

基于条件植被温度指数的干旱预测研究

韩 萍¹ 王鹏新² 张树誉³ 朱德海²

(1 中国农业大学理学院,北京市海淀区圆明园西路 2 号,100193)

(2 中国农业大学信息与电气工程学院,北京市海淀区圆明园西路 2 号,100094)

(3 陕西省农业遥感信息中心,西安市北关正街 36 号,710015)

摘 要:基于遥感量化干旱监测结果,进行了干旱预测的研究。将遥感获得的条件植被温度指数 VTCI 序列应用于陕西关中平原地区,并利用 ARIMA 模型对该地区的 VTCI 时间序列进行分析建模预测。提出由点到面的时空序列预测方法,先对该区域的 36 个气象站所在像素点建立适合的 ARIMA 模型,再对整个区域所有像素点的 VTCI 时间序列进行建模预测。进行 1 步和 2 步预测,显示预测结果较好,1 步预测精度好于 2 步预测;对历史数据进行 AR(1)模型的拟合,拟合误差大部分较小。结果显示 AR(1)模型适合 VTCI 序列。

关键词:VTCI 序列;ARIMA 模型;干旱预测;预测误差;拟合误差

中图法分类号:P237.9

随着遥感技术的发展,出现了利用植被指数和土地表面温度(LST)进行干旱监测的方法。已开发出了基于归一化植被指数(NDVI)的条件植被指数(vegetation condition index, VCI)和距平植被指数(anomaly vegetation index, AVI)以及基于 LST 的条件温度指数(temperature condition index, TCI)等干旱监测方法^[1-4]。这些指数已用于干旱预测^[5-6],如 Liu 等利用巴西东北部平均的 NDVI 距平时间序列和多个 ENSO (El Niño southern oscillation) 指数序列,建立 NDVI 距平值的预测方程,预测该地区 ENSO 干旱的发生^[5]。Seiler 等采用干旱监测指数 VCI 和 TCI 对阿根廷主要农作物产区的天气和干旱状况进行评估^[6]。

王鹏新等基于 NDVI 和 LST 的散点图呈三角形区域分布的条件,提出了条件植被温度指数(vegetation temperature condition index, VTCI)的干旱监测方法^[7-8]。该方法充分利用植被指数与土地表面温度互补特性提供的作物水分亏缺信息来监测旱情,其应用表明该方法适用于监测某一时期的干旱程度^[8-9]。如何利用遥感的 VTCI 数据进行干旱预测就成为本论文的研究重点。本研究以陕西关中平原为例,基于 VTCI 干旱监测

结果,运用 ARIMA (autoregressive integrated moving average) 模型对 VTCI 时空序列进行分析建模,对关中平原干旱状况进行分析预测。

1 原理和数据源

1.1 条件植被温度指数

条件植被温度指数(VTCI)的定义为^[7-8]:

$$VTCI = \frac{LST_{NDVI_i, \max} - LST_{NDVI_i}}{LST_{NDVI_i, \max} - LST_{NDVI_i, \min}} \quad (1)$$

其中, $LST_{NDVI_i, \max} = a + bNDVI_i$; $LST_{NDVI_i, \min} = a' + b'NDVI_i$ 。

式中, $LST_{NDVI_i, \max}$ 和 $LST_{NDVI_i, \min}$ 分别表示在研究区域内,当 NDVI_i 值等于某一特定值时的土地表面温度(LST)的最大值和最小值; a 、 b 、 a' 、 b' 为 4 个待定系数,可通过绘制研究区域的 NDVI 和 LST 的散点图近似地获得。4 个待定系数的确定方法非常重要,这里采用最大-最小值合成方法,即应用多年的旬 NDVI 最大值合成产品和基于每年的旬 LST 最大值合成产品再合成的多年 LST 最小值产品确定冷边界 LST_{\min} 的系数 a' 和 b' ; 使用最大值合成技术,即应用多年的旬 NDVI 和 LST 最大值合成产品确定热边界 LST_{\max} 的系

数 a 和 b 。VTCI 的值越小,表明干旱程度越严重。

条件植被温度指数 VTCI 干旱监测的优点是既考虑了研究区域内 NDVI 的变化,又考虑了 NDVI 值相同时 LST 的变化,适用于一特定年内某一时期的区域内相对于干旱程度的监测。

