

基于非线性特征的干旱影响评估研究

王鹏新¹ 冯明悦¹ 孙辉涛¹ 李俐¹ 张树誉² 景毅刚²

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 陕西省气象局, 西安 710014)

摘要: 客观地认识干旱的非线性特征是干旱影响评估的关键,对制定抗旱减灾策略具有重要指导意义。以陕西省关中平原为研究区域,以核函数方法为非线性算法,基于核主成分分析方法(KPCA),将遥感反演的条件植被温度指数(VTCI)映射到高维特征空间下对其进行特征提取,并结合 Copula 函数构建主成分间的联合分布模型,确定2008—2013年冬小麦主要生育期的综合 VTCI;构建综合 VTCI 与冬小麦单产间的线性回归模型,评估干旱对冬小麦产量的影响。结果表明,相比于传统的主成分分析方法(PCA),KPCA 能有效地提取干旱的非线性特征,且降维效果更好。与 PCA-Copula 方法构建的回归模型相比,应用 KPCA-Copula 方法所建综合 VTCI 与单产间的回归模型的拟合度明显提高,决定系数达到 0.608 ($p < 0.001$),对应模型的估测单产与实测单产之间的均方根误差(RMSE)为 298.1 kg/hm²,相比于 PCA-Copula 的结果降低了 60.1 kg/hm²,且 KPCA-Copula 获取的综合 VTCI 更符合关中平原实际的干旱特征。这表明 KPCA-Copula 方法能够较好地体现干旱的非线性特征,更加适用于干旱影响评估研究。

关键词: 干旱影响评估; 条件植被温度指数; 核主成分分析; 非线性特征; Copula 函数

中图分类号: S127 **文献标识码:** A **文章编号:** 1000-1298(2016)10-0325-07

Drought Impact Assessment Based on Nonlinear Characteristics of Drought

Wang Pengxin¹ Feng Mingyue¹ Sun Huitao¹ Li Li¹ Zhang Shuyu² Jing Yigang²

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Shaanxi Provincial Meteorological Bureau, Xi'an 710014, China)

Abstract: Drought is a typical complex system, and nonlinear characteristics of drought are the concentrated reflection of its complexity. Therefore, objectively understanding of complex nonlinear characteristics of drought is the key approach of assessing the effects of drought, which can provide guideline for making drought mitigation strategies. Guanzhong Plain was chosen as study area, and the kernel method was applied as a nonlinear algorithm. Based on the kernel principal component analysis (KPCA), vegetation temperature condition index (VTCI) retrieved from MODIS was projected into a high-dimensional feature space for feature extraction, and then the joint distribution model of principal components with Copula function was built. Comprehensive values of VTCIs at main growth stages from 2008 to 2013 were determined by using the joint distribution model (the KPCA-Copula method). Linear regression models between the comprehensive VTCIs and wheat yields were established to assess the effect of drought on wheat yields. The results showed that the KPCA could effectively extract the nonlinear characteristics of drought, and it had better performance in dimension reduction compared with the principal component analysis (PCA). Compared with the PCA-Copula method, the determination coefficient of regression model between wheat yields and comprehensive VTCIs with KPCA-Copula method reached 0.608 ($p < 0.001$), which indicated that the fitting degree of the model was improved, and the root mean square error (RMSE) between estimated yields and measured ones was 298.1 kg/hm², which

收稿日期: 2016-04-07 修回日期: 2016-05-18

基金项目: 国家自然科学基金项目(41371390)

作者简介: 王鹏新(1965—),男,教授,博士生导师,主要从事定量遥感及其在农业中的应用研究,E-mail: wangpx@cau.edu.cn

was about 60.1 kg/hm^2 lower than the RMSE by using PCA - Copula method. The comprehensive VTCl with KPCA - Copula method were more in line with actual drought characteristics of Guanzhong Plain. These results indicated that the KPCA - Copula method could well reflect nonlinear characteristics of drought, and it had good applicability in drought impact assessment.

Key words: drought impact assessment; vegetation temperature condition index; kernel principal component analysis; nonlinear characteristics; Copula function

引言

灾害系统是典型的复杂系统,存在着多因子、多系统错综复杂的相互作用,干旱的非线性特征是其复杂性的集中体现^[1],客观地认识干旱的非线性特征是准确地评估干旱对农作物产量影响的关键。核函数方法是一种有效地设计非线性算法的数学手段,其主要思想是通过一个特征映射将输入空间(低维)中的线性不可分数据映射为高维特征空间中的线性可分数据,进而在特征空间进行对应的线性操作^[2-3]。由于运用非线性映射,高维空间的内积运算往往非常复杂,而通过引入由特征空间内积确定的核函数,从而无须直接在高维空间中进行复杂的计算,极大地降低了问题的计算复杂度^[4],并增强了对非线性数据处理的能力。目前,已经提出了大量基于核的特征抽取算法,其应用领域包括模式识别、图像视觉、基于内容的检索及文本分类等^[5-7],其中核主成分分析(Kernel principal component analysis, KPCA)是常用的基于核的非线性特征抽取技术之一^[8]。

