DOI:10.13476/j.cnki.nsbdqk.2020.0113
闫宝伟,肖涵,霍磊,等.干旱致灾临界状态辨识方法[J].南水北调与水利科技(中英文),2020,18(6):27-41. YAN B W, XIAO H, HUO L, et al. Identification of the critical state of drought disaster[J]. South-to-North Water Transfers and Water Science & Technology, 2020, 18(6): 27-41. (in Chinese)

干旱致灾临界状态辨识方法

闫宝伟[1],肖涵[1,2],霍磊[1],杨文发[3],张俊[3],许银山[3]

(1.华中科技大学 水电与数字化工程学院,武汉 430074;2.宁波水利水电规划设计研究院有限公司,浙江 宁波 315192;3.长江水利委员会 水文局,武汉 430010)

摘要:由旱至灾的过程存在一个临界态,为辨识干旱致灾临界态,提出了一种基于主成分分析和支持向量机的干旱致灾临界态的辨识方法。通过主成分分析对反映干旱持续时间、严重程度和极值特征的多个指标进行降维处理,剔除指标间相关性引起的冗余信息,再进一步结合当地的历史旱灾记录,基于支持向量机的分类原理寻求最优分类平面,对潜在的干旱样本是否致灾进行分类,由此确定的分类平面就是干旱致灾临界面。该临界面可以直观地揭示干旱的发展过程及趋势,便于在干旱预警中推广应用。以汉江石泉以上流域为例,按照一定原则共选取 107 个潜在干旱样本,并结合历史旱灾记录,采用上述方法对潜在干旱样本中的致灾样本进行了辨识,率定期和验证期的准确率分别达到 88.6%和 78.6%,识别精度较高。

关键词:干旱致灾;临界状态;主成分分析;支持向量机;分类平面

中图分类号:TV121.7　　**文献标志码:**A　　**开放科学(资源服务)标志码(OSID):**

近年来,受全球气候变化影响,干旱在我国频繁发生,呈现加剧之势[1-3]。干旱通常是由大气环流或季风环流异常引起的[4],大气环流或季风环流异常导致降水持续偏少,气象干旱发生,致使地表、地下径流的补给减少,进而诱发水文干旱。气象干旱和水文干旱导致供水水源匮乏,社会经济水资源供需失衡,进一步引发社会经济干旱[5]。干旱的发生发展是一个由量变到质变渐进积累的过程。具象化来说,降水量持续减少,社会和生态环境缺水量逐渐增加,干旱现象初现但尚未致灾;随着旱情的进一步发展,水资源缺乏严重,农作物大面积减产乃至绝产,部分地区饮用水困难,生态环境恶化,进而形成灾害。由此,干旱不一定形成灾害,只有在干旱发展到一定程度时,灾情才出现,即由旱至灾的过程中存在一个临界状态,一旦越过此临界态,旱灾便形成[6]。实际上,旱灾的形成和发展过程包含复杂的动力学过程及多尺度的水分和能量循环机制,目前,虽有很多从大气环流角度分析干旱形成机理的研究成果[7-8],但对区域性重大干旱形成的动力学机制尚不明确,干旱致灾机理和过程特征也还缺乏深入认识[9],至今尚未形成一套完善的理论能够对干旱致灾的临界状态进行物理解析,这给干旱识别和预警带来了困难。

随着计算机技术和人工智能技术的不断发展,许多统计学方法被应用到干旱识别和预测预警领域。Paulo 等[10]提出了一种基于马尔科夫链和 SPI 指数的干旱预测模型,探究了干旱受降雨季节性等影响的可能演变方向,进一步提高了干旱预测精度。林洁等[11]利用马尔科夫链模型预测了湖北省未来干旱的发生概率,为早期干旱预警提供了依据。冯平等[12]采用人工神经网络(ANN)技术,提出了一个干旱程度评估模型,为干旱研究提供一种新的途径。Belayneh 等[13]对传统随机模型(ARIMA)和机器学习模型(ANN 和 SVR)进行了对比研究,结果

收稿日期:2019-11-17　**修回日期:**2020-05-16　**网络出版时间:**2020-05-28
网络出版地址:http://kns.cnki.net/kcms/detail/13.1430.TV.20200527.1747.003.html
基金项目:国家重点研发计划项目(2016YFC0402708);中央高校基本科研业务费资助(HUST:2017KFYXJJ195)
作者简介:闫宝伟(1981—),男,山东滨州人,副教授,主要从事水文基础理论研究。E-mail:bwyan@hust.edu.cn

表明后者对干旱预测的效果优于前者。周靖楠等[14]利用自适应差分进化算法改进极限学习机预测模型,选用海表异常温度作为该模型的输入因子,构建区域干旱预测模型,在一定程度上提高了预测精度与稳定性。这些模型并没有考虑大气、土壤和水文的潜在物理过程,更多地是关注数据本身内在的特性,试图通过一些统计类方法来寻找其中的特征和规则,有时也会得到满意的结果。这种无法用现有的科学理论做出解释,通过概括实验事实而得到规律的理论称为唯象论。唯象方法根据系统的宏观性质,不考虑系统的内部机制,直接利用系统的特点建立演化方程[15]。不同类型的干旱致灾过程不同,不同承载体的临界状态也不同,且旱灾引发的效应具有复杂多元性,涉及农业、生态和社会经济等多个领域,从物理机理的角度研究干旱致灾过程将变得异常复杂。为此,本文不去深究干旱的成因,而是试图从唯象论的视角宏观地去寻求干旱致灾的临界状态,提出一种基于主成分分析和支持向量机的干旱致灾临界状态辨识方法,为干旱预警和风险防范提供科学依据。

1　潜在干旱样本的选取

干旱是由多个气象及水文要素共同作用而产生的,在选取描述干旱特征属性的指标时需要综合考虑蒸发、降雨、径流、土壤含水量、空气湿度和温度等多个要素[16],常用的指标包括标准化降水指数(SPI)、标准化径流指数(SRI)、降水距平百分率、干燥程度、连续无雨天数、径流距平百分率和土壤相对湿度等[17]。在干旱的辨识和预警过程中,指标的选取会对结果产生较大影响,选取的指标个数不宜过多或过少。指标过多,可能会因指标间的相关性使得干旱某一方面的特征被过度强调,而忽视了其他方面的特征,以致不能全面衡量干旱的主要特征;指标过少,干旱的特征不能被全面地描述,可能导致漏判或误判。

一般而言,干旱事件可由干旱历时、干旱强度和干旱峰值等进行描述[17]。干旱历时是指水资源短缺状态的持续时间,可由连续无效降雨天数表征,同时应该将干旱初期前期降雨的影响纳入考虑因素,并将间隔时间较短的相邻干旱事件合并起来。干旱强度是水资源短缺严重程度的反映,其决定于降水、径流和蒸发的多少,可由降水及径流的距平百分率和日平均蒸发量等指标衡量。干旱峰值反映了缺水的极端情况,可以采用最小枯水流量表示。因此,选用干旱历时 D、降水距平百分率 P_a、径流距平百分率 Q_a、日平均蒸发量 $\bar{E}$ 和最小枯水流量 Q_{min} 等5 个指标作为干旱事件的特征指标[6]。将可能发展成干旱事件的样本定义为潜在干旱样本,并按照以下步骤确定潜在干旱样本及其特征指标。

(1)根据中国气象局的划分标准,当日降雨量 $P\leqslant10$ mm 时为小雨,一般而言,此等级的降雨量不能缓解区域干旱状态,因此将日降雨 $P\leqslant10$ mm 定义为无效降雨。当持续无效降雨天数超过 15 d,即干旱历时 $D\geqslant15$ d 时,该样本为潜在干旱样本。

(2)当两相邻样本间隔不超过 5 d 且日均降雨 $P\leqslant10$ mm 时,短暂有效降雨并不能有效缓解本次干旱,因此,将该两相邻样本合并为一次干旱事件,合并后,将第一个样本的开始时间作为本次干旱样本的开始时间,第二个样本的结束时间作为结束时间。

(3)由于前期降雨对径流和土壤含水量有一定的影响,可能会出现所选干旱样本初期并不“干旱”的情况,因此,可根据经验将该干旱样本的开始时间适当延后。

(4)根据所选干旱样本的起始时间和结束时间,确定样本的上述特征指标。其中

$$D=t_e-t_s+1 \tag{1}$$

式中:t_s、t_e 分别为样本的开始和结束时间。

$$P_a=\frac{\sum_{t=t_s}^{t_e}P_t-\sum_{t=t_s}^{t_e}\bar{P}_t}{\sum_{t=t_s}^{t_e}\bar{P}_t}\times100\% \tag{2}$$

