DOI:10.3969/j.issn.1001-5485.2013.11.004 2013,30(11):16-19

基于信息熵的河南省年降水量时空变化研究

王振亚[1,4],吴德波[2],朱余生[3]

(1.河南省气象台,郑州 450003;2.黄河水利委员会 水文水资源局,郑州 450003;3.机械工业第六设计研究院有限公司,郑州 450003;4.中国气象局 河南省农业气象保障与应用技术重点实验室,郑州 450003)

摘要:选取河南省1960—2006年共83个站年降水量,将信息熵理论应用在河南省年降水量时空变化研究上,计算每站的年降水量时间序列和年降水量与其极小值的相对量序列的熵估计值、全省各年年降水量分布和年降水量与其多年平均降水量的相对值分布的熵估计值。计算结果表明:河南省年降水量的年际变化自西北东南方向增大,自南向北减小,自西向东增大,在西南东北方向则没有明显的规律;河南省年降水量相对值的年际变化自西南东北方向增大,在其他方向上则没有明显的规律;1966,1976,1999和2001年的河南省年降水量的空间分布较均匀,2000年的年降水量的空间分布不均匀;1986年的年降水量与其多年平均的相对值空间分布较均匀,2000年的年降水量与其多年平均的相对值空间分布不均匀。

关 键 词:信息熵;年降水量;时空变化;不确定性

中图分类号:TV125 **文献标志码**:A **文章编号**:1001-5485(2013)11-0016-04

降水量始终是气候变化研究的一个重要方面[1]。降水量时空分布的不确定性易形成旱涝灾害,对水资源、工业、农业和生态环境等方面产生重大影响。C. E. Shannon把Boltzmann的概念引入信息论中,把熵作为一个随机事件的不确定性或信息量的度量[2]。这种不确定性应用在年降水量的时间序列上可以研究降水量的年际变化规律,应用在年降水量的空间分布上可以研究年降水在空间上分布是否均匀。张继国[3]将信息熵的理论与方法应用在淮河流域降雨的时空分布研究上,取得较好的研究结果。

河南省地处北亚热带向暖温带过渡的大陆性季风气候区内[4],境内地理条件复杂,降水量的水平及垂直差异显著。一些学者基于统计学方法对河南省降水量的时空变化进行了相关研究[5-6],其主要基于定性描述,笔者将信息熵理论应用在河南省年降水量时空变化研究方面,该方法将时空变化规律转化为对信息熵的定量计算,理论依据明确,计算方法简单。

1 资料和方法

1.1 研究资料

选取河南省1960—2006年共83个站日降水量,逐站求取各站年降水量时间序列。

1.2 研究理论和方法

物理学中,熵是热力学系统的某种状态函数,它是对系统混乱度的度量,熵增使系统的混乱度增加,熵减使系统的有序性增加[7]。随着信息论的发展与应用,信息熵问题得到了不少学者的积极研究,取得了大量的成果,其具有代表性的模型有Renyi的信息熵、模糊集合的信息熵、高艾期的有效信息熵。

1948年,C. E. Shannon提出了基于离散型变量的信息熵。1994年,N. Ebrahimi等[8]提出了熵估计的2个改进式,可以由观测数据直接得到X的熵估计,并且证明了这2个估计式分别以概率收敛到$H(X)$。蒙特卡洛模拟显示这2个改进的估计式有较小的偏差和平均平方误差[3]。本研究采用其中的一个改进式作为计算熵的估计值。

设$x_1,x_2,\cdots,x_n$是来自总体X的一组观测值,则所求熵的估计值为

$$H_c(m,n)=\frac{1}{n}\sum_{i=1}^{n}\ln\frac{y_{i+m}-y_{i-m}}{c_i m/n}\quad 。\qquad (1)$$

其中

$$c_i=\begin{cases}1+\dfrac{i-1}{m}, & 1\leqslant i\leqslant m;\\ 2, & m+1\leqslant i\leqslant n-m;\\ 1+\dfrac{n-i}{m}, & n-m+1\leqslant i\leqslant n。\end{cases}$$

收稿日期:2012-09-18;修回日期:2012-10-22

作者简介:王振亚(1981-),男,河南正阳人,工程师,硕士研究生,主要从事水文气象方面的研究,(电话)18638398338(电子信箱)hhuwzy1981@163.com。

且 $\begin{cases} y_{i-m} = y_1, & i \leqslant m \\ y_{i+m} = y_n, & i \geqslant n-m \end{cases}$。