KPCA作为传统主成分分析(Principal component analysis, PCA)方法的改进算法,是核方法与PCA相结合的一种降维方法,它通过核函数将原始数据隐式地映射到一个高维的特征空间,在保留数据主要信息的前提下,提取数据中的非线性特征信息,去除数据间的相关性,能够弥补PCA难以提取数据集间非线性特征的不足,已成为解决复杂非线性问题的一种有效途径^[9-10]。HUANG等^[11]利用KPCA具有分析非线性相关变量的特点,将其与支持向量数据描述方法相结合,达到故障诊断的目的。彭令等^[12]基于核主成分分析法对滑坡位移影响因子进行特征提取,以获得的主成分作为特征向量,进而对滑坡相对位移进行预测。PAN等^[13]利用传感器的监测数据建立数据之间的关联集,将其与基于KPCA确定的观测数据的特征矩阵进行对比,从而实现卫星电源系统的异常检测。

在归一化植被指数(NDVI)和地表温度(LST)的散点图呈三角形区域分布的基础上,王鹏新等^[14-15]提出条件植被温度指数(VTCl)的干旱监测

方法,并成功应用于干旱的实时监测、预测和影响评估等研究^[16-17]。而以往的干旱影响评估由于较少考虑干旱的非线性特征,因而影响了干旱影响评估的精度,且研究已经证明Copula函数对基于VTCl的干旱影响评估研究具有较好的适用性^[18]。鉴于此,本文以陕西省关中平原冬小麦为研究对象,以KPCA为非线性特征提取方法,获取2008—2013年冬小麦主要生育期VTCl的主成分因子,并通过Copula函数确定冬小麦主要生育期的综合VTCl值,以期获取更适用于评估复杂、非线性的干旱对冬小麦产量影响研究的方法。

1 材料与方法

1.1 研究区域概况

关中平原位于陕西省中部的渭河流域,其行政区域包括西安、铜川、宝鸡、咸阳、渭南5个省辖市和杨凌示范区,覆盖范围为东经 $106^{\circ}22'$ ~ $110^{\circ}24'$,北纬 $33^{\circ}57'$ ~ $35^{\circ}39'$ 。该区土壤肥沃,有泾河、渭河等河流提供水源,十分适宜耕作,是陕西省的农业基地,作物种植模式主要为冬小麦与夏玉米轮作。冬小麦播种时间一般为10月上、中旬,播种后5~6d出苗,根据关中平原冬小麦的实际生长状况,将其主要生育期划分为:返青期(3月上旬—3月中旬)、拔节期(3月下旬—4月中旬)、抽穗—灌浆期(4月下旬—5月上旬)和乳熟期(5月中旬—5月下旬)^[17]。

1.2 试验数据

VTCl是基于遥感反演的NDVI和LST特征空间呈三角形区域分布的特点提出的,主要用于监测旱情。基于Aqua-MODIS数据产品(MYD09GA和MYD11A1),应用最大值合成技术分别生成旬NDVI和LST最大值合成产品,并以此计算VTCl^[14,19],生成2008—2013年每年3—5月份以旬为单位的关中平原4市(渭南、西安、咸阳和宝鸡)的VTCl数据。依据关中平原的行政边界图及冬小麦的实际生长状况,取各市区内所包含像素的VTCl平均值作为该区域该年该旬的VTCl值,再根据生育期包含的旬VTCl的平均值作为该年该生育期的VTCl值。

1.3 KPCA的基本原理

基于核函数的主成分分析法是利用线性代数、

支持向量机等有关理论来实现非线性空间的降维, 通过某种选定的非线性映射将输入矢量 \mathbf{X} 映射到一个高维线性特征空间中, 然后在特征空间中使用 PCA 方法计算得到的线性主成分实质上就是原始输入空间的非线性主元^[20], 不仅能够实现数据降维, 而且可以有效地处理各指标的非线性影响。