式中:P_t 为第 t 天的日降雨量;$\bar{P}_t$ 为第 t 天的多年平均日降雨量。

$$Q_a=\frac{\sum_{t=t_s}^{t_e}Q_t-\sum_{t=t_s}^{t_e}\bar{Q}_t}{\sum_{t=t_s}^{t_e}\bar{Q}_t}\times100\% \tag{3}$$

式中:Q_t 为第 t 天的日径流;$\bar{Q}_t$ 为第 t 天的多年平均日径流。

$$\bar{E}=\frac{1}{D}\sum_{t=t_s}^{t_e}E_t \tag{4}$$

式中:E_t 为第 t 天的日蒸发量;$\bar{E}$ 为干旱历时内的日均蒸发量。

$$Q_{min}=\min\{Q_t, t=t_s\sim t_e\} \tag{5}$$

式中:Q_{min}为最小枯水流量。

2　干旱特征指标的主成分分析

上述干旱指标从不同侧面反映了干旱的基本特征,能够较全面地反映干旱的特征属性,但它们在某种程度上存在信息的重叠,具有一定的相关性[19],

为此，需去除因指标相关而产生的冗余信息，避免因选取的指标过多包含了某一方面的信息而赋予该特征过高的权重，导致部分信息失真而不能真实表征干旱的特征属性[20]。主成分分析可以在数据信息丢失最少的原则下，通过对高维的变量进行空间降维，将多个实测变量转换为少数几个不相关的综合变量（即主成分），从而去除信息冗余和噪声[21]。

根据主成分分析的原理，需要对以上选取的特征指标进行重新组合，令 $u_i(i=1,2,\cdots,5)$ 依次表示指标 D、P_a、Q_a、$\bar{E}$ 和 Q_{min}，则重构后的综合指标为

$$x_i=\alpha_{i1}u_1+\alpha_{i2}u_2+\alpha_{i3}u_3+\alpha_{i4}u_4+\alpha_{i5}u_5 \tag{6}$$

式中：$x_i(i=1,2,\cdots,5)$ 为重构后的综合指标，即原指标的第 i 个主成分；α_{ij} 为主成分系数。

重构后的综合指标之间互不相关，且方差依次递减，若前 $m(<5)$ 个主成分的方差累计贡献率达到 85%，可认为前 m 个主成分基本保留了原来信息，故取前 m 个主成分作为降维后的综合指标，由下式计算

$$x_l=\boldsymbol{\mu}_i^{\mathrm{T}}u'_i \tag{7}$$

式中：$x_l(l=1,2,\cdots,m)$ 为降维后的综合指标；$\boldsymbol{\mu}_i$ 为各特征值 λ_i 对应的特征向量；u'_i 为 u_i 标准化后的指标。

3　基于支持向量机的干旱致灾辨识方法

以唯象论的视角分析，设想干旱的发展是一个质点在空间的运动过程，质点在空间的位置取决于质点的坐标，当质点穿过某一临界面时，灾害形成。与此相对应，一个潜在干旱样本对应一个质点，样本的特征属性对应质点的坐标，由此，临界面将潜在干旱样本区分为致灾样本和未致灾样本，于是问题转化为如何通过观测样本寻求临界面。

基于统计学习理论的 VC 维理论和结构风险最小原理，Vapnik 等[22]提出了支持向量机模型，其根本思想是建立一个分类超平面作为决策曲面，使得对象正例和反例之间的隔离边缘最大化[23-24]。因此，可以借助支持向量机的最优分类平面实现干旱致灾临界面的辨识。根据年鉴记载的历史旱灾情况，将潜在干旱样本中有旱灾记录的标记为致灾，没有记录的意味着本次干旱没有造成灾害，将其标记为未致灾；令 $y=1$和 $y=-1$ 分别代表致灾和未致灾样本，并结合降维特征索引来构建数据集 $D=\{(x_i,y_i)\mid i=1,2,\cdots,n\}$，$x_i\in R^m$ 为输入，$y_i\in\{-1,1\}$为输出。把这 n 个致灾样本点作为 m 维空间中的点，如果这些点线性线性可分离的，则存在一个最优的分类超平面[21]

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b=0 \tag{8}$$

使得分类间隔（$2/\|\boldsymbol{w}\|$）最大。其中，$\boldsymbol{w}$ 为超平面的法向量，b 为位移项。由此，寻求最优超平面的问题就转化为如下优化问题[21]

$$\min_{w,b}\frac{1}{2}\|\boldsymbol{w}\|^2+C\sum_{i=1}^{n}\xi_i,\ s.t.\ y_i[\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b]\geqslant 1-\xi_i \quad (i=1,2,\cdots,n) \tag{9}$$

式中：参数 C 和 ξ_i 可采用粒子群算法进行优化而得。

对于凸二次优化问题，通过引入拉格朗日乘子，将目标函数和约束条件整合到拉格朗日函数中，这样能方便求解最值问题，并采用序列最小优化算法（SMO）进行求解，即可得到上述最优超平面方程。该超平面方程将潜在干旱样本分为致灾样本和未致灾样本，进一步借助该临界方程对干旱是否致灾进行判别，判别函数[21]为

$$f(\boldsymbol{x})=\mathrm{sgn}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b) \tag{10}$$

若 $f(\boldsymbol{x})=1$，表明致灾状态；若 $f(\boldsymbol{x})=0$，表明处于临界状态；若 $f(\boldsymbol{x})=-1$，表明未致灾状态。由此可以实现干旱的实时预警。

4　实例分析

选取汉江石泉以上流域为研究对象。该区域地处亚热带与暖温带的分界线，是典型的生态环境脆弱区和敏感区，流域面积 2.46 万 km^2，多年平均降水量达到 788 mm。受季风气候影响，区域内降水年内年际分布不均，旱灾频繁发生。中国气象数据网的资料显示，1992—2007 年，该流域发生中度干旱程度以上的年份有 8 年，其中最严重的干旱为 1994 年 8 月，农作物受灾百分比达到了 90%以上，对当地的生活、生产造成了很大影响。因此，该区域干旱致灾临界状态的辨识及预警具有重要研究价值和实际意义。

本文采用的干旱致灾临界状态辨识方法所需资料包括历史实测降水、径流、蒸发以及旱灾记录。其中，降水和蒸发资料来源于中国气象数据网，雨量站包括武功、太白、略阳、留坝、汉中、佛坪和石泉，蒸发站包括太白、留坝、汉中、佛坪和石泉。石泉站的径流资料由长江委水文局提供，旱灾记录通过搜集各地的年鉴获取，资料年限统一取为 1987—2015 年，面平均降水量和面平均蒸发量由泰森多边形加权平均得到。

根据 1987—2015 年的日降水、径流和蒸发资料，按照第 2 节介绍的选样方法，选取日降雨量≤10 mm且干旱历时 $D\geqslant 15$ d 的样本作为潜在干旱样本。若两相邻样本间隔少于 5 d 且间隔内日均降雨量≤10 mm，则合并该相邻样本为一个潜在干旱样本，这样共选取了 107 个潜在干旱样本。利用式(1)至(5)，分别得出每个干旱样本的 5 个特征指标值，

图 1 进一步给出了上述指标在不同年份间的分布情况,平均每年 3~4 个样本,其中 1994、2002 和 2014 年出现干旱次数较多,均有 6 个干旱样本。上述干旱指标存有一定的相关性,为消除它们之间的相关性,采用主成分分析对原特征指标进行降维,前 3 个主成分的方差累计贡献率接近 90%,因此,以这前 3 个主成分作为干旱样本的综合指标,分别由下式计算

