>[22]</sup>探究了渭河流域气候水文要素间相关结构及其演化,研究表明蒸发-降水、降水-径流和径流-蒸发有互相关性。区域蒸发量增加则降水量也会随之增加。根据上述关于不同气候要素之间关联特征的研究,结合专家知识,用 netica 软件构建 BN 模型(图 2),表示系统中变量的相互关系。该层次结构模型共 3 层:第一层是径流的间接影响因素温度、比湿、压强和风速( $T, H_u, P, V$ );第二层是径流的直接影响因素蒸发和降水( $E, T_p$ )。

网络结构与训练数据集确定后,参数学习采用完备数据集下的最大似然估计,以避免采用贝叶斯估计造成学习参数过拟合的问题,得出各个节点的条件概率表。BN 中的推理过程是通过这些条件概率表确定的。

##### 3.2 概率推理

基于构建的网络结构与参数学习,用 netica 软件对该网络结构进行层次间后验概率推理<sup>[23]</sup>。首先是第一层的温度、比湿、压强和风速在 3 种状态下对第二、三层变量的概率推理,其次是第二层的蒸发、降水在三种状态下对第三层径流的概率推理。

在全球气候变暖的大背景下,黄河流域年平均温度也表现为明显上升趋势<sup>[24]</sup>,当温度依次确定为 L、M、H 态,这代表一种上升趋势,推理温度对第二层变量的后验概率为  $P(E|T)$ 、 $P(T_p|T)$ ,对第三层径流的后验概率为  $P(R|T)$ ,概率值见表 3。由表

可知:温度从 L 态依次变为 H 态时,  $P(E_H | T)$  由 25% 增长为 49%,  $P(T_{P_H} | T)$  由 27% 增长为 47%, 即蒸发和降水的 H 态概率越来越大,  $P(R_H | T)$  由 24% 增长为 39%, 径流 H 态概率也越来越大, 这是

由于温度升高, 蒸发量和降水增多, 径流也随之增多, 温度升高加速了整个水循环。同理可得第一层其他变量与第二、三层变量的关系: 比湿和第二、三层气候要素成正相关, 相反, 压强、风速成负相关。

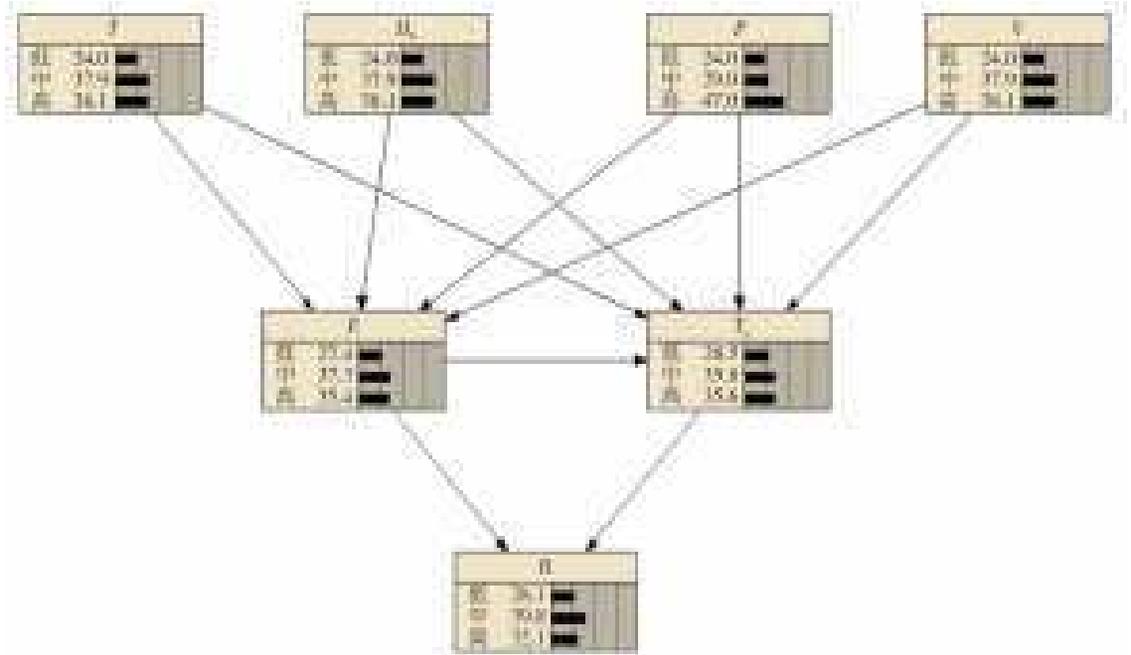


图 2 贝叶斯网络模型

Fig. 2 Bayesian network model

表 3 温度对蒸发、降水和径流的后验概率

Tab. 3 The posterior probability of temperature on evaporative precipitation and runoff