1.2 ARIMA 模型

ARIMA (autoregressive moving average) 模型的基本思想是通过研究时间序列的有序性以及序列观测值之间的依存关系,建立能较精确地反映时间序列中所包含的动态依存关系的数学模型,然后利用该模型预测序列未来的变化^[10, 11]。

对于只含有趋势性的非平稳时间序列,需先进行 d 次逐期差分运算转化为平稳序列,然后再建立 ARMA 模型,称为 ARIMA(p, d, q)模型。对于既含有趋势性又含有季节性变化的非平稳时间序列,可进行 d 次逐期差分以消除趋势性,再进行 D 次季节性差分以消除季节性变化特征,将序列转化为平稳时间序列进行建模。模型表示为 ARIMA(p, d, q) \times (P, D, Q) $_s$, 又称乘积季节模型, P, Q 为季节性因素的自回归移动平均的阶次, S 为季节性周期长度。模型形式为:

$$\nabla^d \nabla_s^D X_t = u + \frac{\theta(B)\Theta_s(B)}{\varphi(B)\Phi_s(B)} a_t \quad (2)$$

式中, $\nabla^d \nabla_s^D$ 表示序列经过 d 阶逐期差分 and 周期为 S 的季节差分; u 为序列的均值; a_t 为白噪声序列。

$$\begin{aligned} \theta(B) &= 1 - \theta_1 B - \dots - \theta_q B^q \\ \varphi(B) &= 1 - \varphi_1 B - \dots - \varphi_p B^p \\ \Theta_s(B) &= 1 - \Theta_1 B^S - \dots - \Theta_Q B^{QS} \\ \Phi_s(B) &= 1 - \Phi_1 B^S - \dots - \Phi_P B^{PS} \end{aligned} \quad (3)$$

式中, B 为后移算子 ($B^k X_t = X_{t-k}$), $\varphi_i, \theta_i, \Theta_i, \Phi_i$ 为模型参数。当 P, D, Q 都为零时,得到 ARIMA (p, d, q) 模型。可见 ARIMA (p, d, q) 模型是 ARIMA (p, d, q) \times (P, D, Q) $_s$ 模型的特例。ARIMA 模型的建模步骤包括:模型的平稳性检验、定阶、参数估计、验证和优化等^[10, 11]。

1.3 数据源

关中地区位于陕西省中部,西起宝鸡,东至潼关,北起北山,南止秦岭。该地区属大陆性季风气候,处于暖温带半湿润与半干旱气候的过渡地带,多年平均降雨量 500~700 mm^[12]。行政区域包括:铜川、宝鸡、咸阳、西安和渭南 5 个市。研究区域气象站的分布如图 1 所示。旱灾是陕西省的主要自然灾害之一,较大区域的旱灾又以关中地区发生灾害的次数最多,平均每 7 a 一次^[13]。



图 1 研究区域气象站的分布

Fig. 1 Locations of the Weather Stations in the Study Area

本研究应用关中平原 1999~2006 年 3~6 月的 AVHRR 多时段卫星遥感反演的 NDVI 和 LST 产品,根据文献[7-8]的方法计算 1999~2006 年 3~6 月各旬的 VTCI 数据,得到关中平原的 VTCI 序列。利用干旱指数 VTCI 对关中平原未来的干旱状况进行分析预测。将 1999 年 3 月上旬到 2006 年 3 月下旬的数据作为分析建模数据,2006 年 4 月上旬和 4 月中旬的数据作为检验数据,对预测模型的有效性和预测精度进行检验。

2 VTCI 序列的 ARIMA 模型预测

研究区域的 VTCI 数据是时空序列,具有时间和空间的变化。对时空变化的 VTCI 序列进行预测较为复杂,需要找到研究区域中各个像素点的 VTCI 随时间的变化规律,建立适合的数学模型。应用 ARIMA 模型作预测,模型的定阶十分重要,它关系到所建模型是否体现了时间序列的变化规律,这往往需要建模者的主观判断。因此对每个像素点对应 VTCI 时间序列进行定阶判断,再建立适合的 ARIMA 预测模型是比较困难的。这里采用的方法是对该区域的某些像素点的 VTCI 时间序列的变化规律进行研究,确定这些像素点适合的 ARIMA 模型,再对整个区域所有像素点的 VTCI 时间序列进行建模和预测。