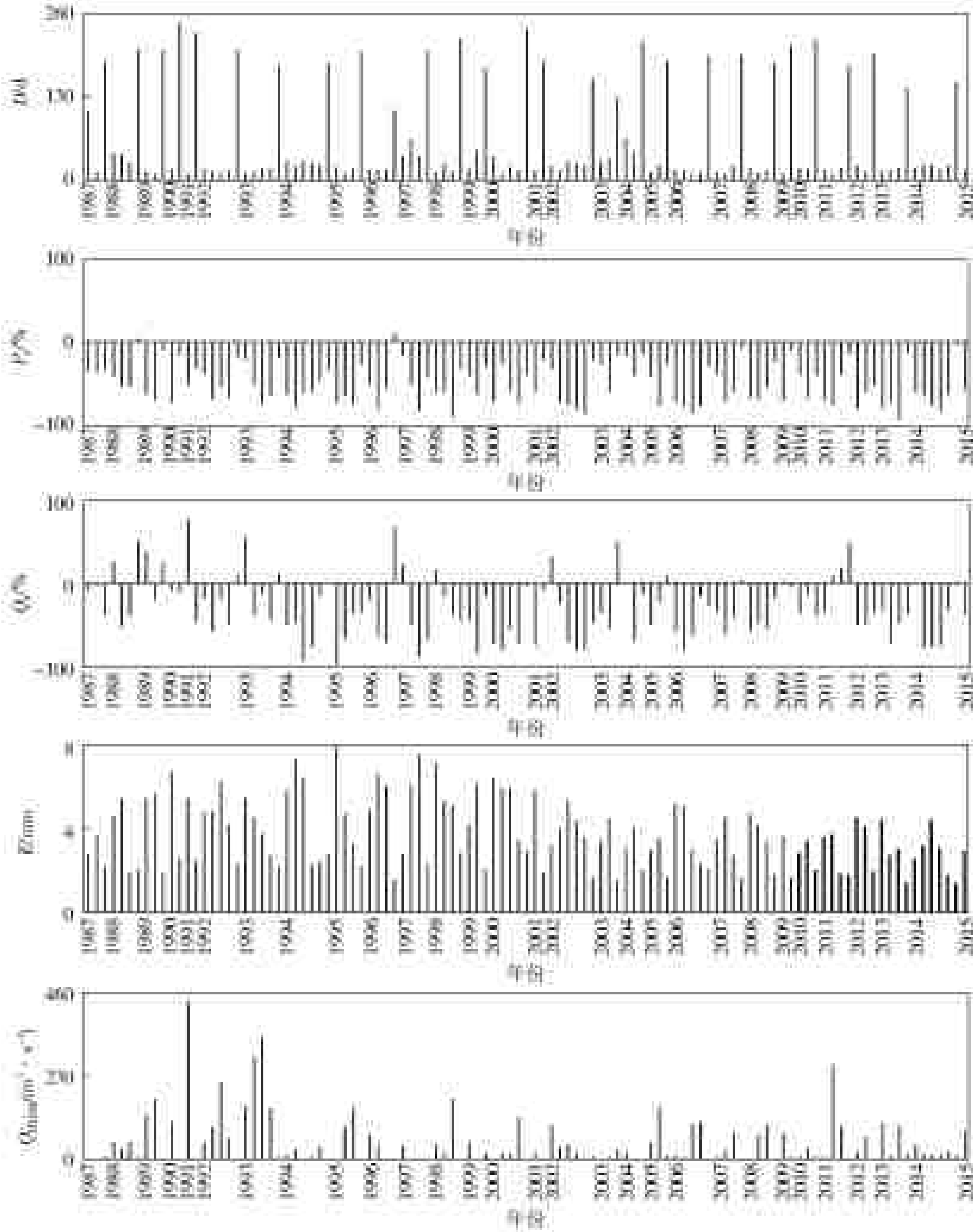

图 1　干旱样本的特征指标

$$x = 0.527\,6u'_1 + 0.557\,0u'_2 + 0.366\,3u'_3 - 0.476\,1u'_4 - 0.225\,0u'_5 \tag{11}$$

$$y = -0.183\,1u'_1 + 0.059\,6u'_2 + 0.623\,0u'_3 - 0.012\,0u'_4 + 0.758\,0u'_5 \tag{12}$$

$$z = 0.014\,6u'_1 + 0.336\,6u'_2 + 0.351\,1u'_3 + 0.821\,1u'_4 - 0.298\,5u'_5 \tag{13}$$

式中:x、y、z 分别代表第 1、第 2、第 3 主成分。

根据各地年鉴的旱灾记录情况,上述 107 个潜在样本中共有 58 个致灾样本,其中 1990 年 9 月 24 日至 1991 年 5 月 26 日为干旱历时最长的一次,持续时间 245 d。选取线性核函数,建立研究区域的二分类支持向量机模型,将 107 个潜在样本中的前 79 个用于分类模型的训练,将其分为致灾样本和未致灾样本,分类准确率为 88.6%,见图 2。进一步还可以得出最优分割平面方程为

$$0.664\,4x + 2.307\,7y + 0.154\,9z - 0.002\,3 = 0 \tag{14}$$

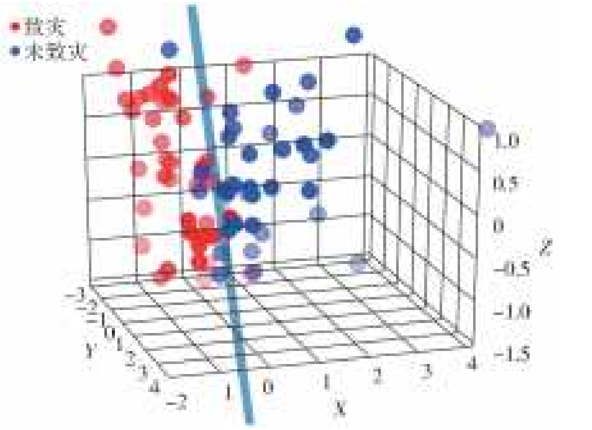

图 2　支持向量机模型训练样本的分类结果

该平面即为研究区域干旱致灾的临界平面，根据样本点到该平面的距离是否大于 0 即可实现样本的分类，将上述潜在样本剩余的 28 个用于模型的检验，分类准确率达到 78.6%，可以用于该区域干旱的实时预警。利用该临界平面方程可以进一步确定各个干旱特征指标的阈值组合，理论上，这样的组合会有无数种，这里仅给出部分组合情况。假定干旱历时和干旱历时内的日平均蒸发量取所有潜在干旱样本的平均值，当最小枯水流量阈值分别取 30、60、90、120 和 150 m^3/s 时，径流距平百分率和降水距平百分率的阈值组合见表 1。可见：在同一最小枯水流量阈值下，随着降水距平阈值的增加，径流距平的阈值在减小，如 $Q_{min}=30\ m^3/s$，$P_a=-20\%$ 和 -40% 时，Q_a 分别等于 -65% 和 -34%；在同一降水距平阈值下，随着最小枯水流量阈值的增加，径流距平的阈值也在增加，如 $P_a=-20\%$，$Q_{min}=30\ m^3/s$ 和 $60\ m^3/s$ 时，Q_a 分别等于 -65% 和 -88%。这些结果显然符合干旱致灾的发展规律。因此，可以利用这些干旱指标的阈值或直接利用致灾平面方程进行干旱的实时预警。

表 1　不同组合情形下径流距平百分率 Q_a 的临界阈值

项目		$Q_{min}/(m^3\cdot s^{-1})$				
		30	60	90	120	150
$P_a/\%$	−20	−65	−88	−111	−134	−157
	−40	−34	−57	−80	−103	−126
	−60	−2	−25	−48	−71	−94
	−80	29	6	−17	−40	−63
	−100	61	38	15	−8	−31

选取该区域 2015 年实际发生的一次旱灾作为例子，说明应用上述临界平面进行干旱预警的过程。按照上述选样方法进行干旱样本的选取，检测到 2015 年 7 月 5 日至 2015 年 8 月 3 日的一次潜在干旱事件，从 7 月 5 日起逐日计算 D、P_a、Q_a、$\bar{E}$ 和 Q_{min} 这 5 个特征指标，根据式(11)至(13)将其转化为综合指标。图 3 绘出了综合指标对应的逐日样本点以及临界平面，图中点的大小随着时间逐渐增大，据此可以直观地看到本次干旱的发展过程，即样本点从开始远离到逐渐靠近，并最终越过临界面转化为致灾状态的过程，与年鉴中记载“7 月至 8 月出现阶段性少雨伏旱天气，降水持续偏少，部分地区达到中度伏旱”一致，其中临界旱灾的发生时间为 2015 年 7 月 13 日。与此对应的各干旱指标的变化过程见图 4：7 月 13 日之前，降水和径流距平百分率变化都比较平稳，比历年同期分别减少大约 90% 和 40%，但持续时间较短，且最小枯水流量还比较大，尚不致灾；之后，无效降雨天数增加，蒸发呈现较大幅度增加，径流距平持续降低，降水距平百分率虽有增加之势，但仍然比历年同期减少 60%～90%，导致灾害形成。

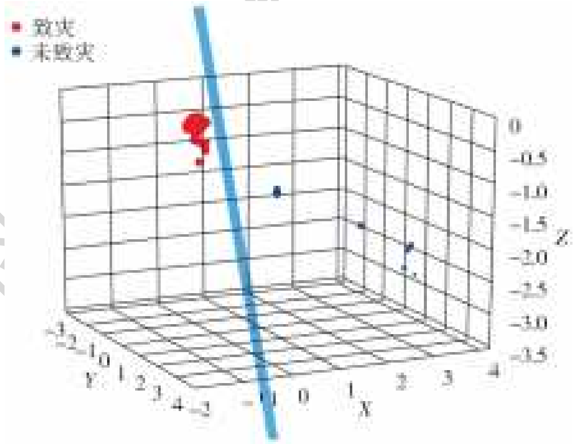

图 3　2015 年 7 月的一次干旱预警过程

5　结　论

干旱是由水文气象多种因素综合作用的结果，其形成的灾害也是多方面的，在造成社会经济损失的同时，也会对生态环境产生破坏。但由旱至灾的过程具有一定的隐蔽性，导致干旱致灾的临界状态不易辨识，很可能使决策者错过最佳救灾时机，以致造成更大的损失。本文提出基于主成分分析和支持向量机的干旱致灾临界状态辨识方法：一方面采用主成分分析对反映干旱特征属性的多重指标进行降维处理，去除因指标相关性而产生的冗余信息，生成的综合指标较全面的反映了干旱的基本特征，从而提高了干旱的识别精度；另一方面基于支持向量机的最优分割超平面提出了干旱致灾临界面的概念，该临界面具有较强的统计意义和几何含义，样本点到该临界面的距离反映了干旱的严重程度，从而实现了干旱致灾临界状态的辨识和预警。

