

MODIS 水稻面积提取中独立成分端元丰度校正方法

刘建红, 朱文泉^{*}, 孙冠楠, 张浚哲, 姜楠

(北京师范大学资源学院, 地表过程与资源生态国家重点实验室, 北京 100875)

摘要: 为了解决独立成分分析中端元丰度校正结果同实际丰度相差较大的问题, 该文提出了一种基于回归分析的独立成分端元丰度校正方法。具体是: 首先应用 ICA 对遥感时序数据进行分解, 获取目标地物的 ICA 分解结果; 再抽选一定量的样本, 将样本目标地物的真实丰度与 ICA 分解结果进行回归; 最后根据回归关系推算每个像元的目标地物丰度。基于 MODIS 时序数据, 将该文方法和线性拉伸方法应用于江苏兴化地区的水稻面积提取, 并将 2 种方法的提取结果同水稻真值图像进行对比。分析结果表明, 该文方法得到的水稻丰度图像的均方根误差、偏差在不同的空间尺度下均小于线性拉伸方法, 而不同空间尺度下的决定系数 (R^2) 均高于线性拉伸方法。与线性拉伸方法相比, 该文方法能获得更接近实际情况的端元丰度校正结果, 增强了 ICA 在农作物面积提取中的应用能力, 为大尺度农作物识别和面积提取提供了依据。

关键词: 遥感, 时间序列分析, 独立成分分析, 生长曲线, 回归, 水稻

doi: 10.3969/j.issn.1002-6819.2012.09.017

中图分类号: TP722

文献标志码: A

文章编号: 1002-6819(2012)-09-0103-06

刘建红, 朱文泉, 孙冠楠, 等. MODIS 水稻面积提取中独立成分端元丰度校正方法 [J]. 农业工程学报, 2012, 28(9): 103-108.

Liu Jianhong, Zhu Wenquan, Sun Guannan, et al. Endmember abundance calibration method for paddy rice area extraction from MODIS data based on independent component analysis[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2012, 28(9): 103-108. (in Chinese with English abstract)

0 引言

水稻是中国最重要的粮食作物之一。中国水稻产量居世界之首, 2010 年为 1.93 亿 t, 约占世界水稻产量的 30%^[1]。除青海外, 全国各省都有水稻分布。根据 2006 年统计数据, 中国水稻种植面积占世界水稻种植总面积的 18.5%^[2]。水稻生产不仅关系到中国和世界的粮食安全, 还对环境和气候变化、经济和政治决策、水资源管理等产生重要影响^[2-3]。

低分辨率遥感时间序列数据 (time-series coarse resolution data, TCRD) 计算的植被指数 (vegetation index, VI) 可以构成反映农作物动态变化的生长曲线, 从而为农作物识别及其面积提取提供了有利条件。国内外学者发展了许多大尺度农作物提取模型, 这些模型基本上都是根据农作物生长曲线的不同特征来设定阈值进行识别^[4-7]。这些模型依赖于多步骤的阈值判断, 阈值的高低和 VI 曲线上的噪声对农作物识别结果影响大。由于农作物受各地的播种时间、物候差异等影响, 阈值具有明显的区域

性特点, 从而降低了这类模型的通用性。受空间分辨率限制, 一个 TCRD 像元可能包含多种地物类型, 尤其是在土地覆盖类型分布结构和农作物种植结构复杂的地区, 农作物分类存在较大误差。有学者认为基于像元的分类适合研究大尺度作物分布趋势, 而区域内不同农作物识别则依赖于混合像元分解^[8]。Lobell 等^[9]应用扩展的线性分解模型进行与农作物相关的耕地分类研究, 结果表明模型精度取决于农作物类型和研究的尺度, 刘姝娣等^[10]应用线性光谱混合模型提高了棉花识别的精度。Rembold 等^[11]采用光谱角制图进行混合像元分解获取了研究区的冬小麦面积, 结果表明冬小麦识别精度受数据源和检验数据质量的影响; Verbeiren 等^[12]通过对区域尺度的农作物面积提取试验发现神经网络分解方法优于线性分解方法。以上都是监督式的混合像元分解方法, 这种方法面临的最主要问题是: 端元数量和端元光谱 (本文是生长曲线) 需要预先确定。为了有效提取一种农作物, 有时需要确定研究区所有主要的农作物类型及其生长曲线。由于现有农作物标准生长曲线数据库不够健全、农作物生长曲线的时空差异较大等因素的影响^[13], 以农作物生长曲线为先验知识的监督式混合像元分解方法难以被有效应用。