T	$P(E T)$			$P(T_P T)$			$P(R T)$		
	$[E_L]$	$E_M$	$E_H]$	$[T_{P_L}$	$T_{P_M}$	$T_{P_H}]$	$[R_L]$	$R_M$	$R_H]$
$T_L$	0.45	0.30	0.25	0.38	0.35	0.27	0.31	0.45	0.24
$T_M$	0.21	0.52	0.28	0.27	0.43	0.30	0.31	0.39	0.31
$T_H$	0.23	0.28	0.49	0.25	0.29	0.47	0.24	0.37	0.39

表 4 蒸发对降水、径流的后验概率

Tab. 4 The posterior probability of evaporative on precipitation and runoff

E	$P(T_P E)$			$P(R E)$		
	$[T_{P_L}$	$T_{P_M}$	$T_{P_H}]$	$[R_L]$	$R_M$	$R_H]$
$E_L$	0.39	0.33	0.28	0.36	0.53	0.11
$E_M$	0.26	0.48	0.26	0.36	0.37	0.27
$E_H$	0.23	0.25	0.52	0.13	0.33	0.54

表 5 降水对径流的后验概率

Tab. 5 The posterior probability of precipitation on runoff

$T_P$	$P(R T_P)$		
	$[R_L]$	$R_M$	$R_H]$
$T_{P_L}$	0.25	0.56	0.19
$T_{P_M}$	0.38	0.39	0.23
$T_{P_H}$	0.20	0.27	0.52

其他气候要素状态未确定, 蒸发对同层降水的后验概率表示为  $P(T_P | E)$ , 对下一层径流的后验概率表示为  $P(R | E)$ , 概率值见表 4。由表 4 可知: 蒸发从 L 态依次变为 H 态时,  $P(T_{P_H} | E)$  由 28% 增长为 52%, 说明蒸发量影响降水, 蒸发量越多降水随之而多的概率越大;  $P(R_H | E)$  也由 11% 增长为 54%, 说明蒸发增多使得径流处于 H 态的概率也会变大, 这可能是蒸发增多使降水增多, 从而间接使得径流增多的可能增大。降水对下一层径流的后验概率表示为  $P(R | T_P)$ , 概率值见表 5。由表 5 可知, 降水和径流成正相关, 降水由 L 态依次变为 H 态时, H 态径流的概率越来越大。表 6 显示的是蒸发、降水对径流的联合后验概率。由表 6 可知, 当  $E_H$  与  $T_{P_H}$  共同作用于径流时,  $R_H$  的概率由 2% 变为 71%, 说明蒸发和降水共同作用对径流影响更大。

列举所有各气候因子(L、M、H)对径流(L、M、H)事件发生的概率, 如图 3 所示, 通过比较同一径流状态概率大小, 确定该状态径流影响最紧密的气候因子。  $R_L$  表示径流深低于 2.22 mm 时, 比较  $R_L$  的 18 个条件概率值大小, 得到 L 态径流的最大影响因子及状态是降水  $T_{P_M}$ , 该事件可表示为  $P(R_L | T_{P_M})$ , 发生的概率为 38%;  $R_M$  表示径流深为 2.22~6.63 mm 时, M 态径流的最大影响因子是降水  $T_{P_L}$ , 该事件可

表示为  $P(R_M | T_{P_L})$ , 其概率为 56%;  $R_H$  表示径流深大于 6.63 mm 时, 由于  $P(R_H | T_{P_H})$  和  $P(R_H | E_H)$  的概率值非常接近, 分别为 52% 和 54%, 得到 H 态径流的最大影响因子是降水和蒸发。综上所述, 径流始终与降水关系最密切且 H 态径流受蒸发和降水影响都大。

因果链是结点与父结点之间的有序关系<sup>[25]</sup>。20 世纪 70 年代以来, 黄河径流下降趋势逐渐显著<sup>[26]</sup>, 层层逆向推理 L 态径流的因果链探明因果关系, 逆向推理的因果链属于最大后验假设问题。从径流节点出发, 比较第二层气候要素对 L 态径流的后验概率, 逆向依次找到对子节点发生影响最大的父节点, 得到:  $R_L$  的最大影响因子是  $T_{P_M}$ , 其发生概率为 38.4%;  $T_{P_M}$  的最大影响因子是  $H_{UM}$ , 其发生概率为 48%。经过逆向溯源得  $R_L$  的因果链为  $H_{UM} \rightarrow$