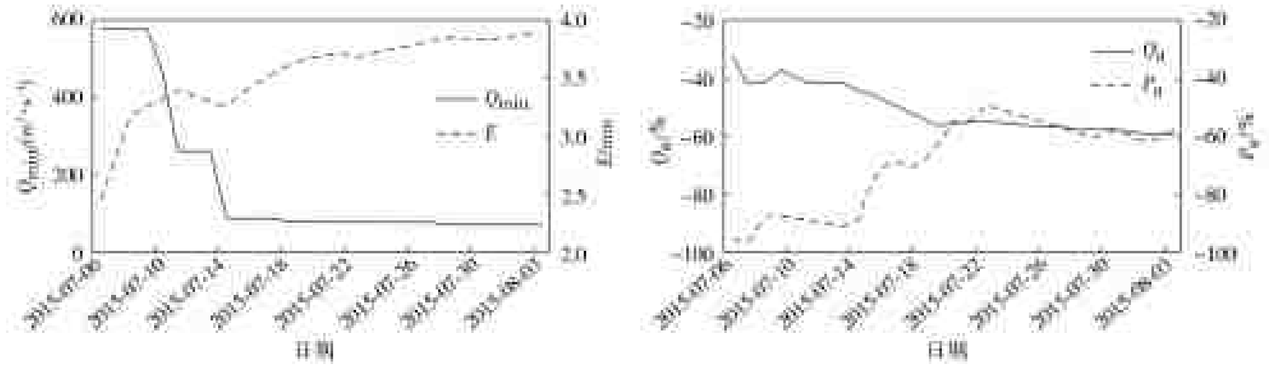

图 4　各指标变化过程

参考文献:

[1] 韩兰英,张强,姚玉璧,等. 近 60 年中国西南地区干旱灾害规律与成因[J]. 地理学报,2014,69(5):632-639. DOI:10. 11821/dlxb201405006.

[2] 金菊良,宋占智,崔毅,等. 旱灾风险评估与调控关键技术研究进展[J]. 水利学报,2016,47(3):398-412. DOI:10. 13243/j. cnki. slxb. 20150945.

[3] 曹博,张勃,马彬,等. 基于 SPEI 指数的长江中下游流域干旱时空特征分析[J]. 生态学报,2018,38(17):6258-6267. DOI:10. 5846/stxb201707011185.

[4] 屈艳萍,郦建强,吕娟,等. 旱灾风险定量评估总体框架及其关键技术[J]. 水科学进展,2014,25(2):297-304.

[5] 裴源生,蒋桂芹,翟家齐. 干旱演变驱动机制理论框架及其关键问题[J]. 水科学进展,2013,24(3):449-456.

[6] 肖涵. 干旱致灾临界状态辨识及汉江上游未来气候情景下干旱预测研究[D]. 武汉:华中科技大学,2019.

[7] 卫捷,张庆云,陶诗言. 1999 及 2000 年夏季华北严重干旱的物理成因分析[J]. 大气科学,2004,28(1):125-137. DOI:10. 3878/j. issn. 1006-9895. 2004. 01. 12.

[8] 陈权亮,华维,熊光明,等. 2008—2009 年冬季我国北方特大干旱成因分析[J]. 干旱区研究,2010,27(2):182-187.

[9] 李耀辉,周广胜,袁星,等. 干旱气象科学研究:“我国北方干旱致灾过程及机理”项目概述与主要进展[J]. 干旱气象,2017,35(2):165-174. DOI:10. 11755/j. issn. 1006-7639(2017)-02-0165.

[10] PAULO A A,PEREIRA L S. Prediction of SPI drought class transitions using Markov chains[J]. Water Resources Management,2007,21(10):1813-1827. DOI:10. 1007/s11269-006-9129-9.

[11] 林洁,夏军,佘敦先,等. 基于马尔科夫链模型的湖北省干旱短期预测[J]. 水电能源科学,2015,33(4):6-9,51.

[12] 冯平,钟翔,张波. 基于人工神经网络的干旱程度评估方法[J]. 系统工程理论与实践,2000,20(3):141-144. DOI:10. 3321/j. issn:1000-6788. 2000. 03. 026.

[13] BELAYNEH A,ADAMOWSKI J,et al. Long-term SPI drought forecasting in the Awash River basin in Ethiopia using wavelet neural network and wavelet support vector regression models[J]. Journal of Hydrology,2014,508(2):418-429. DOI:10. 1016/j. jhydrol. 2013. 10. 052.

[14] 周靖楠,刘振男. 基于自适应差分进化算法优化极限学习机的干旱预测方法[J]. 水电能源科学,2018,36(6):6-9. DOI:CNKI:SUN:SDNY. 0. 2018-06-002.

[15] 王泽农. 物理学中的唯象方法与唯象理论的认识论意义[J]. 现代物理知识,2008,20(3):52-54.

[16] 张迎,黄生志,黄强,等. 基于 Copula 函数的新型综合干旱指数构建与应用[J]. 水利学报,2018,49(6):703-714. DOI:10. 13243/j. cnki. slxb. 20171207.

[17] 潘妮,卫仁娟,詹存,等. 干旱指数在四川省的适用性分析[J]. 南水北调与水利科技,2017,15(4):71-78. DOI:10. 13476/j. cnki. nsbdqk. 2017. 04. 013.

[18] DAYAL K S,DEO R C,APAN A A. Investigating drought duration-severity-intensity characteristics using the Standardized Precipitation-Evapotranspiration Index:Case studies in drought-prone southeast Queensland[J]. Journal of Hydrologic Engineering,2018,23(1):05017029. DOI:10. 1061/(ASCE)HE. 1943-5584. 0001593.

[19] 常文娟,梁忠民,马海波. 基于主成分分析的干旱综合指标构建及其应用[J]. 水文,2017,37(1):33-38. DOI:10. 3969/j. issn. 1000-0852. 2017. 01. 006.

[20] 闫宝伟,潘增,薛野,等. 论水文计算中的相关性分析方法[J]. 水利学报,2017,48(9):1039-1046. DOI:10. 13243/j. cnki. slxb. 20161280.

[21] HOTELLING H. Analysis of a complex of statistical variables into principal components[J]. Journal of Educational Psychology,1933,24(6):417-520. DOI:10. 1037/h0071325.

[22] VAPNIK V. The Nature of statistical learning theory[M]. New York:Springer,1995. DOI:10. 1007/978-1-4757-2440-0.

[23] 张岩,杨明祥,雷晓辉,等. 基于 PCA-PSO-SVR 的丹江口水库年径流预报研究[J]. 南水北调与水利科技,2018,16(5):35-40. DOI:10. 13476/j. cnki. nsbdqk. 2018. 0070.

[24] ZUO W,WANG F,ZHANG D,et al. Distance metric learning via iterated support vector machines[J]. IEEE Transactions on Image Processing,2017,26(10):4937-4950. DOI:10. 1109/TIP. 2017. 2725578.

·译文(Translation)·

DOI:10.13476/j.cnki.nsbdqk.2020.0113

Identification of the critical state from drought to disaster

YAN Baowei[1], XIAO Han[1,2], HUO Lei[1], YANG Wenfa[3], ZHANG Jun[3], XU Yinshan[3]

(1. School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; 2. Ningbo Water Conservancy & Hydropower Planning and Design Institute Co., Ltd, Ningbo 315192, China; 3. Bureau of Hydrology, Changjiang Water Resources Commission, Wuhan 430010, China)

Abstract: There is a critical stage in the process from drought to disaster. In order to identify the critical state of drought disaster, a method based on principal component analysis and support vector machine is proposed. The principal component analysis method is used to reduce the dimensionality of several indexes that can reflect the duration, severity, and extremum of drought, and to remove the redundant information of the correlated multiple indicators. The support vector machine is used to find the optimal classification plane based on the historical drought records. The potential drought samples are classified into drought samples and disaster samples based on this classification plane. This classification plane can be phenomenologically defined as the critical state plane of drought-disaster transition. The development process and trend of drought can be revealed more intuitively, which is convenient to popularize and apply to assess drought warnings. The sub-basin above Shiquan in the Hanjiang River basin is selected as a case study, and 107 potential drought samples are selected according to certain principles. The above-proposed method is used to identify the disaster samples according to the historical drought records, with accuracies of 88.6% and 78.6% in calibration and validation periods, respectively.