通过引入非线性映射函数 Φ , 使输入空间的 VTCI 数据 x_1, x_2, \dots, x_p 变换为特征空间中的样本点 $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_p)$, 并假设 $\sum_{j=1}^p \Phi(x_j) = 0$, 则在特征空间中的协方差矩阵为

$$\mathbf{C} = \frac{1}{p} \sum_{j=1}^p \Phi(x_j) \Phi(x_j)^T \quad (1)$$

对于一般 PCA 方法, 即通过求特征方程 $\lambda' \mathbf{v}' = \frac{1}{p} \sum_{j=1}^p \mathbf{x}_j \mathbf{x}_j^T \mathbf{v}'$, 获得特征值 λ' 及对应的特征向量 \mathbf{v}' , 而 KPCA 是求解方程 $\lambda \mathbf{V} = \mathbf{C} \mathbf{V}$ 在特征空间中的特征值 λ 及对应的特征向量 \mathbf{V} , 进而有

$$\lambda (\Phi(x_i) \mathbf{V}) = \Phi(x_i) \mathbf{C} \mathbf{V} \quad (i=1, 2, \dots, p) \quad (2)$$

由于 \mathbf{V} 是 $\Phi(x_j)$ 的线性组合, 必定存在一组 a_j 使得 $\mathbf{V} = \sum_{j=1}^p a_j \Phi(x_j)$ 成立, 将其代入式(2)可得

$$\begin{aligned} & \lambda \sum_{j=1}^p a_j (\Phi(x_j)^T \Phi(x_i)) = \\ & \frac{1}{p} \sum_{j=1}^p a_j \sum_{i=1}^p (\Phi(x_i)^T \Phi(x_j)) (\Phi(x_j)^T \Phi(x_i)) \end{aligned} \quad (3)$$

定义 $p \times p$ 矩阵 \mathbf{K} , 即核函数为

$$\mathbf{K}_{ij} = \Phi(x_i)^T \Phi(x_j) = k(x_i, x_j) \quad (4)$$

将式(4)代入式(3), 可得 $p\lambda \mathbf{K} \mathbf{a} = \mathbf{K}^2 \mathbf{a}$, 即

$$\eta \mathbf{a} = \mathbf{K} \mathbf{a} \quad (5)$$

式中 η ——矩阵 \mathbf{K} 对应的特征值

\mathbf{a} ——特征值 η 对应的特征向量

由式(5)可以看出, 高维空间的内积运算可以转化为核函数, 而无须直接在高维空间中进行复杂的计算。

为了提取主要生育期 VTCI 的主成分, 需要计算其对应的映射数据 $\Phi(x)$ 在特征空间中的特征向量 \mathbf{V}^k 上的投影

$$\begin{aligned} y_k = \mathbf{V}^k \Phi(x) &= \sum_{i=1}^p a_{k,i} (\Phi(x_i) \Phi(x)) = \\ & \sum_{i=1}^p a_{k,i} \mathbf{K}(x_i, x) \end{aligned} \quad (6)$$

式中 \mathbf{V}^k —— \mathbf{V} 中的第 k 个特征向量

$a_{k,i}$ ——矩阵 \mathbf{K} 的第 k 个特征值对应的特征向量 \mathbf{a}_k 的第 i 个数

由于上述推导均是在假定映射数据的均值为零

的前提下进行的, 如这一假设条件难以成立, 此时式中的 \mathbf{K} 用 \mathbf{K}' 代替, 即

$$\mathbf{K}' = \mathbf{K} - \mathbf{I}_p \mathbf{K} - \mathbf{K} \mathbf{I}_p + \mathbf{I}_p \mathbf{K} \mathbf{I}_p \quad (7)$$

式中 \mathbf{I}_p —— $p \times p$ 的矩阵, 且满足 $\mathbf{I}_p = (1/p) \mathbf{1}_{p \times p}$

1.4 基于 KPCA - Copula 的联合分布模型

由于核函数的形式和参数的变化会隐式地改变从输入空间到特征空间的映射, 从而对特征空间的性质产生影响, 最终改变各种核函数方法的性能, 因而需要选取适当的核函数, 求出矩阵 \mathbf{K} , 基于此, 通过式(6)获取关中平原冬小麦主要生育期 VTCI 的主成分因子, 进而基于 Copula 函数构建各主成分因子的联合分布。

1.4.1 核函数的选择

核函数是满足 Mercer 条件的任意对称函数, 只要选定核函数, 就意味着选定某种映射。常用的核函数有径向基 (Radial basis function, RBF) 核函数、多项式核函数、感知器核函数等, 本研究选择运用最广泛的 RBF, 无论低维、高维、小样本、大样本等情况, RBF 核函数均适用^[21-22], 其表达式为

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\theta^2}\right) \quad (8)$$

式中 θ ——高斯径向基函数的宽度参数, 控制函数的径向作用范围

$\|\cdot\|$ ——2 范数

基于 RBF 核函数, 通过比较不同核参数值 (即 θ 的取值) 对应各主成分的贡献率, 选择能用较少的主成分概括数据本身信息所对应的参数值, 从而确定参数, 获取主要生育期 VTCI 的主成分因子。