$T_{P_M} \rightarrow R_L$ , 见图 4。

表 6 蒸发、降水对径流的后验概率

Tab. 6 The posterior probability of precipitation and evaporative on runoff

E	$T_P$	$P(R E, T_P)$		
		$[R_L]$	$R_M$	$R_H$
$E_L$	$T_{P_L}$	0.25	0.73	0.02
	$T_{P_M}$	0.50	0.46	0.04
	$T_{P_H}$	0.33	0.33	0.33
$E_M$	$T_{P_L}$	0.18	0.57	0.25
	$T_{P_M}$	0.44	0.31	0.25
	$T_{P_H}$	0.42	0.26	0.32
$E_H$	$T_{P_L}$	0.33	0.33	0.33
	$T_{P_M}$	0.16	0.47	0.37
	$T_{P_H}$	0.04	0.25	0.71

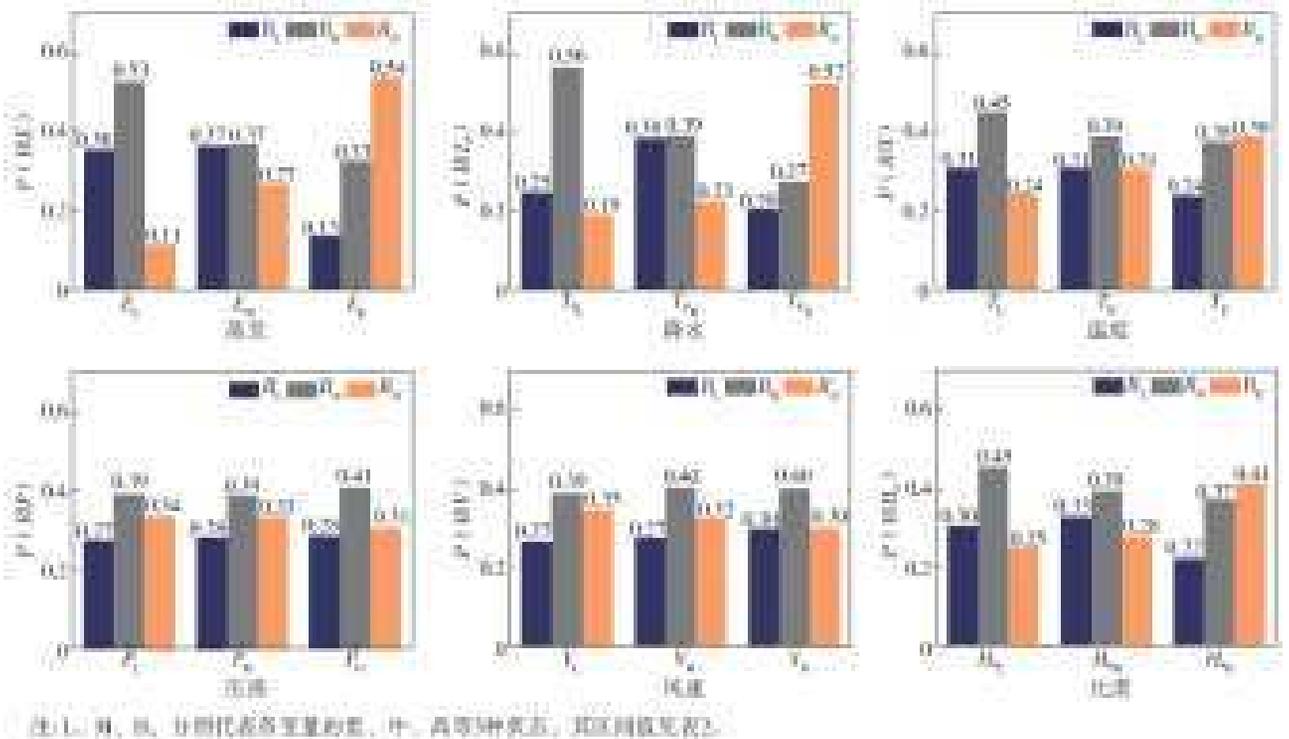


图 3 所有气候因子对径流的影响

Fig. 3 Influence of all meteorological factors on runoff

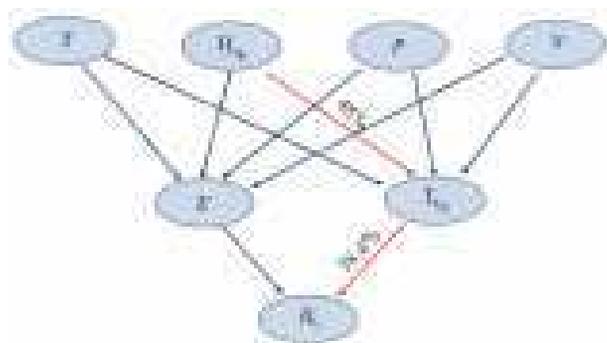


图 4 L 态径流的因果链分析

Fig. 4 Causal chain analysis of L- state runoff

### 3.3 模型预测

随着各种全球气候模式数据产生, 大范围流域水文预报应运而生。但由于全球气候情景数据具有很强的不确定性, 故采取多种模式集合的方法预测径流。在预测径流量的贝叶斯网络模型中, 将概率推理时的 3 个离散区间增加为 12 个以提高预测精度, 离散后的气候月均值数据集输入 BN 预测模型能计算径流区间的概率, 累加 12 个区间的概率与区间的径流之积即得月径流深。该预测方法结合了集合预报与概率预报, 极大降低了预报的不确定性。