Key words: drought to disaster; critical state; principal component analysis; support vector machine; classification plane

In recent years, affected by global climate change, droughts have frequently occurred in China and showed a trend of intensification[1-3]. Drought is usually caused by atmospheric circulation or monsoon circulation anomaly[4]. The atmospheric circulation or monsoon circulation anomalies can lead to persistent low precipitation and meteorological drought, which reduces the replenishment of surface and underground runoff and induces hydrological drought. Meteorological drought and hydrological drought lead to the lack of water supply and the imbalance between the supply and demand of social and economic water resources, which further cause socio-economic drought[5]. The occurrence and development of drought is a gradual accumulation process from quantitative change to qualitative change. To be more specific, the amount of precipitation continues to decrease; the water shortage of the social and ecological environment is gradually increasing; the drought has begun to appear but has not yet caused a disaster. However, with the further development of the drought, the shortage of

Received: 2019-11-17 **Revised**: 2020-05-16 **Online publishing**: 2020-05-28
Online publishing address: http://kns.cnki.net/kcms/detail/13.1430.TV.20200527.1747.003.html
Funds: National Key Research and Development Program of China (2016YFC0402708); Fundamental Research Funds for the Central Universities of Ministry of Education of China (HUST: 2017KFYXJJ195)
Author brief: Yan Baowei (1981-), male, associate professor, Binzhou, Shandong Province, mainly engaged in basic hydrology theory. E-mail: bwyan@hust.edu.cn

water resources is serious, the production of crops has been reduced or even stopped, drinking water is in short supply in some areas, and the ecological environment is deteriorating and in turn causes disasters. Therefore, drought will become disaster only when it develops to a certain degree, namely there is a critical state in the process from drought to disaster. Once this critical state is broken, the drought disaster will happen[6]. In fact, the formation and development process of drought includes complex dynamic processes and cycle mechanisms of multi-scale water and energy. At present, although many studies[7-8] analyze the mechanism of drought formation from the perspective of atmospheric circulation, the dynamic mechanism of the formation of major regional droughts is still unclear, and the mechanism and process characteristics from drought to disasters are still lacking better understanding[9]. So far, a complete theory has not been formed to physically analyze the critical state from drought to disasters, which brings difficulties to drought identification and warning.

With the continuous development of computer technology and artificial intelligence technology, many statistical methods have been applied to the field of drought identification, prediction, and warning. Paulo et al.[10] proposed a drought prediction model based on Markov chain and SPI. They explored the possible evolution direction of drought affected by rainfall seasonality and further improved the accuracy of drought prediction. Lin et al.[11] used Markov chain model to predict the probability of future drought in Hubei Province in China, which lays the groundwork for early drought warning. Feng et al.[12] used artificial neural network (ANN) technology to propose a drought assessment model, which provides a new way of drought research. Belayneh et al.[13] conducted a comparative study on traditional stochastic model, namely autoregressive integrated moving average (ARIMA) and machine learning model(ANN and support vector machine (SVR), and the results showed that the latter is better than the former for drought prediction. Zhou et al.[14] used adaptive differential evolution algorithm to improve the extreme learning machine prediction model, and selected abnormal sea surface temperature as the input factor of the model to construct a regional drought prediction model, improving the prediction accuracy and stability to a certain extent. These models do not consider the potential physical processes of the atmosphere, soil, and hydrology, but pay more attention to the inherent characteristics of the data itself, and try to find the characteristics and rules through some statistical methods, and sometimes, satisfactory results can be obtained. This kind of theory that cannot be explained by the existing scientific theories and obtains the law by summarizing experimental facts is called phenomenology. The phenomenological method, based on the macroscopic nature of the system, ignores the internal mechanism of the system, and directly uses the characteristics of the system to establish the evolution equation[15]. Different types of droughts have distinct disaster-causing processes, and the critical states of different carriers also vary. What's more, the effects caused by droughts are complex and diverse, involving many fields such as agriculture, ecology, and socio-economics. Studying the drought-causing process from the perspective of physical mechanisms will become extremely complicated. For this reason, this paper does not go into the causes of drought, but tries to find the critical state from drought to disaster from the perspective of phenomenology macroscopically, and proposes a method of identifying the critical state from drought to disaster based on principal component analysis and SVM, so as to provide scientific basis for drought warning and risk prevention.

1 Selection of potential drought samples

Drought is caused by the joint action of multiple meteorological and hydrological elements. When selecting indexes to describe the characteristics of drought, it is necessary to comprehensively consider evaporation, rainfall, runoff, soil water content, air humidity and temperature, and other factors[16]. The commonly used indexes include standardized precipitation index (SPI), standardized runoff index (SRI), anomaly percentage of precipitation degree of dryness, number of consecutive days without rain, anomaly percentage of runoff, and relative soil humidity[17]. In the process of drought identification and warning, the selection of indexes will have a great impact on the results, and the number of indexes selected should not be too many or too few. Too many indexes may cause the one aspect of characteristics of drought to be overemphasized due to the correlation between the indexes, with other aspects of the characteristics ignored, so that the main characteristics of drought cannot be fully measured. If there are too few indexes, drought characteristics will not be comprehensively described, which will lead to misjudgments.

In generally, drought events can be described by the duration intensity, and peak[17]. Drought duration refers to the duration of water shortage, which can be characterized by the number of consecutive invalid rainfall days. At the same time, the influence of early rainfall in the early stage of the drought should be taken into consideration, and adjacent drought events with short intervals should be combined. Drought intensity is a reflection of the severity of water shortage, which depends on the amount of precipitation, runoff and evaporation. It can be measured by the anomaly percentages of precipitation and runoffand daily average evaporation. The drought peak reflects the extreme condition of water shortage and can be expressed by the minimum low water flow. Therefore, five indexes including drought duration D, anomaly percentage of precipitation P_a, anomaly percentage of runoff Q_a, daily average evaporation $\bar{E}$, and minimum low water flow Q_{min} are selected as the characteristic indexes of drought events[6]. The samples that may develop into drought events are defined as potential drought samples, and the potential drought samples and their characteristic indexes are defined according to the following steps.

(1) According to the classification standard of the China Meteorological Administration, the daily rainfall $P \leqslant 10$ mm shows light rain. Generally speaking, this level of rainfall cannot alleviate the regional drought. Therefore, the daily rainfall $P \leqslant 10$ mm is defined as invalid rainfall. When the invalid rainfall lasts for 15 days continuously namely when the drought duration $D \geqslant 15$ days, the sample is a potential drought sample.

(2) When the interval between two adjacent samples is no more than 5 days with average daily rainfall $P \leqslant 10$ mm, the short-term effective rainfall cannot effectively alleviate the drought. Therefore, the two adjacent samples are combined into a drought event. After that, the start time of the first sample is taken as the start time of this drought sample, and the end time of the second sample is taken as the end time.

(3) Since the previous rainfall has a certain impact on runoff and soil water content, the selected drought samples may not in a "drought" state at the initial stage. Therefore, the start time of the drought samples can be appropriately delayed based on experience.

(4) According to the start time and end time of the selected drought samples, the above-mentioned characteristic indexes of the samples can be determined. Among them, there are

$$D = t_e - t_s + 1 \tag{1}$$

where: t_s and t_e are the start time and the end time of the samples, respectively.

$$P_a = \frac{\sum_{t=t_s}^{t_e} P_t - \sum_{t=t_s}^{t_e} \bar{P}_t}{\sum_{t=t_s}^{t_e} \bar{P}_t} \times 100\% \tag{2}$$

where: P_t is the daily rain fall on day t and $\bar{P}_t$ is the perennial average daily rainfall on day t.

$$Q_a = \frac{\sum_{t=t_s}^{t_e} Q_t - \sum_{t=t_s}^{t_e} \bar{Q}_t}{\sum_{t=t_s}^{t_e} \bar{Q}_t} \times 100\% \tag{3}$$

where: Q_t is the daily runoff on day t and $\bar{Q}_t$ is the perennial average daily runoff on day t.

$$\bar{E} = \frac{1}{D} \sum_{t=t_s}^{t_e} E_t \tag{4}$$

where: E_t is the daily evaporation and $\bar{E}$ is the daily average evaporation in the drought duration.

$$Q_{min} = \min\{Q_t, t = t_s \sim t_e\} \tag{5}$$

where: Q_{min} is the minimum low water flow.