1.4.2 边缘分布函数的优选

目前常用的几种分布线型有皮尔逊 III 型分布、指数分布、正态分布、Gamma 分布、广义极值分布等, 其中, 依据广义极值分布的形状参数将其分为 Gumbel、Frechet 和 Weibull 3 种极值分布^[23]。对于各主成分的边缘分布函数的选择, 首先通过极大似然法估计边缘分布的参数, 然后通过 Kolmogorov - Smirnov (K - S) 方法检验边缘分布函数拟合效果, 最后通过比较理论频率与经验频率之间的均方根误差 (RMSE) 以及信息准则值 AIC (Akaike information criterion) 优选边缘分布函数。边缘分布优选的原则是 RMSE 及 AIC 越小拟合效果越好^[18], 基于此, 优选出各主成分对应的边缘分布。

1.4.3 干旱影响评估模型的构建

Copula 函数可以通过边缘分布和相关性结构两部分来构造联合分布。KPCA 降维方法获取的主成分因子间相互独立, 基于各主成分优选的边缘分布函数, 采用独立 Copula 函数构建各主成分因子之间

的联合分布,获取综合 VTCI 值(C_{VTCI})

$$C_{VTCI} = C(u_1, u_2, \dots, u_m) = u_1 u_2 \dots u_m \quad (9)$$

式中 m ——选择的主成分个数

u_m ——第 m 个主成分因子优选的边缘分布函数

基于确定的综合 VTCI 值,建立其与冬小麦单产之间的一元线性回归模型,并评估干旱对冬小麦产量的影响。

2 结果与分析

2.1 各生育期 VTCI 的 KPCA 分析

核函数存在未知参数,不同的核参数值会得到不同的核矩阵,对 KPCA 方法的效果有较大影响。本研究采用的 RBF 核函数存在一个参数 θ , 参数值的选择原则是尽量使用较少的主成分来概括数据本身信息,从而使降维效果更明显。基于 RBF 核函数的不同参数值 θ , 获取对应的各主成分贡献率(图 1)。结果表明,随着 θ 值的增大,第 1 主成分的贡献率明显增大,同时前 3 个主成分的累计贡献率也不断增大,当参数值大于 5 时,这种变化不再明显,各主成分贡献率趋于稳定。将 KPCA 核参数值为 5 的累计贡献率与 PCA 方法的结果对比,可以看出, KPCA 的第 1 主成分的贡献率达到 78.02%, 且前 2 个主成分的累计贡献率达到 97.54%; 而采用 PCA 获取的前 3 个主成分的贡献率分别为 62.02%、21.11%、14.50%, 累计贡献率为 97.63%, 意味着经过 KPCA 获取的前 2 个主成分涵盖的信息相当于 PCA 获取的前 3 个主成分的结果。可见,只要选择合适的核参数,就能使 KPCA 方法较 PCA 方法用更少的主成分来表达原始数据信息,降维效果有所提高。因此,本研究选择核函数参数值为 5 对应的前 2 个主成分因子构建干旱影响评估模型。

2.2 干旱影响评估模型的构建与分析

2.2.1 边缘分布函数的确定

基于目前常用的分布线型,选取正态分布、对数分布和广义极值分布分别拟合基于 KPCA 获取的关中原冬小麦 4 个生育期 VTCI 的前 2 个主成分因子值,利用 K-S 检验对各主成分因子的边缘分布函数进行拟合优度评价。结果表明,前 2 个主成分的广义极值分布和正态分布均能通过 K-S 检验,其中,第 1 主成分属于广义极值分布中的 Frechet 分布,第 2 主成分属于广义极值分布中的 Weibull 分布。基于正态分布及对应的极值分布类型,结合 RMSE 及 AIC 越小拟合效果越好的优选原则,进一步获取各边缘分布对应的 RMSE 和 AIC(表 1),可以看

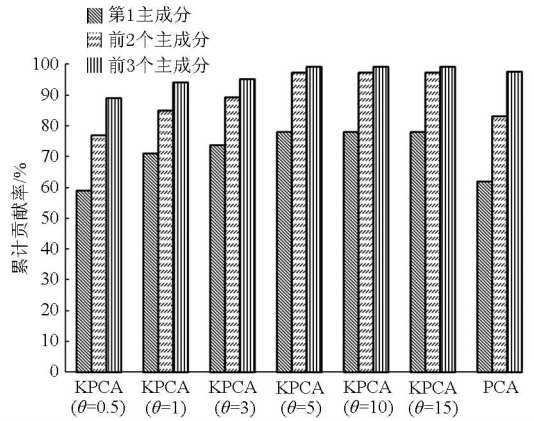


图 1 基于不同参数值的 KPCA 与 PCA 方法的各主成分累计贡献率比较

Fig. 1 Comparison of cumulative contribution rates of principal components between KPCA and PCA methods based on different parameter values