首先选用 36 个气象站点所在像素的 VTCI 时间序列的变化规律进行 ARIMA 模型的建模分析,分析方法依照 ARIMA 模型建模步骤^[10-11]。选用气象站点的优点是气象站点的分布是大致均匀的,这些站点具有一定的区域代表性,可认为其邻近像素点干旱状况与其遵循同样的变化规律。其次是可以参照这些像素点的气象数据,更好地分析 VTCI 时间序列的变化规律。

根据 ARIMA 的建模方法,对 1999 年 3 月上旬至 2006 年 3 月下旬的 36 个气象站点的 VTCI 时间序列进行建模分析。结果表明,这些序列都

是平稳的(图2为部分气象站像素点时间序列的自相关函数图),均适合AR(1)模型,即

$$(X_t - \mu) - \varphi_1(X_{t-1} - \mu) = a_t \quad (4)$$

式中, X_t 为监测结果中某像素点的时间序列; μ 为该像素点时间序列的均值; a_t 为白噪声序列; φ_1 为AR(1)模型的惟一参数,反映零均值序列 $X_t - \mu$ 对 $X_{t-1} - \mu$ 的依赖程度。

φ_1 的最小二乘估计为:

$$\varphi_1 = \frac{\sum_{t=1}^{N-1} (X_t - \mu)(X_{t+1} - \mu)}{\sum_{t=1}^{N-1} (X_t - \mu)^2} \quad (5)$$

式中, N 为时间序列的总序列数。

根据预测误差最小原则,可以得到 X_t 的未来值的 l 步预测:

$$Y_{t+l} - \mu = \varphi_1^l (Y_t - \mu) \quad (6)$$

式中, Y_{t+l} 为未来 $t+l$ 时刻的 VTCI; Y_t 为已知历史数据中最后 t 时刻的 VTCI 值。

考虑气象站点的分布是大致均匀的,它们相应的 VTCI 时间序列都适合 AR(1) 模型,可以认为其他像素点的 VTCI 时间序列也是按照同样的 AR(1) 模型的规律变化。因此,对 1999 年 3 月上旬至 2006 年 3 月下旬研究区域所有的像素点的 VTCI 时间序列建立 AR(1) 模型,并依照式(6)进行 1 步和 2 步预测,即预测 2006 年 4 月上旬和 4 月中旬的 VTCI 变化,并将预测结果与实际的 VTCI 数据进行对比分析。

图 3 为所建 AR(1) 模型的参数 φ_1 分布图,该图显示参数呈现区域性分布,南部地区的 φ_1 较北部地区的偏大。

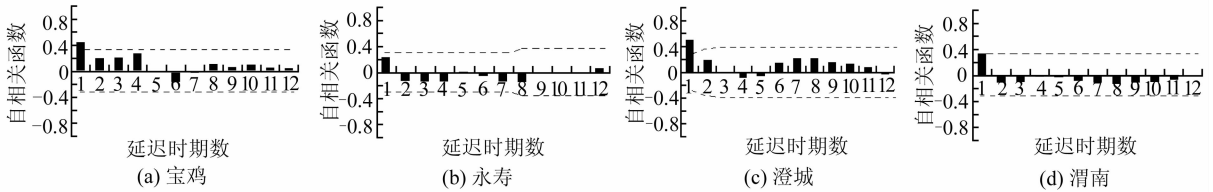


图 2 部分气象站像素点时间序列的自相关函数

Fig. 2 Parts of ACF Plots of VTCI Series from 36 Pixels Where Weather Stations Are Located



图 3 φ_1 参数分布

Fig. 3 Distribution of the Parameter φ_1

图 4 和图 5 分别为预测的 2006 年 4 月上旬和 4 月中旬的 VTCI 影像及其频数分布,图 6 为预测误差图(监测结果中各像素点的 VTCI 值减去预测结果中对应的值)及其频数分布。表 1 给出相应的统计结果。从中可以分析得出:

1) 1 步和 2 步的预测结果都基本反映了监测结果的整体特征(图 4 和图 5),即关中平原的北部地区较南部地区干旱(VTCI 值越小反映干旱程度越严重),黄河两岸地区较为干旱。由监测和