计算月径流量的公式为

$$Q=1\ 000F\sum_{k=1}^{12}r_k P_k \quad (4)$$

式中:  $Q$  为径流量,  $\text{m}^3$ ;  $F$  为黄河流域面积,  $F=79.5$  万  $\text{km}^2$ ;  $r_k$  为  $k$  区间的径流深,  $\text{mm}$ ;  $P_k$  为  $k$  区间的预报概率。

为了验证离散型贝叶斯网络具有一定的预测效果,划分 1979—2018 年的 ERA-interim 数据预测月径流量来验证模型,并采用决定系数  $R^2$  与均方根误差 RMSE 来评价预测效果。以 1979—2008 年的 ERA 再分析数据集(月值)做训练集,2009—2018 年做验证集,得到该验证期的预测值与实际值对比见图 5,模型的  $R^2$  为 0.514, RMSE 为 33.16 亿  $\text{m}^3$ ,可以看出模型的拟合效果较好,预测值与实际值的偏差较小。

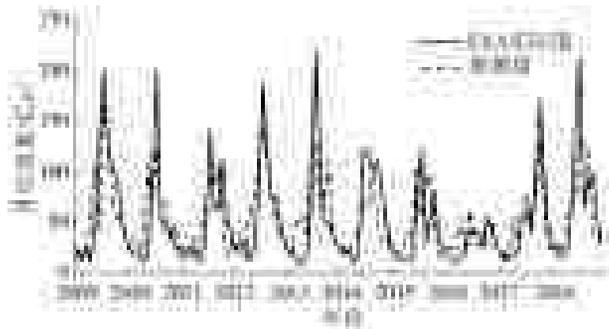


图 5 验证期(2009—2018 年)的预测值与实际值对比  
Fig. 5 Comparison chart of predicted and actual values during the verification period

将 ERA 再分析数据(1979—2018 年)作为训练

集,利用 CMIP5 10 个全球气候模式数据预测的黄河流域历史及未来年径流结果见图 6。左边是 CMIP5 历史数据(1979—2005 年)预测的黄河流域径流,能够捕捉到下降趋势,与同时期 ERA 再分析径流的趋势一致,并且径流量的拟合范围(图 6 灰色区域)可以覆盖再分析数据的滑动平均值,因此预测结果比较可靠。右边是 CMIP5 3 种 RCP 情景数据预测的黄河流域未来 20 年和 60 年年径流量的趋势和周期平均值。3 种 RCP 情景方案的预测趋势均呈微弱减少趋势,但各时段有差异:在 RCP2.6 情景方案中,2020—2026 年呈下降趋势,2026—2070 年呈上升趋势,2070—2080 年呈下降趋势;在 RCP4.5 情景方案中,2020—2040 年和 2040—2080 年两个时段都呈下降趋势;在 RCP8.5 情景方案中,2020—2040 年呈上升趋势,2040—2080 年时段呈下降趋势。在 RCP2.6 情景下,黄河流域未来 20 年、60 年天然径流的周期平均值为 585.50 亿、588.57 亿  $\text{m}^3$ ;在 RCP4.5 情景下,其值为 585.42 亿、587.53 亿  $\text{m}^3$ ;在 RCP8.5 情景下,其值为 593.50 亿、585.11 亿  $\text{m}^3$ 。李晓宇等<sup>[27]</sup>在忽略下游孙口、利津等水文站径流的基础上,用花园口站代表黄河流域预测 2050 年和 2100 年的天然径流量为 533.0 亿  $\text{m}^3$  和 501.2 亿  $\text{m}^3$ ;姚文艺等<sup>[28]</sup>推求的花园口断面 2020、2030、2040 和 2050 年的天然径流量分别为 513.7 亿、534.8 亿、563.5 亿和 537.0 亿  $\text{m}^3$ ,由此可以看出整个黄河流域的天然径流结果在合理范围内。