独立成分分析 (independent component analysis, ICA) 能够在很少先验知识的情况下实现端元光谱和丰度的同时提取, 因此近年来成为高光谱遥感领域研究的热点^[3,14]。然而 ICA 要求各成分之间相互统计独立且为非高斯分布, 导致其结果是非约束的, 不能直接作为端元的丰度^[3]。Wang 等^[15]提出了一种基于最大值-最小值标准化的线性

收稿日期: 2011-07-07 修订日期: 2011-10-31

基金项目: 国家高技术研究与发展计划项目 (2006AA120101), 国家自然科学基金资助项目 (40871194)

作者简介: 刘建红 (1985-), 女, 江西玉山人, 博士生, 研究方向为资源与环境遥感。北京 北京师范大学资源学院地表过程与资源生态国家重点实验室, 100875。Email: jhliu_85@163.com

*通信作者: 朱文泉 (1975-), 男, 湖南永兴人, 北京师范大学资源学院副教授, 博士, 主要从事植被生态参数遥感反演、遥感数据处理等方面的研究。北京 北京师范大学资源学院地表过程与资源生态国家重点实验室, 100875。Email: zhuwq75@bnu.edu.cn

拉伸方法（以下称为线性拉伸方法），拉伸结果可以直接作为端元的丰度。这种方法得到的结果只能近似反映目标地物的丰度，但与真实丰度还有较大差异。

为提高地物丰度的提取精度，本文提出了一种独立成分端元丰度校正的新方法，通过加入先验知识改进端元丰度校正过程。具体是：1) 进行 ICA 分解，获取目标地物的 ICA 分解结果；2) 将样本目标地物的真实丰度与 ICA 分解结果回归，获取 ICA 分解结果与目标地物真实丰度之间的关系；3) 将这个回归关系应用到整个研究区从而得到整个研究区目标地物的丰度。以期为大尺度农作物识别和面积提取提供参考。

1 研究方法

本文的技术路线如图 1 所示。首先，应用 ICA 对低分辨率遥感时间序列数据进行分解以获取目标地物独立成分；再将目标地物的中分辨率遥感识别结果作为准真值；然后抽选一定量的样本，将样本目标地物的实际丰度与 ICA 分解结果进行回归，进而推算每个像元的目标地物丰度。为了进一步阐明本文方法的应用效果及优势，本文还应用线性拉伸方法提取水稻面积。最后，对 2 种方法的提取结果进行对比分析。

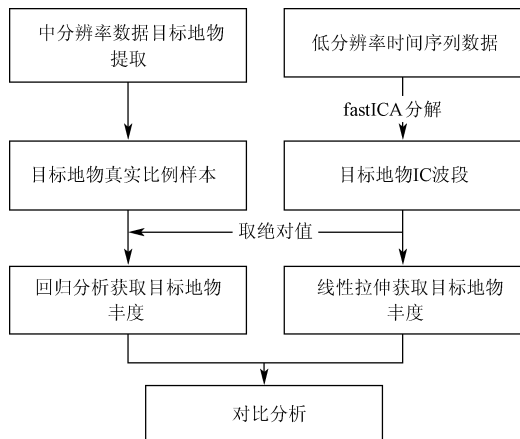


图 1 研究技术路线图
Fig.1 Flowchart of research method

1.1 ICA 模型

ICA 是近几年才发展起来的一种基于信号高阶统计特性的分析方法，其目的是将观测数据进行某种线性分解，得到统计独立的成分^[16-18]。ICA 已经成功地应用于盲信号分离。设观测数据 $X=\{x_1, x_2, \dots, x_n\}$ 是源信号 $S=\{s_1, s_2, \dots, s_m\}$ 的观测值，假设第 i 个观测信号是由 n 个独立分量线性混合而成^[19]