预测结果的频数分布对比来看,监测与预测结果呈现相似的频数分布,最大频数值很接近,1 步预测为 0.65 和 0.58,2 步预测为 0.73 和 0.59。不同的是 VTCI 的取值范围,比较频数大于 50 的 VTCI 的取值范围,1 步预测为 [0.55, 0.93] 与 [0.52, 0.73],2 步预测为 [0.63, 0.9] 与 [0.47, 0.67]。表 1 显示监测结果和预测结果中统计的 VTCI 最大、最小、平均值、标准误差也很接近。

2) 对细节的反映上,1 步预测的精度明显高于 2 步预测的精度(图 4 和图 5)。1 步预测结果中除了少数区域外,基本反映了实际的 VTCI 影像的分布。图 6 的预测误差分析中也显示了这一点,1 步预测最大频数预测误差值为 0.06,2 步预测最大频数预测误差值 0.15。表 1 显示 1 步预测的平均预测误差为 0.10,2 步预测的平均预测误差为 0.17,也反映出 1 步预测精度好于 2 步预测。

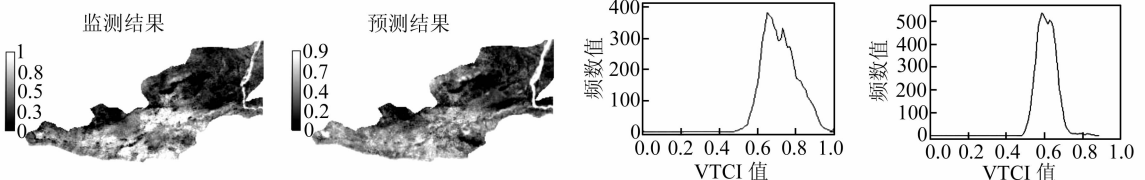


图 4 2006 年 4 月上旬 VTCI 监测和预测结果及其频数分布

Fig. 4 Monitoring, Forecasting Results and Their Frequency Distributions for the First Ten Days of April in 2006

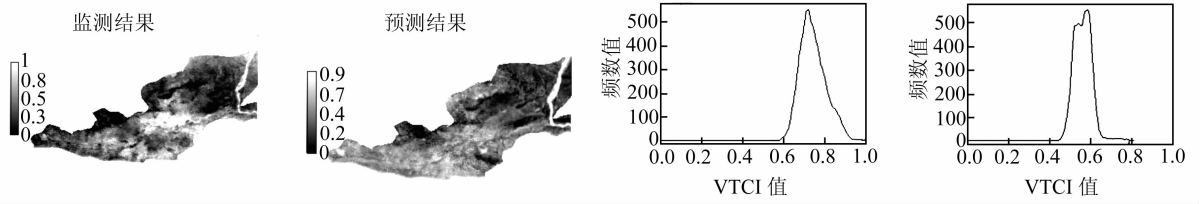


图 5 2006 年 4 月中旬 VTCI 监测和预测结果及其频数分布

Fig. 5 Monitoring and Forecasting Results and Their Frequency Distributions for the Middle Ten Days of April in 2006

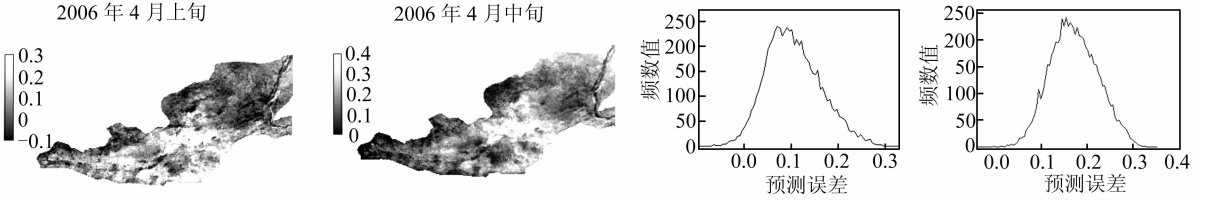


图 6 2006 年 4 月上旬、4 月中旬的预测误差及其频数分布

Fig. 6 Forecasting Errors and Their Frequency Distributions for the First and Middle Ten Days of April in 2006

表 1 监测结果、预测结果、预测误差图相应的统计结果

Tab. 1 Statistics for Monitoring and Forecasting Results, and the Forecasting Errors