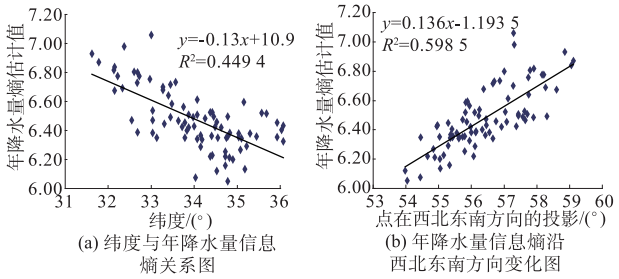
式中：$y_1 \leqslant y_2 \leqslant \cdots \leqslant y_n$ 是 $x_1, x_2, \cdots, x_n$ 按照从小到大排列后的顺序值；m 为正整数且 $0 < m \leqslant n/2$。

2　应用结果

2.1　年降水量空间分布研究

采用1.1节中的年降水量序列，将单站的年降水量时间序列看成总体 X 的一组观测值，求得该站的熵估计值，以此类推，分别求得83个站的熵估计值，计算过程中发现，修武、武陟、三门峡等12站年降水量有相等的情况出现，取 $m=2$；83个站经纬度坐标在一个直角坐标系中，分别求得它们在与 X 轴（0°纬线）成45°（即西南东北方向）和 −45°（即西北东南方向）夹角的投影坐标；通过将各站的熵估计值与所处的经度、纬度、西南东北方向和西北东南方向投影坐标分别作线性相关性分析，了解年降水量在时间上的不确定性的空间结构，希望找出其分布规律。

年降水量的熵估计值与西北东南方向投影坐标成线性正相关，相关系数为0.77，年降水量的熵估计值与纬度成线性负相关，相关系数为 −0.67（见图1），年降水量的熵估计值与经度成弱线性正相关，相关系数为0.45，年降水量的熵估计值与西南东北方向投影坐标无相关关系。换句话讲，年降水量在时间上的不确定性自西北东南方向增加，自南向北减弱，自西向东增加，在西南东北方向则没有明显的规律。从图2可看出，年降水量的熵估计值在东南部有个极大值区，西部有个极小值区，说明年降水量时间上的不确定性在东南部达到了极强，在西部达到了极弱。



(a) 纬度与年降水量信息熵关系图　(b) 年降水量信息熵沿西北东南方向变化图

图1　年降水量熵估计与纬度和西北东南方向投影坐标线性相关分析

Fig.1　The linear correlation between information entropy of annual precipitation respectively with latitude, and projection coordinate in northwest-southeast direction

2.2　年降水量相对值空间分布研究

单站的熵估计值可以说明这个站点的降水不确定性程度，但也存在着2个站点降水不确定性具有相同的程度，降水量级差异性显著这种可能。比如

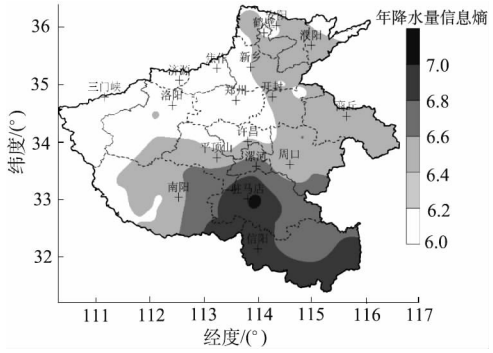


图2　年降水量信息熵分布图

Fig.2　Distribution of the information entropy of annual precipitation

存在着下面2组数据：100，200，300，400和2 100，2 200，2 300，2 400，它们的熵估计值都为0.43。如果将每组数据都除以它们的极小值，得到：1，2，3，4和1，1.047 619，1.10和1.142 857，得到的熵估计值分别为0.21和0.06。气候条件的不同引起降水量级的差异，同样的增幅对湿润和干旱地区的影响差异很大，从这个角度讲对数据适当处理后进行研究具有一定的意义。本研究将单站每年的年降水量除以单站年降水量时间序列的极小值得到新的时间序列，其它计算过程同2.1节。

年降水量相对值的熵估计值与西南东北方向投影坐标成线性弱正相关，相关系数为0.49（图3），年降水量的熵估计值与纬度、经度、西北东南方向投影坐标没有相关关系。换句话讲，年降水量相对值在时间上的不确定性自西南东北方向增加，在自南向北、自西向东和西北东南方向上则没有明显的规律。从图4可以看出：年降水量相对值的熵估计值在驻马店附近和河南的东北部有2个极大值区，西部有个极小值区，说明年降水量相对值时间上的不确定性在驻马店附近和河南的东北部达到了极强，在西部达到了极弱。

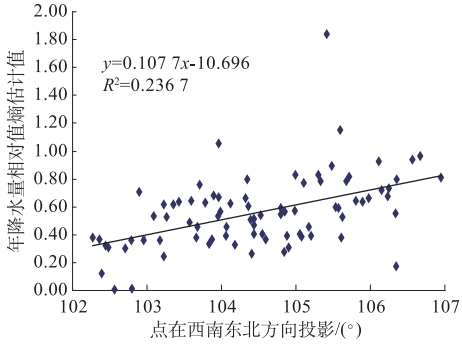


图3　年降水量相对值熵估计与西南东北方向投影坐标线性相关分析

Fig.3　Linear correlation between the information entropy of annual precipitation's relative value and the projection coordinate in southwest-northeast direction