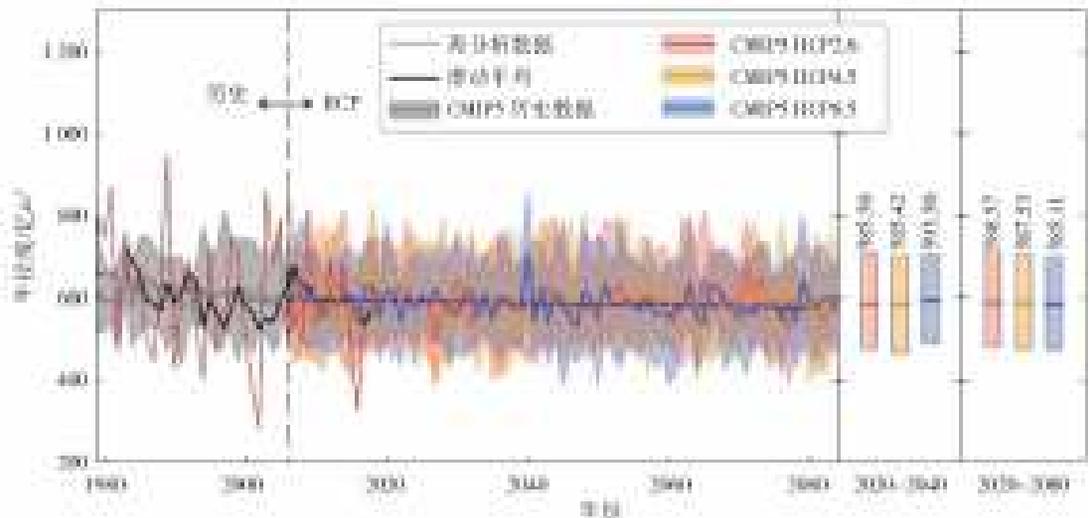


图 6 基于气候-径流贝叶斯模型的黄河流域历史及未来径流预测结果

Fig. 6 The historical and future runoff prediction results of the Yellow River basin based on the climate-runoff Bayesian model

## 4 结 论

本文利用 ERA 再分析数据与 CMIP5 数据,采用 BN 模型捕获气候因素之间的依赖关系并推断其对径流的影响,对 RCP2.6、RCP4.5、RCP8.5 等 3

种排放情景下黄河未来径流的变化情况进行预测和分析,主要结论如下:

(1)气候-径流系统中任意变量状态改变都会引起其他变量变化,如温度状态由低到高,蒸发量和降水增多概率变大,径流随之增多的概率也变大,表明

温度、蒸发、降水与径流成正相关。其他气候要素与径流关系:比湿与径流成正相关,压强和风速与径流成负相关。

(2)比较所有气候因子对径流事件发生的概率发现,黄河流域的不同径流量区间对气候的敏感程度不同,但都与降水关系最密切。

(3)经过层层逆推理得到的  $H_{\text{VM}} \rightarrow T_{\text{PM}} \rightarrow R_{\text{L}}$  因果链表明,比湿对降水的影响较为密切,降水相对于蒸发对径流影响更大。

(4)利用 CMIP5 的 10 种模式历史数据集平均计算了 1979—2005 年黄河流域径流趋势和径流范围,结果与实际较符合,表明利用贝叶斯网络研究黄河流域径流的变化情况是合适的。RCP2.6、RCP4.5、RCP8.5 等 3 种情景预测的未来径流结果存在趋势差异,总体上 3 种情景都呈微弱减少趋势;未来 60 年,黄河流域的天然径流量预计在 587 亿  $\text{m}^3$  左右。