VTCI 影像	最小值	最大值	平均值	标准误差
2006/4/上(监测)	0	1	0.72	0.89
2006/4/上(预测)	0	0.88	0.61	0.051
2006/4/中(监测)	0	1	0.73	0.12
2006/4/中(预测)	0	0.82	0.57	0.046
2006/4/上(误差)	-0.096	0.33	0.10	0.059
2006/4/中(误差)	-0.038	0.35	0.17	0.052

3 VTCI 序列的 ARIMA 模型拟合

对 VTCI 序列所建 AR(1)模型,是由 36 个气象站的单点时间序列的建模分析而确定的,ARIMA 模型的建模方法决定了数据拟合最高的参数是最优模型的参数,因此对这些像素点的历史数据的拟合精度较好。然而 AR(1)模型对其他的像素点的拟合精度并不能确定,需要对整个区域的历史数据的拟合情况进行分析,以此判断所建模型对整个区域的 VTCI 时间序列拟合的优劣。

依据以下公式,对 1999 年 3 月上旬到 2006 年 3 月下旬每旬的逐个 VTCI 影像进行拟合,并对拟合结果进行统计分析:

$$\begin{cases} \hat{X}_i = \mu \\ \hat{X}_i = \mu + \varphi_1(X_{i-1} - \mu) \end{cases} \quad (6)$$

式中, \hat{X}_i 为拟合影像中某像素点 VTCI 拟合序列; μ 为序列 X_i 的均值; \hat{X}_1 为 1999 年 3 月上旬的赋予值; X_{i-1} 为监测结果中相应像素点的 VTCI 序列; φ_1 为 AR(1)模型的参数。

图 7 是部分 VTCI 影像的拟合情况。可以看出拟合影像基本反映了原来监测结果的特征。各旬的拟合误差统计分析显示各旬的拟合误差大多在 -0.2, 0.2 区间,最大频数的误差在 -0.15 和 0.05 之间,表明拟合精度较好。

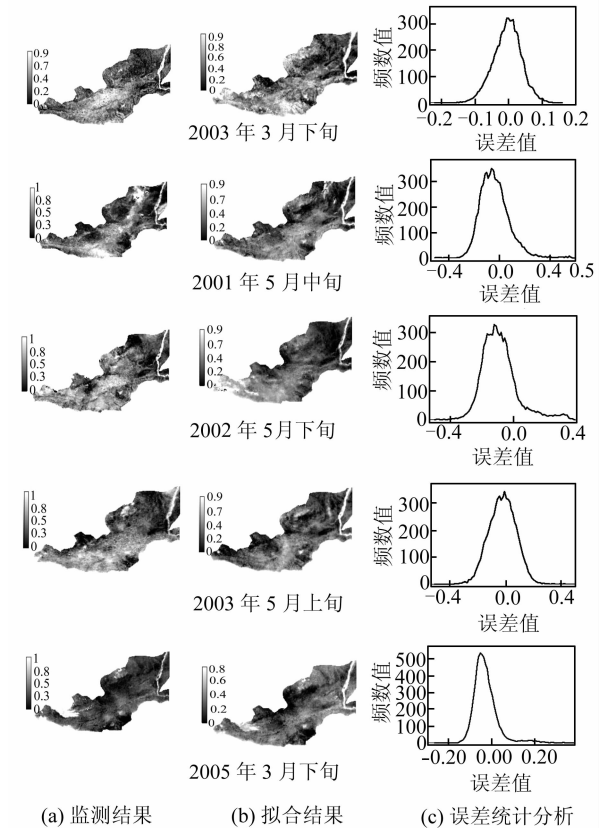


图 7 部分 VTCI 影像的拟合

Fig. 7 Parts of Monitoring Results, Simulating Results and the Simulating Errors

图8是各旬的平均拟合误差的散点图,从图中可以看出平均拟合误差的分布情况。图中显示平均拟合误差大部分分布在 $[-0.2, 0.2]$ 区域,只有5旬分布此区域外,误差大于0.2。表2是平均拟合误差的分布统计,平均拟合误差在 -0.1 和 0.1 之间的占65.7%,在 -0.2 和 0.2 之间的占92.6%,表明AR(1)模型对VTCI时空序列的拟合较好。