2 Principal component analysis of drought characteristic indexes

The above drought indexes reflect the basic characteristics of drought from different sides, and can more comprehensively reflect the characteristic attributes of drought, but they overlap in information to a certain extent and have certain relevance[19]. For this reason, redundant information generated by correlation of indexes should be removed to avoid assigning excessive weight to the characteristics because the selected index contains too much information of a certain aspect, resulting in partial information distortion and failure to truly represent the characteristic attributes of drought[20]. Principal component analysis can convert multiple measured variables into a few unrelated comprehensive variables (namely principal components) through spatial reduction of high-dimensional variables under the principle of the least loss of data information, thereby removing information redundancy and noise[21].

According to principal component analysis, the above selected characteristic indexes need to be recombined, if u_i ($i=1,2,\cdots,5$) represents indexes D, P_a, Q_a, $\bar{E}$, and Q_{min} in turn, the reconstructed aggregative index can be obtained

$$x_i=\alpha_{i1}u_1+\alpha_{i2}u_2+\alpha_{i3}u_3+\alpha_{i4}u_4+\alpha_{i5}u_5 \tag{6}$$

where: x_i ($i=1,2,\cdots,5$) are the reconstructed aggregative indexes, namely the principal component i of the original indicator; α_{ij} is the principal component coefficient.

The reconstructed aggregative indexes are not correlated with each other, and the variances decrease sequentially. If the cumulative contribution rate of the variance of the first m (<5) principal components reaches 85%, it can be considered that the first m principal components basically retain the original information. Therefore, the first m principal components are taken as the aggregative indexes after dimensionality reduction, which is calculated by the following equation

$$x_l=\boldsymbol{\mu}_i^{\mathrm{T}}u'_i \tag{7}$$

where: x_l ($l=1,2,\cdots,m$) is the aggregative index after dimensionality reduction; $\boldsymbol{\mu}_i$ is the feature vector corresponding to each eigenvalue λ_i; u'_i is the index after u_i standardization.

3 Drought disaster identification method based on support vector machine

Analyzed from the perspective of phenomenology, it is assumed that the development of drought is a movement process of a particle in space. The position of a particle in space depends on the coordinates of the particle. When the particle crosses a certain critical state plane, disasters are formed. Correspondingly, a potential drought sample corresponds to a particle, and the characteristic attributes of the sample correspond to the coordinates of the particle. Therefore, the critical state plane divides the potential drought sample into a disaster-causing sample and a non-disaster-causing sample, so the question is transformed into how to find the critical state plane through observing the sample.

Based on the Vapnik-Chervonenkis Dimension theory of statistical learning theory and the principle of structural risk minimization, Vapnik et al. [22] proposed a SVM model. The basic idea is to establish a classification hyperplane as a decision surface to maximize isolated edge between the positive and negative examples of the object[23-24]. Therefore, the optimal classification hyperplane of SVM can be used to identify the critical state plane of drought disaster. According to the historical drought conditions recorded in the yearbook, the potential drought samples with drought records are marked as disaster-causing, and the absence of records means that the drought does not cause disasters, and they are marked as non-disaster-causing. We let $y=1$ and $y=-1$ represent the disaster-causing and non-disaster-causing samples, and combine the dimensionality-reduction characteristic index to construct a data set $D=\{(x_i,y_i)\mid I=1,2,\cdots,n\}$. Among them, $x_i\in R^m$ is the input and $y_i\in\{-1,1\}$ is the output. These n disaster-causing sample points are taken as points in the m-dimensional space. If these points are linearly separable, there is an optimal classification hyperplane[21]

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b=0 \tag{8}$$

The classification interval ($2/\|\boldsymbol{w}\|$) is set to the maximum. Among them, $\boldsymbol{W}$ is the normal vector of the hyperplane, and b is the displacement term. Therefore, the problem of seeking the optimal hyperplane is transformed into the following optimization problem[21]

$$\min_{w,b}\frac{1}{2}\parallel \boldsymbol{w}\parallel^2+C\sum_{i=1}^{n}\xi_i, s.t.\ y_i[\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b]\geqslant 1-\xi_i \quad (i=1,2,\cdots,n) \tag{9}$$

where: parameters C and ξ_i can be optimized by particle swarm algorithm.

For convex quadratic optimization problems, with Lagrangian multipliers introduced, the objective function and constraints are integrated into the Lagrangian function, which can facilitate the solution of the extreme value. Moreover, the above-mentioned optimal hyperplane equation can be obtained through the sequential minimum optimization (SMO) algorithm. The hyperplane equation divides the potential drought samples into disaster-causing samples and non-disaster-causing samples, and further uses the critical equation to determine whether the drought is disaster-causing, and the discriminant equation is[21]

$$f(\boldsymbol{x})=\mathrm{sgn}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b) \tag{10}$$

$f(\boldsymbol{x})=1$ indicates a disaster-causing state; $f(\boldsymbol{x})=0$ denotes a critical state; $f(\boldsymbol{x})=-1$ means a non-disaster-causing state. Thus, real-time early warning of drought can be realized.

4 Case study

Sub-basin above Shiquan in the Hanjiang River basin is selected as a case study. Located at the boundary between subtropical and warm temperate zones, this area is a typical ecological environmentally fragile and sensitive area. The basin area is 24,600 km^2, and the perennial average precipitation reaches 788 mm. Affected by the monsoon climate, the inter-annual distribution of precipitation in the region is uneven during the year, and droughts occur frequently. According to the data of the China Meteorological Data Network, from 1992 to 2007, there were eight years of moderate drought or more in the basin, and the most severe drought took place in August 1994. The percentage of crops affected by the disaster reached more than 90%, which had a great impact on local life and production. Therefore, the identification and warning of the critical state from drought to disaster in this region has important research value and practical significance.

The data required for the identification method of critical state from drought to disasters used in this paper includes historical measured precipitation, runoff, evaporation, and drought records. The precipitation and evaporation data come from the China Meteorological Data Network; the rainfall stations include Wugong County, Taibai County, Lueyang County, Liuba County, Hanzhong County, Fuoping County, and Shiquan County; the evaporation stations include Taibai County, Liuba County, Hanzhong County, Fuoping County, and Shiquan County. The runoff data of Shiquan Station is provided by the Hydrological Bureau of the Yangtze River Committee. The drought records are obtained with the yearbooks of various places collected. The data period is uniformly taken from 1987 to 2015. The planar average precipitation and average evaporation are obtained by the weighted average of Thiessen polygon.

According to the daily rainfall, runoff, and evaporation data from 1987 to 2015, according to the sampling method introduced in Section 2, samples with daily rainfall $\leqslant$10 mm and drought duration $D\geqslant$15 days are selected as potential drought samples. If the interval between two adjacent samples is less than five days and the average daily rainfall in the interval is less than or equal to 10 mm. The adjacent samples are merged into a potential drought sample, so a total of 107 potential drought samples are selected. With Eqs. (1)-(5), the five characteristic indexes of each drought sample are obtained. Fig. 1 further shows the distribution of the above indexes in different years. On average, there are three to four samples per year. There are more droughts in 1994, 2002 and 2014, with six drought samples. The above drought indexes have a certain correlation. Principal component analysis is used to perform the dimensionality reduction on the original characteristic indexes in order to eliminate the correlation between them. The cumulative contribution rate of the variance of the first three principal components is close to 90%. Therefore, the first three principal components are used as aggregative indexes of drought samples, and they are calculated by the following equations:

$$x=0.5276u'_1+0.5570u'_2+0.3663u'_3-0.4761u'_4-0.2250u'_5 \tag{11}$$

$$y=-0.1831u'_1+0.0596u'_2+0.6230u'_3-0.0120u'_4+0.7580u'_5 \tag{12}$$

$$z=0.0146u'_1+0.3366u'_2+0.3511u'_3+0.8211u'_4-0.2985u'_5 \tag{13}$$

where: x, y, and z represent the first, second, and third principal component, respectively.