本文没有考虑人类活动对流域径流量的影响,可在后续的研究工作中进一步进行深入分析。

#### 参考文献(References):

- [1] WU C S,JI C C,SHI B W,et al. The impact of climate change and human activities on streamflow and sediment load in the Pearl River basin[J]. International Journal of Sediment Research, 2019, 34(4): 307-321. DOI:10.1016/j.ijsrc.2019.01.002.
- [2] 王国庆,张建云,管晓祥,等. 中国主要江河径流变化成因定量分析[J]. 水科学进展, 2020, 31(3): 313-323. (WANG G Q,ZHANG J Y,GUAN X X,et al. Quantifying attribution of runoff change for major rivers in China[J]. Advances in Water Science, 2020, 31(3): 313-323. (in Chinese)) DOI: 10.14042/j.cnki.32.1309.2020.03.001.
- [3] DOLL P,SCHIMMEL H M. How is the impact of climate change on river flow regimes related to the impact on mean annual runoff: A global-scale analysis[J]. Environmental Research Letters, 2012, 7(1): 14-37. DOI: 10.1088/1748-9326/7/1/014037.
- [4] 姚文艺. 黄河流域水沙变化情势分析与评价[M]. 郑州:黄河水利出版社,2011. (YAO W Y. Analysis and evaluation of water and sediment variation in the Yellow River basin[M]. Zhengzhou: Yellow River Water Conservancy Press,2011. (in Chinese))
- [5] 徐文馨,陈杰,顾磊,等. 长江流域径流对全球升温 1.5 °C 与 2.0 °C 的响应[J]. 气候变化研究进展, 2020, 16(6): 690-705. (XU W X,CHEN J,GU L,et al. The response of runoff in the Yangtze River basin to global warming 1.5 °C and 2.0 °C [J]. Climate Change Research Progress, 2020, 16(6): 690-705. (in Chinese))
- [6] AHMADI M,VAZIRI B M,AHMADI H,et al. Assessment of climate change impact on surface runoff, statistical downscaling and hydrological modeling[J]. Physics and Chemistry of the Earth, Parts A/B/C, 2019, 114:102800. DOI:10.1016/j.pce.2019.09.002.
- [7] AZARI M,MORADI H,SAGHAFIAN B,et al. Climate change impacts on streamflow and sediment yield in the north of Iran[J]. Water and Energy International, 2015, 61(1): 1-11. DOI: 10.1080/02626667.2014.967695.
- [8] 赵瑞娟,龙昱帆,赵凯,等. 贝叶斯网络在水利工程中的应用[J]. 东北水利水电, 2020, 38(4): 62-65, 72. (ZHAO R J, LONG Y F, ZHAO K, et al. Application of Bayesian network in hydraulic project[J]. Water Resources & Hydropower of Northeast China, 2020, 38(4): 62-65, 72. (in Chinese)) DOI: 10.14124/j.cnki.dbslsd22-1097.2020.04.022.
- [9] 张铭,李承军,张勇传. 贝叶斯概率水文预报系统在中长期径流预报中的应用[J]. 水科学进展, 2009, 20(1): 40-44. (ZHANG M, LI C J, ZHANG Y C. Application of the Bayesian statistic hydrological forecast system to middle and long-term runoff forecast[J]. Advances in Water Science, 2009, 20(1): 40-44. (in Chinese)) DOI: 10.3321/j.issn:1001-6791.2009.01.006.
- [10] 梁忠民,戴荣,王军,等. 基于贝叶斯模型平均理论的水文模型合成预报研究[J]. 水力发电学报, 2010, 29(2): 114-118. (LIANG Z M, DAI R, WANG J, et al. Study on forecast combination of different hydrological models by Bayesian model averaging[J]. Journal of Hydroelectric Engineering, 2010, 29(2): 114-118. (in Chinese))
- [11] 周文翀,韩振宇. CMIP5 全球气候模式对中国黄河流域气候模拟能力的评估[J]. 气象与环境学报, 2018, 34(6): 42-55. (ZHOU W C, HAN Z Y. Evaluation of the capability of CMIP5 global climate model to simulate climate in the Yellow River basin of China[J]. Journal of Meteorology and Environment, 2018, 34(6): 42-55. (in Chinese)) DOI: 10.3969/j.issn.1673-503X.2018.06.005.
- [12] 王国庆,张建云,金君良,等. 基于 RCP 情景的黄河流域未来气候变化趋势[J]. 水文, 2014, 34(2): 8-13. (WANG G Q, ZHANG J Y, JIN J L, et al. Future climate change trend of the Yellow River basin based on RCP scenario[J]. Hydrology, 2014, 34(2): 8-13. (in Chinese)) DOI: 10.3969/j.issn.1000-0852.2014.02.003.
- [13] LI H B, SHEFFIELD J, WOOD F E. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching[J]. 2010, DOI: 10.12006/j.issn.1673-1719.2019.207.