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \dots + a_{im}s_m \quad (i = 1, 2, \dots, n) \quad (1)$$

用矢量表示式 (1)，则有

$$X=AS \quad (2)$$

也可以写成

$$X = \sum_{j=1}^m a_j s_j \quad (3)$$

式 (1) ~ (3) 中， x_i 表示第 i 个观测信号， s_m 表

示第 m 个源信号， a_{im} 表示第 i 个观测信号中包含第 m 个源信号的比例， $A=\{a_1, a_2, \dots, a_m\}$ 为混合矩阵， a_i 为混合矩阵基向量。ICA 就是仅通过观测数据 X 估计出未知独立源 s_j 或估计混合矩阵 A 。ICA 在解决信号源分解问题时，一般假设源信号之间统计独立、至多只有一个源信号服从高斯分布及各源信号具有零均值。这样的 ICA 模型存在 2 个不确定性：一是不能确定独立成分的顺序，二是不能确定独立成分的方差^[19]。

ICA 实现算法包括最大似然估计法、最大熵方法、最小互信息、最大负熵法等。这些算法都是利用分离后的各成分之间的最大独立性来建立对照函数，通过寻找一种迭代算法来进行信号的分离。快速 ICA 算法 (FastICA) 是一种基于负熵最大化的定点算法，每次从混合信号中分离出一个独立成分，具有计算简单、收敛快等特点，被广泛用于遥感图像分析^[20]。FastICA 采用峭度 (Kurtosis) 对负熵做近似计算^[21,22]，以此作为随机变量的非高斯性度量

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (4)$$

式中 $kurt(y)$ 表示随机变量 y 的峭度， $E\{y^n\}$ 表示 n 阶中心矩。对于一个高斯分布的随机变量，它的峭度等于 0；而大多数非高斯随机变量的峭度不等于 0。由于峭度有正有负，一般利用峭度的绝对值或平方作为非高斯性度量，值为 0 的是高斯变量，大于 0 的是非高斯变量。

1.2 端元确定及丰度校正

由于独立成分必须满足统计独立性，因此 ICA 得到的结果是非约束的，不能直接作为端元丰度。Wang 等^[15] 提出了一种端元提取及丰度校正方法，使 ICA 能够获得相对确定的结果。具体是：用 $IC_i(r)$ 表示第 i 个独立成份 (IC) 波段上像元 r 的值，那么第 i 个 IC 波段的端元 e_i 定义为第 i 个 IC 波段的绝对值最大的像元，其光谱作为端元光谱

$$|e_i| = \max_r |IC_i(r)| \quad (5)$$

然后根据 $|e_i|$ 采用最小值-最大值标准化准则对 ICA 分解结果进行线性拉伸

$$\alpha_{IC_i(r)} = \frac{|IC_i(r)| - \min_r |IC_i(r)|}{|e_i| - \min_r |IC_i(r)|} \quad (6)$$

式中， $\alpha_{IC_i(r)}$ 表示第 i 个 IC 波段上像元 r 的丰度； $|IC_i(r)|$ 表示第 i 个 IC 波段上像元 r 的绝对值，拉伸后的数值作为目标地物的丰度。

1.3 样本支持下的端元丰度校正

尽管线性拉伸方法得到的丰度可以近似表示目标地物的分布，但还不能直接作为每个像元内目标地物的真实比例。本文方法的思路是，采用一定量的样本，将样本目标地物的真实丰度与 ICA 分解结果进行回归，获取与像元内目标地物真实比例更接近的结果。

1.4 精度评价

为对本文方法和线性拉伸方法的提取结果进行评价，本文采用了以下 3 个统计指标：均方根误差 (RMSE)、偏差 (Bias) 和决定系数 (R^2)。偏差是指误差的平均值，