出,第 1 主成分优选 Frechet 分布;第 2 主成分优选 Weibull 分布。尽管 2 个主成分的边缘分布各异,但 Copula 函数不限制单变量边缘分布的形式,能够灵活地用于构建 2 个主成分间的联合分布。

表 1 冬小麦主要生育期 VTCI 主成分因子的分布线型拟合效果

Tab. 1 Distribution of fitting results of principle components of VTCI at main growth stages of winter wheat

边缘分布	第 1 主成分		第 2 主成分		
	RMSE	AIC	边缘分布	RMSE	AIC
正态分布	0.07	-21.74	正态分布	0.09	-19.34
Frechet	0.03	-31.30	Weibull	0.06	-25.35

2.2.2 综合 VTCI 与单产间线性回归模型的对比

基于 KPCA 的冬小麦主要生育期 VTCI 的 2 个主成分的优选边缘分布,结合 Copula 函数建立各主成分间的联合分布,获取综合 VTCI,并建立其与冬小麦单产之间的一元线性回归模型(表 2)。可以看出,综合 VTCI 与冬小麦单产间的相关性达极显著水平($p < 0.001$),且基于 KPCA - Copula 方法确定的综合 VTCI 与单产间线性回归模型的决定系数($R^2 = 0.608$)明显高于采用 PCA - Copula 确定的回归模型的决定系数($R^2 = 0.490$),模型拟合度较高,将该回归模型作为冬小麦单产的估测模型,对应的估测单产与实测单产之间的均方根误差为 298.1 kg/hm²,相比于 PCA - Copula 的结果降低了 60.1 kg/hm²;平均相对误差为 5.42%,较 PCA - Copula 方法降低了 1.02%,表明 KPCA - Copula 方法确定的模型估测精度较高,将其进一步用于分析 2008—2013 年 4 个市估测单产与实际单产间的均方根误差(图 2),可见,基于 KPCA - Copula 方法,各市的均方根误差均

小于 $420 \text{ kg}/\text{hm}^2$, 且除宝鸡市外, 其余各市的均方根误差较 PCA - Copula 的结果均减小, 表明 KPCA - Copula 方法可更好地用于开展关中平原干旱对冬小麦生产及其产量的影响评估研究, 这从另一角度说明了 KPCA 对 PCA 方法的改进起到了较好的效果。

表 2 KPCA - Copula 与 PCA - Copula 方法确定的综合 VTCI 与单产之间回归模型的对比

Tab. 2 Comparison of linear regression models between wheat yields and comprehensive values of VTCI determined by KPCA - Copula and PCA - Copula methods

方法	回归模型	决定系数 R^2	显著性 p 值
KPCA - Copula	$y = 1796.3x + 3746.1$	0.608	< 0.001
PCA - Copula	$y = 3065.9x + 3960.0$	0.490	< 0.01

注: x 表示综合 VTCI 值; y 表示单产, kg/hm^2 。

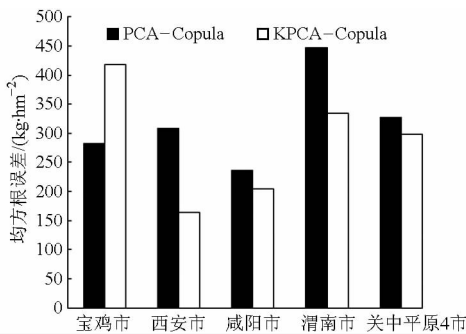


图 2 基于 KPCA - Copula 和 PCA - Copula 估产结果的精度比较

Fig. 2 Comparison of accuracies of estimated yields by using KPCA - Copula and PCA - Copula methods

2.2.3 综合 VTCI 的干旱特征分析

将基于 KPCA 降维获取的综合 VTCI 与采用 PCA 降维获取的综合 VTCI 进行比较(图 3), 可以看出, 基于 KPCA 方法降维的计算结果与 PCA 方法的结果分布趋势大体一致, 宝鸡市、西安市和咸阳市 2 种方法的综合 VTCI 的最大值和最小值对应的时间基本一致, 即综合 VTCI 均在 2010 年达到最大, 而在 2013 年综合 VTCI 最小。结合陕西省历年气候条件和小麦单产统计资料, 2010 年气象条件属冬小麦适宜年份, 产量为丰产年份; 2013 年的干旱程度严重, 不利于冬小麦生长, 对应的单产均低于其他各年, 进一步表明基于 KPCA 与 PCA 2 种降维方法获取的综合 VTCI 与该 3 市区的实际干旱特征均较相符。