- 115(D10);D10101. DOI:10.1029/2009JD012882.
- [14] 高谦. 多模式动力降尺度与偏差订正相结合的中国区域极端温度模拟及预估[D]. 南京:南京信息工程大学, 2017. (GAO Q. Simulation and prediction of regional extreme temperature in China based on multi model dynamic downscaling and deviation correction [D]. Nanjing: Nanjing University of Information Engineering, 2017. (in Chinese))
- [15] 刘晓明, 李盼池, 刘显德, 等. 贝叶斯网络参数学习中的连续变量离散化方法研究[J]. 计算机与数字工程, 2018, 46(5): 992-996. (LIU X M, LI P C, LIU X D, et al. Research on discretization method of continuous variables in Bayesian network parameter learning[J]. Computer and Digital Engineering, 2018, 46(5): 992-996. (in Chinese)) DOI: 10.3969/j.issn.1672-9722.2018.05.029.
- [16] KERBER R. ChiMerge: Discretization of numeric attributes[J]. Learning, Inductive, 1992: 123-128.
- [17] 张连文, 郭海鹏. 贝叶斯网引论[M]. 北京: 科学出版社, 2006. (ZHANG L W, GUO H P. Introduction to Bayesian Nets[M]. Beijing: Science Press, 2006. (in Chinese))
- [18] LEE S H, KANG J E, PARK C S, et al. Multi-risk assessment of heat waves under intensifying climate change using Bayesian networks [J]. International Journal of Disaster Risk Reduction, 2020, 50: 101704 DOI: 10.1016/j.ijdrr.2020.101704.
- [19] 郭晨露, 马耀明, 马伟强, 等. 青藏高原珠峰地区戈壁下垫面上实际蒸散发量和蒸发皿蒸发量的关系研究[J]. 高原气象, 2017, 36(1): 79-86. (GOU C L, MA Y M, MA W Q, et al. Relationship between actual evapotranspiration and pan evaporation in the Gobi land surface of the Qomolangma region of the Qinghai-Xizang Plateau[J]. Plateau Meteorology, 2017, 36(1): 79-86. (in Chinese)) DOI: 10.7522/j.issn.1000-0534.2016.00020.
- [20] 柳春, 王守荣, 梁有叶, 等. 1961—2010 年黄河流域蒸发皿蒸发量变化及影响因子分析[J]. 气候变化研究进展, 2013, 9(5): 327-334. (LIU C, WANG S R, LIANG Y Y, et al. Analysis of pan evaporation change and the influence factors in the Yellow River basin in 1961-2010[J]. Climate Change Research, 2013, 9(5): 327-334. (in Chinese)) DOI: 10.3969/j.issn.1673-1719.2013.05.00.
- [21] 姜帅, 张亮, 李娇娇, 等. 基于 ECMWF 集合预报的一次大范围暴雨敏感因子分析[J]. 广东气象, 2019, 41(5): 1-5. (JIANG S, ZHANG L, LI J J, et al. Analysis of factors sensitive to a large-scale heavy rain based on ECMWF ensemble forecast[J]. Guangdong Meteorology, 2019, 41(5): 1-5. (in Chinese)) DOI: 10.3969/j.issn.1007-6190.2019.05.001.
- [22] 韦晓伟, 张洪波, 辛琛, 等. 变化环境下流域气象水文要素的相关性演化[J]. 南水北调与水利科技(中英文), 2020, 18(6): 17-26. (WEI X W, ZHANG H B, XING C, et al. Correlation evolution analysis of hydro-meteorological elements of the basin under changing environment[J]. South-to-North Water Transfers and Water Science & Technology, 2020, 18(6): 17-26. (in Chinese)) DOI: 10.13476/j.cnki.nsbtdqk.2020.0112.
- [23] 侯夷, 任小平, 王长城, 等. 基于动态贝叶斯网络的空袭目标威胁评估[J]. 兵工自动化, 2019, 38(12): 77-81. (HOU Y, REN X P, WANG C C, et al. Threat assessment of air attack target based on dynamic Bayesian network [J]. Ordnance Industry Automation, 2019, 38(12): 77-81. (in Chinese)) DOI: 10.7690/bg-zdh.2019.12.018.
- [24] 姚宛艳, 吴迪. 近 50 a 来黄河流域温度和降水基本特征和变化趋势分析[J]. 中国农村水利水电, 2014(8): 104-109, 114. (YAO W Y, WU D. Characteristics and variation trends of temperature and precipitation in the Yellow River basin during 1961-2010 [J]. China Rural Water and Hydropower, 2014(8): 104-109, 114. (in Chinese)) DOI: 10.3969/j.issn.1007-2284.2014.08.028.
- [25] 鲍振鑫, 严小林, 王国庆, 等. 1956—2016 年黄河流域河川径流演变规律[J]. 水资源与水工程学报, 2019, 30(5): 52-57. (BAO Z X, YAN X L, WANG G Q, et al. The trend in streamflow of the Yellow River basin during 1956-2016 [J]. Journal of Water Resources and Water Engineering, 2019, 30(5): 52-57. (in Chinese)) DOI: 10.11705/j.issn.1672-643X.2019.05.09.
- [26] 王毅, 张楠, 陈建民, 等. 电力系统故障诊断的因果链分析法[J]. 电力科学与工程, 2013, 29(3): 23-28. (WANG Y, ZHANG N, CHEN J M, et al. Research on power system fault diagnosis based on cause-effect chain method [J]. Electric Power Science and Engineering, 2013, 29(3): 23-28. (in Chinese)) DOI: 10.3969/j.issn.1672-0792.2013.03.005.
- [27] 李晓宇, 李焯, 袁华, 等. 气候变化情景下黄河天然径流预测研究[J]. 人民黄河, 2012, 34(3): 27-29, 33. (LI X Y, LI Z, YUAN H, et al. Study on natural runoff forecasting of the Yellow River under future climate change scenarios [J]. Yellow River, 2012, 34(3): 27-29, 33. (in Chinese)) DOI: 10.3969/j.issn.1000-1379.2012.03.011.
- [28] 姚文艺, 冉大川, 陈江南. 黄河流域近期水沙变化及其趋势预测[J]. 水科学进展, 2013, 24(5): 607-616. (YAO W Y, RAN D C, CHEN J N. Recent changes in runoff and sediment regimes and future projections in the Yellow River basin [J]. Advances in Water Science, 2013, 24(5): 607-616. (in Chinese)) DOI: 10.14042/j.cnki.32.1309.2013.05.013.