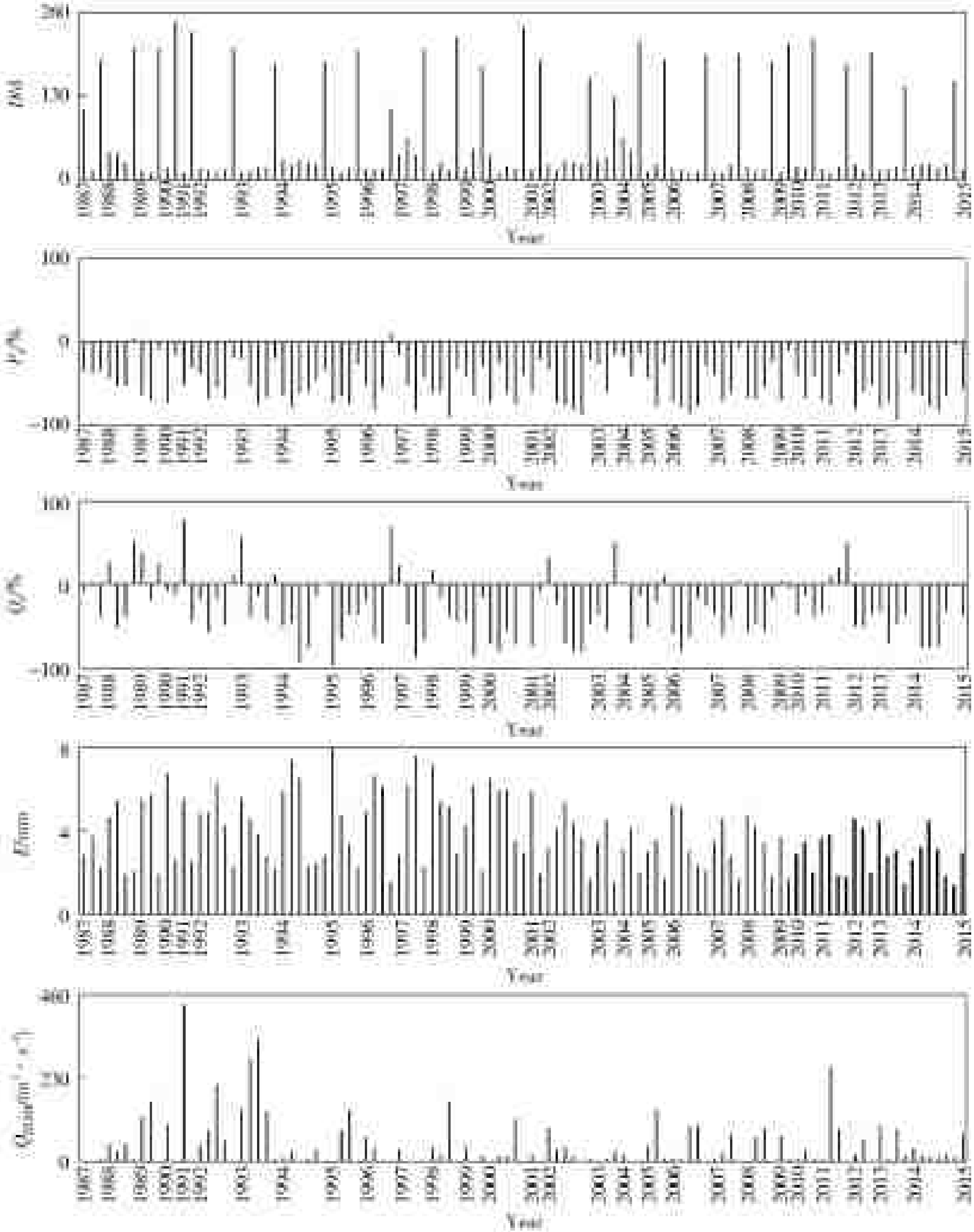

Fig. 1 The drought indicators of the samples

According to the drought records in local yearbooks, there are a total of 58 disaster-causing samples among the 107 potential samples mentioned above. Among them, the longest drought lasted 245 days from September 24, 1990 to May 26, 1991. The linear kernel function is selected to establish the two-class support vector machine model of the research area, and the first 79 of the 107 potential samples are used for the training of the classification model. The classification accuracy is 88.6% after they are divided into disaster-causing samples and non-disaster-causing samples as shown in Fig. 2. Furthermore, the optimal classification plane can be obtained as

$$0.6644x+2.3077y+0.1549z-0.0023=0 \quad (14)$$

This plane is the critical plane for drought disaster in the study area. The classification of samples can be realized according to whether the distance between the sample point and the plane is greater than 0. The remaining 28 potential samples above are used for model testing, and the classification accuracy reaches 78.6%, which can be used for real-time warning of drought in the region. The critical plane equation can be used to further determine the threshold combination of each drought characteristic index. In theory, there will be countless kinds of such combinations but they are not fully given here. Assuming that the drought duration and the daily average evaporation within the drought duration are taken as the average of all potential drought samples, when the minimum low water flow threshold is 30, 60, 90, 120, and 150 m^3/s, respectively, the threshold combinations of the anomaly percentages of runoff and precipitation are shown in Tab. 1. It can be seen that under the same minimum low water flow threshold, as the precipitation anomaly threshold increases, the runoff anomaly threshold decreases. For example, for $Q_{min}=30$ m^3/s, $P_a=-20\%$ and -40% lead to $Q_a=-65\%$ and -34%, respectively. Under the same precipitation anomaly threshold, with the growth of the minimum low water flow threshold, the runoff anomaly threshold also augments. For instance, for $P_a=-20\%$, $Q_{min}=30$ m^3/s and 60 m^3/s result in $Q_a=-65\%$ and -88%, respectively. These results are obviously in line with the development law of drought disasters. Therefore, the thresholds of these drought indexes or the disaster-causing plane equation can be used for real-time warning of drought.

A drought that actually occurred in the region in 2015 is selected as an example to illustrate the process of applying the above critical plane for drought warning. According to the above sample selection method, the drought samples are selected, and a potential drought event from July 5, 2015 to August 3, 2015 is detected. The five characteristic indexes D, P_a, Q_a, $\bar{E}$, and Q_{min} are calculated daily from July 5 and are transformed into aggregative indexes according to Eqs. (11)−(13). Fig. 3 presents the daily sample points and critical planes corresponding to the aggregative indexes. The size of the points in the figure gradually goes up with time. Based on this, the development process of this drought can be intuitively observed, namely the process of the sample points moving away from the beginning to gradually approaching and finally crossing the critical state plane and transforming into a disaster-causing state. It is consistent with the record in the yearbook of "phased low rainfall and summer droughts occurs from July to August; precipitation continues to be low; some areas are plagued by moderate summer droughts". Among them, the critical disaster happened on July 13, 2015. The corresponding changes in the drought indexes are shown in Fig. 4. Before July 13, the changes in the anomaly percentages of precipitation and runoff are relatively stable, with a decrease of about 90% and 40% respectively over the same period of the previous year, but the duration is relatively short, and the minimum low water flow is still relatively high, and it is not yet a disaster. After that, the number of invalid rainfall days rises, the evaporation surges, and the runoff anomaly continues to decrease. Although the anomaly percentage of precipitation has risen, it is still 60% - 90% less than those in the same period in the previous years, triggering disasters.

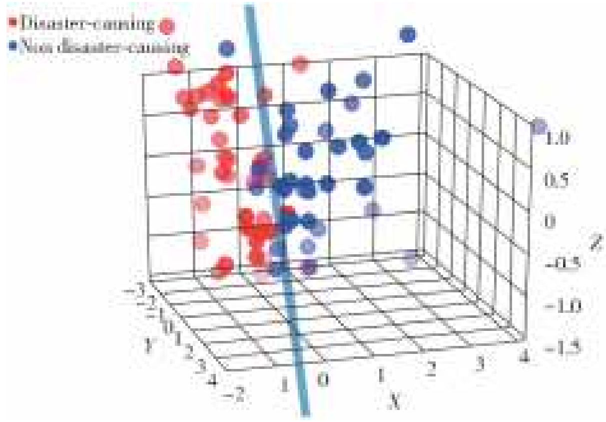

Fig. 2　Classification results of the training samples for the support vector machine model

Tab. 1　Critical thresholds of anomaly percentages of runoff Q_a in different combinations

Project		$Q_{min}/(m^3 \cdot s^{-1})$				
		30	60	90	120	150
$P_a/\%$	−20	−65	−88	−111	−134	−157
	−40	−34	−57	−80	−103	−126
	−60	−2	−25	−48	−71	−94
	−80	29	6	−17	−40	−63
	−100	61	38	15	−8	−31

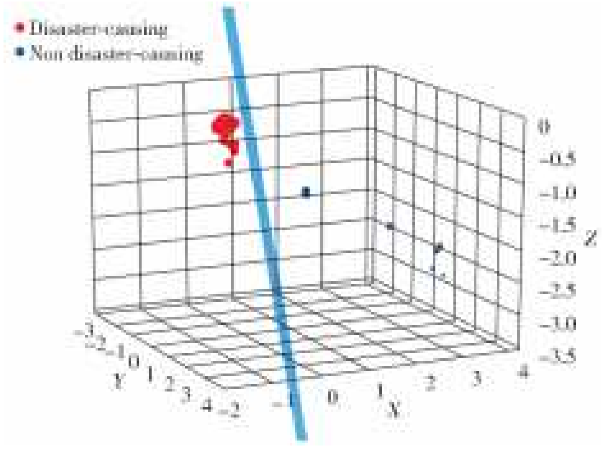

Fig. 3 A drought early warning process in July 2015

5 Conclusions

Drought is the result of the combined effects of many hydrological and meteorological factors, and the caused disasters are also harmful in many aspects such as society, economy, and ecological environment. However, the process from drought to disaster might not be seen, which makes it difficult to identify the critical state from drought to disaster. It is probably that decision makers may fail to seize best time for disaster relief and have an encounter with greater losses. This paper proposes a method for identifying critical states from drought to disaster based on principal component analysis and SVM. On the one hand, principal component analysis is used to reduce the dimensionality of several indexes that reflect the drought characteristics attributes, and to remove redundant information generated by the correlation of indexes. The generated aggregative index more comprehensively reflects the basic characteristics of drought, thereby improving the accuracy of drought identification. On the other hand, based on the optimal classification hyperplane of SVM, the concept of the critical state plane from drought to disaster is proposed, which has strong statistical significance and geometric meaning. The distance from the sample point to the critical state plane reflects the severity of drought, thereby realizing the identification and early warning of the critical state from drought to disaster.