用于指示过高或过低估计。 $RMSE$ 是估计偏差与方差相结合的表达式,用于衡量方法估算的精度。

$$RMSE_{(T)} = \sqrt{\sum_{i=1}^N (\hat{y}_i - y_i)^2 / N} \quad (7)$$

$$Bias_{(T)} = \sum_{i=1}^N (\hat{y}_i - y_i) / N \quad (8)$$

式中, T 表示窗口大小, N 表示窗口 T 下的子区个数, y_i 和 \hat{y}_i 分别表示第 i 个子区水稻实际丰度和由 2 种方法计算得到的水稻丰度。决定系数用公式 (9) 计算

$$R^2 = \frac{\text{cov}(y, \hat{y})^2}{\text{var}(y) \text{var}(\hat{y})} \quad (9)$$

式中, $\text{cov}(y, \hat{y})^2$ 表示水稻丰度准真值图像与估算图像之间的协方差, $\text{var}(y)$ 和 $\text{var}(\hat{y})$ 分别表示水稻丰度准真值图像的方差和水稻丰度估算图像的方差。准真值图像与估算图像之间的 R^2 反映了 2 种方法能够获取水稻丰度空间分布的一致程度。

通过改变不同的空间尺度(窗口大小),可以推测出水稻提取精度与评价单元大小之间的关系。本文在对 2 种提取水稻面积的方法进行评估时,窗口大小从 1 像元×1 像元尺度一直到 10 像元×10 像元尺度。

2 研究区概况与数据获取

本文研究区位于江苏省中部,覆盖兴化市及周边地区(119°30'29"~120°21'16"E, 32°35'45"~33°15'8"N),面积为 4 442 km²。区域内以冬小麦-水稻的种植模式为主,主要夏粮作物为水稻,同期作物有夏玉米、棉花、大豆、甘薯及花生等。

本文用于提取水稻面积的 TCRD 数据是 8 天合成的陆地表面反射率数据集 MODIS09A1(<https://lpdaac.usgs.gov/>),空间分辨率为 500 m。该数据集已经过大气校正,并根据观测覆盖、视角、云及阴影等条件选择 8d 之中的最佳观测值进行合成。MODIS09A1 包含了可见光、近红外和短波红外等 7 个波段。用于本文分析的增强型植被指数(enhanced vegetation index, EVI)数据是根据 2010 年 5 月 1 日到 2010 年 11 月 1 日反射率数据计算得到的,共 23 个时相。在这个时段内,水稻经历 3 个不同阶段:一是灌水移栽期,二是水稻生长期,三是水稻成熟及收割期,水田的植被覆盖和含水率也会随之变化^[4]。

为了检验基于 MODIS 数据的水稻面积提取精度,试验还用到了该研究区 2010 年 7 月 31 日的 SPOT4 多光谱影像,该影像的空间分辨率为 20 m。对 SPOT 影像进行几何校正、投影转换等预处理后,对其进行水稻面积提取。水稻面积提取流程包括:结合 2010 年 9 月野外调查拍摄的 GPS 点的照片进行目视判断,建立居民地、荷塘、水稻、旱地作物、水体共 5 种土地覆盖类型的解译标志;在 ENVI 软件中提取训练样本并采用自动阈值的分类回归树(classification and regression tree, CART)算法分类获取水稻面积;将水稻面积提取结果重采样到 500 m,作为水稻丰度的准真值。国外已有研究将未经验证的

LANDSAT 分类结果作为真值来检验 MODIS 分类结果^[9,11],由于本文在对 SPOT 图像分类时已经参考了野外调查记录,且相对于 MODIS 数据,SPOT 图像分辨率足够高,因此本文并未对 SPOT 分类结果进行野外验证。

在对 ICA 提取结果进行回归分析之前,需要选择样本。为了得到对全局的无偏估计,每个丰度区间都应该有样本分布。因此,在水稻丰度准真值图像上按照水稻丰度把所有像元划为 10 个区间,分别是 ≤10%, >10%~20%, >20%~30%, >30%~40%, >40%~50%, >50%~60%, >60%~70%, >70%~80%, >80%~90%, >90%~100%,然后在每个区间分别随机抽取 10 个样本(像元),总共得到 100 个样本用于后续分析。样本的真值即样本对应的水稻丰度准真值。