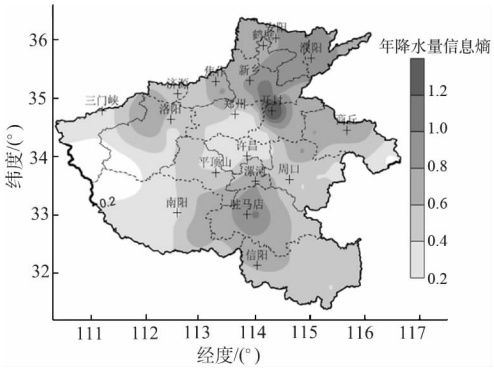


图 4 年降水量相对值信息熵分布图

Fig. 4 Distribution of the information entropy of annual precipitation's relative value

2.3 年降水量时间变化研究

采用节 1.1 中的年降水量序列,将同一年的全省 83 个站的年降水量看成总体 X 的一组观测值,求得熵估计值,分别求得 47 a 的熵估计值;将每站的年降水量除以该站 47 a 的年平均降水量再乘以 100 得到新的数据时间序列,分别求得其 47 a 的熵估计值,了解年降水量在空间上的不确定性随时间变化情况。

从图 5(a)可以看出:年降水量绝对值熵估计值时间变化曲线没有明显的变化规律,1966,1976,1999,2001 年熵估计值较小,2000 年熵估计值较大,则 1966,1976,1999,2001 年的年降水量的空间分布不确定性较小,2000 年的年降水量的空间分布不确定性较大;从图 5(b)可以看出:年降水量相对值熵估计值时间变化曲线没有明显的变化规律,1986 年熵估计值较小,2000 年熵估计值较大,则 1986 年的年降水量相对值的空间分布不确定性较小,2000 年的年降水量相对值的空间分布不确定性较大。

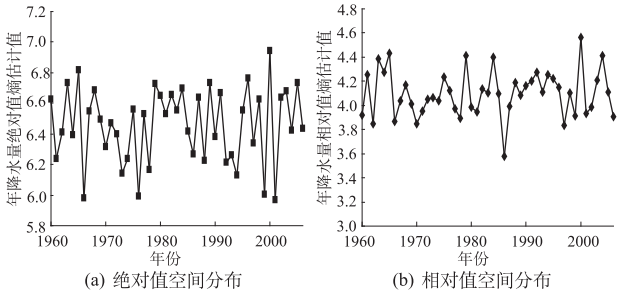


图 5 年降水量空间分布的熵估计值的时间变化曲线

Fig. 5 Time-history curves of the information entropy of annual precipitation's spatial distribution

3 结 论

(1) 各站年降水量的年际变化自西北东南方向增大,自南向北减小,自西向东增大,在西南东北方向则没有明显的规律;各站年降水量相对其极小值相对量的年际变化自西南东北方向增大,在自南向北、自西向东和西北东南方向上则没有明显的规律。

(2) 各站点年降水量的年际变化在河南省的东南部达到了极大,在西部达到了极小;各站年降水量相对其极小值相对量的年际变化在驻马店附近和河南的东北部达到了极大,在西部达到了极小。

(3) 结论(1)和结论(2)两者的变化规律呈现出一定的反向性,各站点年降水量的年际变化自南向北、自西向东和西北东南方向上规律明显,西南东北方向无规律,而各站年降水量相对其极小值相对量的年际变化在西南东北方向上有规律,在自南向北、自西向东和西北东南方向上无规律;对比图 2 和图 4 可以看出两者在驻马店附近都有极大值区,在三门峡地区的值较小,最明显的不同之处在大值区分布,各站点年降水量的年际变化在河南的东南部,而各站年降水量相对其极小值相对量的年际变化则在河南的东北部。

(4) 1966,1976,1999 和 2001 年的年降水量的空间分布较均匀,2000 年的年降水量的空间分布较不均匀;1986 年的年降水量相较多年平均的相对值空间分布较均匀,2000 年的年降水量相较多年平均的相对值空间分布较不均匀。

(5) 信息熵在河南省年降水量的时空分布研究上具有一定意义,该方法将河南省年降水量时空分布变化定量化,理论依据明确,计算方法简便。

参考文献:

[1] 左洪超,吕世华,胡隐樵. 中国近 50 年气温及降水量的变化趋势分析[J]. 高原气象,2004,23(2):238 - 244. (ZUO Hong-chao, LV Shi-hua, HU Yin-qiao. Variation Trend of Yearly Mean Air Temperature and Precipitation in China in the Last 50 Years[J]. Plateau Meteorology, 2004,23(2):238 - 244. (in Chinese))