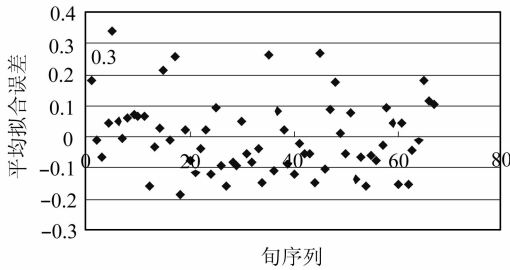


图8 平均拟合误差的散点图

Fig. 8 Scatter Plot of Mean Simulating Errors

表2 平均拟合误差分析

Tab. 2 Statistics of Mean Simulating Errors

平均拟合误差 σ 分布	旬的数目	百分比/%
$ \sigma < 0.01$	44	65.7
$0.1 < \sigma < 0.2$	18	26.9
$ \sigma > 0.2$	5	7.5

4 结 语

以陕西关中平原地区为例,基于干旱指数VTCI序列,采用时间序列ARIMA模型的建模预测方法,对该地区的干旱状况进行分析预测。VTCI序列为时空序列,这里采用的是由点到面的分析方法,先对36个气象站点的像素点建立适合的ARIMA模型,再对整个区域所有像素点的VTCI时间序列进行建模预测。研究表明,整个区域每个像素点的VTCI时间序列都适合AR(1)模型。将预测结果与原始的监测结果进行对比,显示预测结果大致反映了实际监测结果的变化,1步预测的平均预测误差为0.10,2步预测的平均预测误差为0.17,1步预测精度好于2步预测。对原始序列进行历史拟合,反映出模型拟合情况较好,平均拟合误差大部分分布在 $[-0.2, 0.2]$ 区域,平均拟合误差绝对值小于0.1的占总体的65.7%,显示AR(1)模型适合VTCI序列,可以利用它进行1~2步预测。

参 考 文 献

- [1] Boken V k, Cracknell A P, Heathcote R L. Monitoring and Predicting Agricultural Drought; a Global study[M]. New York: Oxford University Press, 2005: 57-87
- [2] 陈维英,肖乾广,盛永伟. 距平植被指数在1992年特大干旱监测中的应用[J]. 环境遥感, 1994, 9: 106-112
- [3] Kogan F N. Remote Sensing of Weather Impacts on Vegetation in Non-homogeneous Areas [J]. International Journal of Remote Sensing, 1990, 11: 1405-1420
- [4] Kogan F N. Application of Vegetation Index and Brightness Temperature for Drought Detection[J]. Advances in Space Research, 1995, 15: 9-100
- [5] Liu W T, Negrón Juárez R I. ENSO Drought Onset Prediction in Northeast Brazil Using NDVI[J]. International Journal of Remote Sensing, 2001, 22 (17): 3483-3501
- [6] Seiler R A, Kogan F, Wei Guo. Monitoring Weather Impact and Crop Yield from NOAA AVHRR Data in Argentina[J]. Advances in Space Research, 2000, 26(7): 1177-1185
- [7] 王鹏新, 龚健雅, 李小文. 条件植被温度指数及其在干旱监测中的应用[J]. 武汉大学学报·信息科学版, 2001, 26(5): 412-418
- [8] 王鹏新, Wan Zhengming, 龚健雅, 等. 基于植被指数和土地表面温度的干旱监测模型[J]. 地球科学进展, 2003, 18(4): 527-533
- [9] Wan Z, Wang P, Li X. Using MODIS Land Surface Temperature and Normalized Difference Vegetation Index Products for Monitoring Drought in the Southern Great Plains, USA [J]. International Journal of Remote Sensing, 2004, 25(1): 61-72
- [10] Box G E P, Jenkins G M. Time Series Analysis Forecasting and Control [M]. San Francisco: Holden-Day, 1976: 525
- [11] 易丹辉. 统计预测——方法与应用[M]. 北京: 中国统计出版社, 2001
- [12] 陕西历史自然灾害简要记实编委会. 陕西历史自然灾害简要记实[M]. 北京: 气象出版社, 2002
- [13] 温克刚, 翟佑安. 中国气象灾害大典(陕西卷)[M]. 北京: 气象出版社, 2005

第一作者简介: 韩萍, 博士生, 主要研究领域: 干旱监测和预测。
E-mail: wangpx@cau.edu.cn