对于渭南市 2 种方法确定的综合 VTCI 峰值对应的时间存在差异, 即基于 PCA - Copula 的综合 VTCI 峰值出现在 2011 年, 而基于 KPCA - Copula 的峰值对应 2010 年, 结合陕西省统计局发布的 2008—2013 年关中平原各市冬小麦的单产数据, 渭南市 2010 年的单产最高, 干旱程度较低, 意味着该

市的综合 VTCI 应在 2010 年达到峰值, 表明基于 KPCA - Copula 获取的综合 VTCI 在关中平原整体及局部均更符合研究区域的实际干旱特征, 也验证了 KPCA 方法的干旱影响评估结果较 PCA 方法更准确。

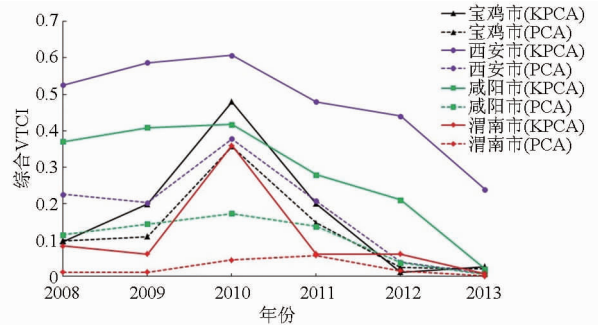


图 3 KPCA - Copula 和 PCA - Copula 2 种方法确定的关中平原 2008—2013 冬小麦综合 VTCI 比较

Fig. 3 Comparison of comprehensive VTCIs of winter wheat during 2008—2013 in Guanzhong Plain determined by KPCA - Copula and PCA - Copula methods

3 讨论

采用 KPCA 方法对冬小麦主要生育期 VTCI 数据进行降维处理, 获取主成分因子, 应用 Copula 函数确定的综合 VTCI 较 PCA - Copula 的结果更符合研究区域的干旱情况。这主要是由于两者对数据的降维处理方式不同, PCA 的本质是通过线性变换将高维数据向低维空间投影进行降维, 且得到的各个主成分反映的只是数据集中变量之间的线性相关性, 而许多数据集变量之间存在的关系往往是非线性的, 如果采用 PCA 降维可能会造成数据中的非线性高维特征信息的丢失。而 KPCA 是在 PCA 基础上, 对 PCA 进行改进以处理非线性问题, 即通过引入核函数将数据非线性映射到高维特征空间, 使得数据在高维特征空间中线性可分, 再进行主成分分析, 且 KPCA 方法只需计算用作内积的核函数, 无需知道非线性映射函数的具体形式, 也无需计算非线性变换, 能够克服一般映射方法中非线性函数结构及其参数值的确定以及特征空间维数的限制, 不仅能实现数据降维, 而且能用核方法有效地处理各生育期 VTCI 的非线性影响, 克服 PCA 方法难以反映数据中变量之间的非线性关系, 计算结果更客观, 且更符合干旱的非线性本质。

采用 KPCA 方法对 VTCI 数据进行降维处理, 结果表明, 第 1 个主成分的贡献率接近 80%, 意味着该主成分能够表达大部分原始变量信息, 且前 2 个主成分的累计贡献率高达 98%, 相当于 PCA 获取的前 3 个主成分表达的信息量, 与 PCA 相比, KPCA 不

仅能够抽取非线性特征,而且具有更优的识别性能,降维效果更佳。通过优选各主成分因子的边缘分布,进而利用 Copula 函数对于互相独立的变量无需求解参数的优越性,构建综合 VTCl 与单产之间的回归模型,结果表明,KPCA - Copula 方法较 PCA - Copula 方法能够更好地评估关中平原干旱对冬小麦生产及其产量的影响。

4 结论

(1)采用 KPCA 方法获取 2008—2013 年关中平原 4 市主要生育期 VTCl 的主成分因子,并将其与 PCA 方法的结果进行对比,结果表明,KPCA 能更有效地处理干旱的非线性特征,其对应的第 1 个主成分的贡献率明显高于 PCA,且前 2 个主成分的累计贡献率相当于 PCA 方法对应的前 3 个主成分,相比

于 PCA 需要提取前 3 个主成分构建干旱影响评估模型,KPCA 仅需选择前 2 个主成分,降维效果更佳。

(2)基于 KPCA - Copula 获取的综合 VTCl,建立其与冬小麦单产之间的回归模型,并将回归结果与 PCA - Copula 的结果进行对比,结果表明,采用 KPCA - Copula 方法确定的综合 VTCl 与单产之间回归模型的决定系数达到 0.608,较 PCA - Copula 的结果提高了 0.118,且对应模型的估测单产与实测单产间的均方根误差与平均相对误差均低于 PCA - Copula 的结果,精度较高,说明 KPCA - Copula 方法能够更好地基于干旱的非线性特征评估其对冬小麦产量的影响,且该方法确定的综合 VTCl 更符合关中平原实际的干旱特征。