[2] 王 彬. 熵与信息[M]. 西安:西北工业大学出版社,1994. (WANG Bin. Entropy and Information[M]. Xi'an: Northwest Industrial University Press,1994. (in Chinese))

[3] 张继国. 降水时空分布的信息熵研究[D]. 南京:河海大学,2004. (ZHANG Ji-guo. Information Entropy of Precipitation Distribution in Time and Space[D]. Nanjing: Hohai University, 2004. (in Chinese))

[4] 王纪军,裴铁璠,苏爱芳,等. 河南省降水集中程度研究[J]. 人民黄河,2010,32(10):84 - 86. (WANG Ji-jun, PEI Tie-fan, SU Ai-fang, *et al.* Research of Precipitation Concentration Degree in Henan[J]. Yellow River, 2010,32

(10): 84 - 86. (in Chinese))

[5] 张红卫,陈怀亮. 基于地理信息的河南省 3 ~ 4 月降水时空分布及变化[J]. 安徽农业科学,2009,37(24):11643 - 11644. (ZHANG Hong-wei, CHEN Huai-liang. Temporal-Spatial Distribution Analysis of Precipitation in March and April in Henan Area Supported by Geographic Information System[J]. Journal of Anhui Agricultural Sciences,2009,37(24):11643 - 11644. (in Chinese))

[6] 张红卫,陈怀亮,张　宏. 河南省春节降水与温度变化的时空分布 - 基于地理信息系统的分析[J]. 自然灾害学报,2012,21(1):170 - 173. (ZHANG Hong-wei, CHEN Huai-liang, ZHANG Hong. Spatiotemporal Distribution of Precipitation and Temperature Change in Henan Province in Spring: An Analysis Based on GIS[J]. Journal of Natural Disasters, 2012, 21(1): 170 - 173. (in Chinese))

[7] 夏　军. 灰色系统水文学:理论、方法及应用[M]. 武汉:华中理工大学出版社, 2000. (XIA Jun. Grey System Hydrology: Theory, Methods and Applications[M]. Wuhan: Huazhong University of Science and Technology Press, 2000. (in Chinese))

[8] EBRAHIMI N, PFLUGHOEFT K, SOOFI E S. Two Measures of Sample Entropy[J]. Statistics & Probability Letters,1994,20(3):225 - 234.

(编辑:赵卫兵)

Research on Spatio-Temporal Variation of Annual Precipitation in Henan Province by Information Entropy

WANG Zhen-ya[1,4], WU De-bo[2], ZHU Yu-sheng[3]

(1. Henan Meteorological Observatory, Zhengzhou 450003, China; 2. Hydrology Bureau of Yellow River Conservancy Commission, Zhengzhou 450003, China; 3. SIPPR Engineering Group Co., Ltd., Zhengzhou 450003, China; 4. Henan Provincial Key Laboratory of Agrometeorological Ensuring and Applied Technique under China Meteorological Administration, Zhengzhou 450003, China)

Abstract: The temporal and spatial variation of annual precipitation in Henan province was researched based on the theory of information entropy. The annual precipitation time series at 83 stations from 1960 to 2006 was selected. The estimated information entropy of the annual precipitation time series and the relative value of annual precipitation to its minimum at each station, as well as the estimated information entropy of the annual precipitation's spatial distribution and the relative value of annual precipitation to its multi-year average in each year, were calculated. Results show that the interannual variation of annual precipitation increases from northwest to southeast and from west to east, and reduces from south to north, but has no obvious regularity from southwest to northeast; while the interannual variation of the relative value of annual precipitation to its minimum increases from southwest to northeast but has no obvious regularity in the other directions. The spatial distribution of annual precipitation was uniform in 1966, 1976, 1999 and 2001 and nonuniform in 2000; the spatial distribution of the relative value of annual precipitation to its multi-year average is uniform in 1986 and nonuniform in 2000.

Key words: information entropy; annual precipitation; temporal and spatial variation; uncertainty

(上接第 15 页)

tree-ring index sequence. The results show that there are three positive phase stages and four negative phase stages in the low-frequency information of reconstruction sequence. Except for transitory positive phase in mid-1970s and around 1990s, the overall trend of I_{APO} is negative phase from 1914 to 1948, positive phase from 1948 to 1965, and generally negative phase since the mid-1960s. Comparative analysis shows that the reconstructed summer I_{APO} sequence is well synchronic with another summer I_{APO} reconstruction sequence.

Key words: summer Asian-Pacific Oscillation index (I_{APO}); tree-ring width of *Pinus tabulaeformis*; Xiaowutai Mountain