## The Yellow River runoff forecast based on Bayesian network

ZHAO Feifei<sup>1,2</sup>, ZHANG Qingqing<sup>1,2</sup>, ZHANG Yu<sup>3</sup>, SHI Xufang<sup>1,2</sup>, ZHONG Deyu<sup>1,2,3</sup>

(1. School of Water Resources and Electric Power, Qinghai University, Xining 810016, China; 2. State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China; 3. State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China)

**Abstract:** With the development of social economy, the demand for water resources of the Yellow River basin is increasing. And most areas of the Yellow River basin are in arid and semi-arid areas and the ecological environment is fragile, aggravating the sensitivity of the water resources system to climate change. Although many researches have been conducted on the impact of climate factors on hydrological phenomena, the causality and the probability of interaction between climate factors and runoff are still very vague. Besides, in the study of hydrological forecast, the water balance model, the wavelet analysis, the neural network, and the fuzzy inference method can merely provide a deterministic forecast of hydrological process, while they can not quantitatively describe the uncertainty of forecast results. Based on the correlation and uncertainty of climate and hydrological systems, Bayesian network (BN) is then used to quantify the impact of climate factors on runoff and forecast the future runoff in the Yellow River basin.

Based on expert knowledge bases and other scholars' research results on the relationship between climate and runoff in the Yellow River basin, six climate factors including temperature, pressure, wind speed, specific humidity, evaporation, and precipitation were determined to form the variable node of BN with runoff, and the BN model of climate runoff was constructed by the Netica. The ChiMerge method was used to discretize the ERA-Interim reanalysis of climate and hydrological data from 1979 to 2018 into three sections. After the determination of network structure and training data set, the conditional probability table of each node can be obtained by the maximum likelihood estimation, and the Bayesian influence probability between variables can be calculated by the variable elimination method. In the BN model for predicting runoff, all the data in the prediction model is divided into twelve intervals to improve the prediction accuracy. The ERA reanalysis data in years of 1979–2018 is used as the training set. The history climatic data of CMIP5 ten climate models in years of 1979–2005 is used as a validation set. The reliability verification of the model is established by comparing the predicted range and trend of runoff in years of 1979–2005 with ERA runoff data over the same period. Finally, the future runoff data of the Yellow River basin is predicted by the climate variables of several typical concentration emission scenarios (RCP 2.6, RCP 4.5, and RCP 8.5) in years of 2006–2080.

From the climate runoff Bayesian network, it is found that any change in the state of any variable can cause other variables to change. For example, as the temperature state from low to high, the probability of high state evaporation and precipitation can increase, and the probability of high state runoff can also increase, indicating temperature, evaporation, and precipitation are positively correlated with runoff. The relationship among other climatic factors and runoff can also be found that specific humidity is positively correlated, pressure and wind speed are negatively correlated, and precipitation has the closest relationship with runoff. From years of 1979–2018, the natural runoff of the Yellow River showed a decreasing trend, and the runoff based on the probability prediction from BN also showed a decreasing trend. In the RCP2.6 scenario, the Yellow River basin's runoff in the next twenty and sixty years will be reached 58.55 billion m<sup>3</sup> and 58.857 billion m<sup>3</sup>. While in the RCP4.5 scenario, the Yellow River basin's runoff in the next twenty and sixty years will be reached 58.542 billion m<sup>3</sup> and 58.753 billion m<sup>3</sup>; and in the RCP8.5 scenario, the associated values will be reached 59.35 billion m<sup>3</sup> and 58.511 billion m<sup>3</sup>.

Climate change is of great significance to the change of surface runoff. The climate-runoff BN constructed in this paper is a network of uncertain relationships between climate and hydrological elements. It explores the climatic reasons for the reduction of runoff and conducts mid-and long-term predictions of future runoff. It is concluded that the main reason for the decrease in runoff is the decrease in precipitation, and the specific humidity is a key climatic element affecting precipitation. The future runoff forecast of the Yellow River in the next sixty years is estimated to be around 58.7 billion m<sup>3</sup> by BN.

**Key words:** Bayesian network; climate factor; runoff; uncertainty problem; Yellow River basin