参 考 文 献

- 郭毅,赵景波. 1368—1948 年陇中地区干旱灾害事件序列分形特征研究[J]. 地球科学进展, 2010, 25(6): 630 - 637.
GUO Yi, ZHAO Jingbo. Characteristics of fractal dimension of drought disaster sequences from 1368 to 1948 in Longzhong area [J]. Advances in Earth Science, 2010, 25(6): 630 - 637. (in Chinese)
- ZHANG Y W. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM[J]. Chemical Engineering Science, 2009, 64(5): 801 - 811.
- 王霞,王占岐,金贵,等. 基于核函数支持向量回归机的耕地面积预测[J]. 农业工程学报, 2014, 30(4): 204 - 211.
WANG Xia, WANG Zhanqi, JIN Gui, et al. Land reserve prediction using different kernel based support vector regression[J]. Transactions of the CSAE, 2014, 30(4): 204 - 211. (in Chinese)
- CUI P L, LI J H, WANG G Z. Improved kernel principal component analysis for fault detection[J]. Expert Systems with Applications, 2008, 34(2): 1210 - 1219.
- ZHONG S P, CHEN D Y, XU Q F, et al. Optimizing the Gaussian kernel function with the formulated kernel target alignment criterion for two-class pattern classification[J]. Pattern Recognition, 2013, 46(7): 2045 - 2054.
- LIU Q S, JIN H L, TANG X O, et al. A new extension of kernel feature and its application for visual recognition[J]. Neurocomputing, 2008, 71(10 - 12): 1850 - 1856.
- 石礼娟,谢彪彪,谢新港,等. 基于组合核函数的籼稻重度不宜存监测模型[J]. 农业机械学报, 2013, 44(8): 165 - 168.
SHI Lijuan, XIE Biaobiao, XIE Xin'gang, et al. Identification model of severely unstorable indica paddy based on combined kernel function[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(8): 165 - 168. (in Chinese)
- LIU C J. Gabor-based kernel PCA with fractional power polynomial models for face recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 572 - 581.
- LI Z, KRUGER U, XIE L, et al. Adaptive KPCA modeling of nonlinear systems[J]. IEEE Transactions on Signal Processing, 2015, 63(9): 2364 - 2376.
- 冯永玖,刘妙龙,童小华,等. 基于核主成分元胞模型的城市演变重建与预测[J]. 地理学报, 2010, 65(6): 665 - 675.
FENG Yongjiu, LIU Miaolong, TONG Xiaohua, et al. Kernel principal components analysis based cellular model for restructuring and predicting urban evolution[J]. Acta Geographica Sinica, 2010, 65(6): 665 - 675. (in Chinese)
- HUANG J, YAN X F. Related and independent variable fault detection based on KPCA and SVDD[J]. Journal of Process Control, 2016, 39: 88 - 99.
- 彭令,牛瑞卿,赵艳南,等. 基于核主成分分析和粒子群优化支持向量机的滑坡位移预测[J]. 武汉大学学报: 信息科学版, 2013, 38(2): 148 - 152.
PENG Ling, NIU Ruiqing, ZHAO Yannan, et al. Prediction of landslide displacement based on KPCA and PSO - SVR[J]. Geomatics and Information Science of Wuhan University, 2013, 38(2): 148 - 152. (in Chinese)
- PAN D W, LIU D T, ZHOU J, et al. Anomaly detection for satellite power subsystem with associated rules based on kernel principal component analysis[J]. Microelectronics Reliability, 2015, 55(9 - 10): 2082 - 2086.
- 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.
WANG Pengxin, GONG Jianya, LI Xiaowen. Vegetation temperature condition index and its application for drought monitoring [J]. Geomatics and Information Science of Wuhan University, 2001, 26(5): 412 - 418. (in Chinese)
- 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.
 WANG Pengxin, WAN Zhengming, GONG Jianya, et al. Advances in drought monitoring by using remotely sensed normalized difference vegetation index and land surface temperature products[J]. *Advances in Earth Science*, 2003, 18(4): 527 – 533. (in Chinese)
- 16 田苗, 王鹏新, 韩萍, 等. 基于 SARIMA 模型和条件植被温度指数的干旱预测[J]. *农业机械学报*, 2013, 44(2): 109 – 116.
 TIAN Miao, WANG Pengxin, HAN Ping, et al. Drought forecasts based on SARIMA models and vegetation temperature condition index[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(2): 109 – 116. (in Chinese)