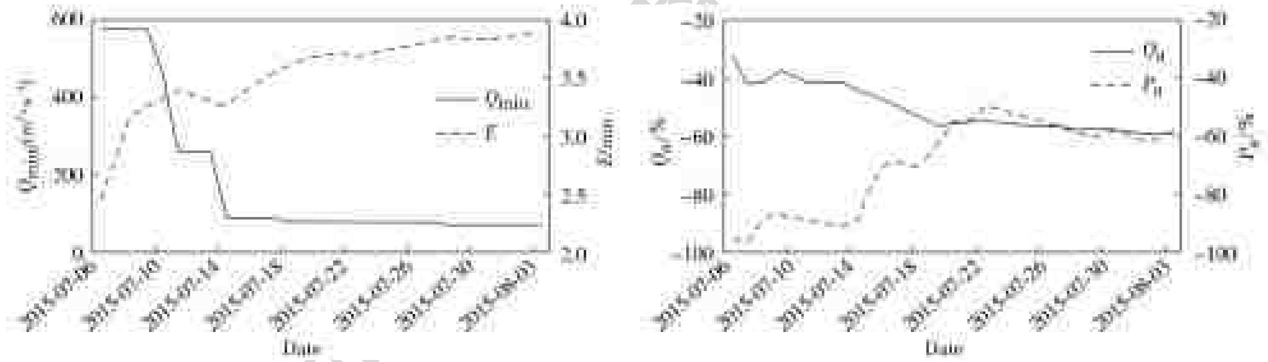

Fig. 4 Change process of each index

References:

[1] HAN L Y, ZHANG Q, YAO Y B, et al. Characteristics and origins of drought disasters in southwest China in nearly 60 years[J]. Acta Geographica Sinica, 2014, 69 (5): 632-639. (in Chinese) DOI: 10. 11821/dlxb201405006.

[2] JIN J L, SONG Z Z, CUI Y, et al. Research progress on the key technologies of drought risk assessment and control[J]. Journal of Hydraulic Engineering, 2016, 47 (3): 398-412. (in Chinese) DOI: 10. 13243/j. cnki. slxb. 20150945.

[3] CAO B, ZHANG B, MA B, et al. Spatial and temporal characteristics analysis of drought based on SPEI in the Middle and Lower Yangtze basin[J]. Acta Ecologica Sinica, 2018, 38 (17): 6258-6267. (in Chinese) DOI: 10. 5846/stxb201707011185.

[4] QU Y, LI J, LYU J, et al. A quantitative framework for assessing drought disaster risk and key techniques[J]. 2014, 25 (2): 297-304. (in Chinese)

[5] PEI Y, JIANG G, ZHAI J. Theoretical framework of drought evolution driving mechanism and the key problems[J]. 2013, 24 (3): 449-456. (in Chinese)

[6] XIAO H. Identification of critical state of drought disaster and drought prediction in future climate scenarios of the upper basin of the Hanjiang River[D]. Wuhan: Huazhong University of Science and Technology, 2019. (in Chinese)

[7] WEI J, ZHANG Q Y, TAO S Y. Physical causes of the 1999 and 2000 summer severe drought in north China

[J]. Chinese Journal of Atmospheric Sciences, 2004, 28 (1): 125-137. (in Chinese) DOI: 10.3878/j.issn.1006-9895.2004.01.12

[8] CHEN Q L, HUA W, XIONG G M, et al. Analysis on the causes of severe drought in north China in winter of 2008-2009[J]. Arid Zone Research, 2010, 27 (2): 182-187. (in Chinese)

[9] LI Y H, ZHOU G S, YUAN X, et al. Summary of the main progress of drought meteorology scientific research: The mechanism and disaster formation processes of drought in the north region of China[J]. Journal of Arid Meteorology, 2017, 35 (2): 165-174. (in Chinese) DOI: 10.11755/j.issn.1006-7639(2017)-02-0165.

[10] PAULO A A, PEREIRA L S. Prediction of SPI drought class transitions using Markov chains[J]. Water Resources Management, 2007, 21 (10): 1813-1827. DOI: 10.1007/s11269-006-9129-9.

[11] LIN J, XIA J, SHE D X, et al. Short-term drought prediction of Hubei Province based on Markov chain model[J]. Water Resources and Power, 2015, 33 (4): 6-9, 51. (in Chinese)

[12] FENG P, ZHONG X, ZHANG B. A drought degree evaluation method based on the artificial neural network[J]. Systems Engineering-Theory & Practice, 2000, 20 (3): 141-144. (in Chinese) DOI: 10.3321/j.issn: 1000-6788.2000.03.026.

[13] BELAYNEH A, ADAMOWSKI J, et al. Long-term SPI drought forecasting in the Awash River basin in Ethiopia using wavelet neural network and wavelet support vector regression models[J]. Journal of Hydrology, 2014, 508 (2): 418-429. DOI: 10.1016/j.jhydrol.2013.10.052.

[14] ZHOU J N, LIU Z N. Drought prediction method based on self-adaptive differential evolutionary algorithm and extreme learning machine[J]. Water Resources and Power, 2018, 36 (6): 6-9. (in Chinese) DOI: CNKI: SUN: SDNY.0.2018-06-002.

[15] WANG Z N. The phenomenological method in physics and the epistemological significance of phenomenological theory[J]. Modern Physics, 2008, 20 (3): 52-54. (in Chinese)

[16] ZHANG Y, HUANG S Z, LI P, et al. Construction and application of a new comprehensive drought index based on Copula function Yangtze River[J]. Journal of Hydraulic Engineering, 2018, 49 (6): 703-714. (in Chinese) DOI: 10.13243/j.cnki.slxb.20171207.

[17] PAN N, WEI R J, ZHAN C, et al. Study about application analysis of drought index in Sichuan Province [J]. South-to-North Water Transfers and Water Science & Technology, 2017, 15 (4): 71-78. (in Chinese) DOI: 10.13476/j.cnki.nsbdqk.2017.04.013.

[18] DAYAL K S, DEO R C, APAN A A. investigating drought duration-severity-intensity characteristics using the standardized precipitation-evapotranspiration index: Case studies in drought-prone southeast Queensland[J]. Journal of Hydrologic Engineering, 2018, 23 (1): 05017029. DOI: 10.1061/(ASCE)HE.1943-5584.0001593.

[19] CHANG W J, LIANG Z M, MA H B. Construction of drought composite indicator based on principal component analysis and its application[J]. Journal of China Hydrology, 2017, 37 (1): 33-38. (in Chinese) DOI: 10.3969/j.issn.1000-0852.2017.01.006.

[20] YAN B W, PAN Z, XUE Y, et al. Modeling dependence and correlation in hydrological calculation[J]. Journal of Hydraulic Engineering, 2017, 48 (9): 1039-1046. (in Chinese) DOI: 10.13243/j.cnki.slxb.20161280.

[21] HOTELLING H. Analysis of a complex of statistical variables into principal components[J]. Journal of Educational Psychology, 1933, 24 (6): 417-520. DOI: 10.1037/h0071325.

[22] VAPNIK V. The nature of statistical learning theory [M]. New York: Springer, 1995. DOI: 10.1007/978-1-4757-2440-0.

[23] ZHANG Y, YANG M X, LEI X H, et al. Research on annual runoff forecast of Danjiangkou reservoir based on PCA-PSO-SVR[J]. South-to-North Water Transfers and Water Science & Technology, 2018, 16 (5): 35-40. (in Chinese) DOI: 10.13476/j.cnki.nsbdqk.2018.0070.

[24] ZUO W, WANG F, ZHANG D, et al. Distance metric learning via iterated support vector machines [J]. IEEE Transactions on Image Processing, 2017, 26 (10): 4937-4950. DOI: 10.1109/TIP.2017.2725578.