- 17 李艳, 王鹏新, 刘峻明, 等. 基于条件植被温度指数的冬小麦主要生育期干旱监测效果评价——II. 改进层次分析法和变异系数法组合赋权[J]. *干旱地区农业研究*, 2014, 32(1): 236 – 239, 262.
 LI Yan, WANG Pengxin, LIU Junming, et al. Application of temperature condition index to evaluate the drought monitoring effect in main growing period of winter wheat: II. Improved analytic hierarchy process and variation coefficient method[J]. *Agricultural Research in the Arid Areas*, 2014, 32(1): 236 – 239, 262. (in Chinese)
- 18 王鹏新, 冯明悦, 孙辉涛, 等. 基于主成分分析和 Copula 函数的干旱影响评估研究[J]. *农业机械学报*, 2016, 47(9): 334 – 340.
 WANG Pengxin, FENG Mingyue, SUN Huitao, et al. Drought impact assessment based on principle component analysis and Copula function[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(9): 334 – 340. (in Chinese)
- 19 孙威, 王鹏新, 韩丽娟, 等. 条件植被温度指数干旱监测方法的完善[J]. *农业工程学报*, 2006, 22(2): 22 – 26.
 SUN Wei, WANG Pengxin, HAN Lijuan, et al. Further improvement of the approach to monitoring drought using vegetation and temperature condition indexes from multi-years' remotely sensed data[J]. *Transactions of the CSAE*, 2006, 22(2): 22 – 26. (in Chinese)
- 20 SCHOLKOPF B, SMOLA A, MULLER K. Nonlinear component analysis as a kernel eigenvalue problem [J]. *Neural Computation*, 1998, 10(5): 1299 – 1319.
- 21 殷勇, 薛俊莉, 于慧春, 等. 基于 KFDA 的食醋电子鼻鉴别方法[J]. *农业机械学报*, 2014, 45(9): 236 – 240.
 YIN Yong, XUE Junli, YU Huichun, et al. Identification method of electronic nose based on KFDA for different vinegar samples [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45(9): 236 – 240. (in Chinese)
- 22 ROMERO A, GATTA C, CAMPS V G. Unsupervised deep feature extraction for remote sensing image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(3): 1349 – 1362.
- 23 GOUBANOVA K, LI L. Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations[J]. *Global and Planetary Change*, 2007, 57(1 – 2): 27 – 42.

~~~~~

(上接第 357 页)

- 11 周勇钧, 俞守华, 区晶莹. 多特征 Camshift 和 Kalman 滤波结合的猪只智能跟踪[J]. *广东农业科学*, 2013(9): 174 – 177, 188.  
 ZHOU Yongjun, YU Shouhua, OU Jingying. Pig intelligent tracking based on multi-feature Camshift algorithm combined with Kalman filter[J]. *Guangdong Agricultural Sciences*, 2013(9): 174 – 177, 188. (in Chinese)
- 12 YU Shouhua, CHEN Zicheng, OU Jingying. Tracking algorithm based on multi-feature detection and target association of pigs on large-scale pig farms[J]. *Journal of Information & Computational Science*, 2015, 12(10): 3837 – 3844.
- 13 周丽萍, 陈志, 陈达, 等. 基于改进 Otsu 的生猪耳根特征区域检测方法[J]. *农业机械学报*, 2016, 47(4): 1 – 7.  
 ZHOU Liping, CHEN Zhi, CHEN Da, et al. Pig ear root detection based on adapted Otsu[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(4): 1 – 7. (in Chinese)
- 14 OHTSU N. A threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems Man & Cybernetics*, 1979, 9(1): 62 – 66.
- 15 KASHIHUA M, BAHR C, OTT S. Automatic monitoring of pig locomotion using image analysis[J]. *Livestock Science*, 2014, 159: 141 – 148.
- 16 段玉瑶, 马丽, 刘刚. 猪舍场景下的生猪目标跟踪和行为检测方法研究[J]. *农业机械学报*, 2015, 46(增刊): 187 – 193.  
 DUAN Yuyao, MA Li, LIU Gang. Target tracking and behavior detection method in piggery scenarios [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2015, 46(Supp.): 187 – 193. (in Chinese)
- 17 FITZGIBBON A W, FISHER R B. A buyer's guide to conic fitting[C] // *Proceedings of the 6th British Conference on Machine Vision*, 1995, 2: 513 – 522.
- 18 刘龙申, 沈明霞, 姚文, 等. 基于加速度传感器的母猪产前行为特征采集与分析[J]. *农业机械学报*, 2013, 44(3): 192 – 196, 191.  
 LIU Shenlong, SHEN Mingxia, YAO Wen, et al. Acquisition and analysis of sows' behavior before farrowing based on acceleration sensor[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(3): 192 – 196, 191. (in Chinese)