3 结果与分析

3.1 端元提取结果

在 ENVI 4.7 下对 MODIS EVI 进行 ICA 分析,选择了特征向量占 95% 的前 8 个独立成分。对每一独立成分的时间曲线与水稻的 EVI 曲线进行对比,由此确定第一独立成分(IC_1)反映的是水稻丰度信息(图 2)。

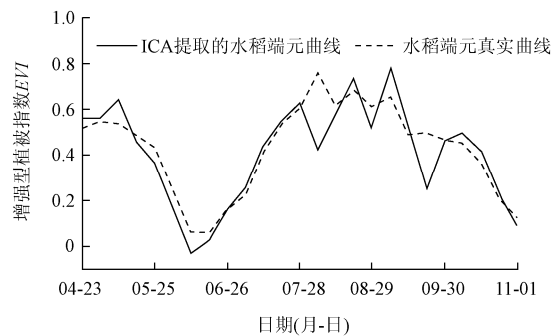


图 2 ICA 提取的水稻端元曲线与真实的水稻端元曲线
Fig.2 Observed and modeled EVI time series for paddy rice endmember

从图 2 中可以看出,4 月下旬随冬小麦、油菜收割以及水田灌水, EVI 迅速降低;6 月上中旬,秧苗移栽后开始生长, EVI 又逐渐升高。8 月上旬到 9 月中旬是水稻生长期,由于研究区地处北亚热带湿润气候区,夏季降水多,造成水稻生长曲线的波动。9 月中旬后,水稻成熟收割, EVI 逐渐降低。

3.2 水稻面积提取结果

确定水稻独立成分图像后,将选中的 100 个样本的水稻丰度准真值与其对应的 IC_1 绝对值进行回归(图 3),然后再根据回归公式,对 IC_1 的绝对值进行计算,得到结果如图 4b 所示。

对 IC_1 采用线性拉伸方法处理,得到水稻丰度分布初步结果见图 4c。相对于水稻丰度准真值图像(图 4a),线性拉伸的结果总体上高估了研究区的水稻丰度,而本文方法的估算结果在水稻低覆盖区和无覆盖区与实际情况极为接近,并且在水稻高覆盖区的分布趋势也与水稻的真实分布相符。3 种丰度图像的直方图也与这一结论吻合(图 5)。

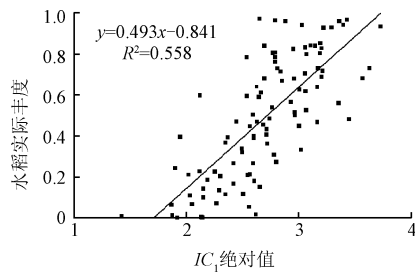
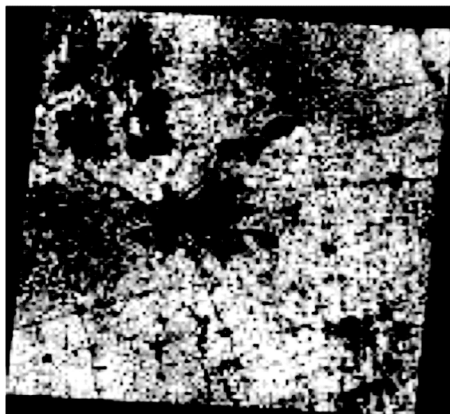
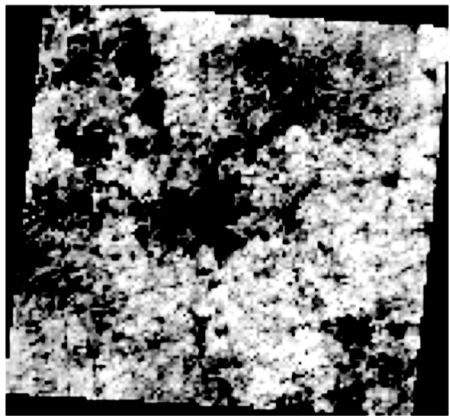


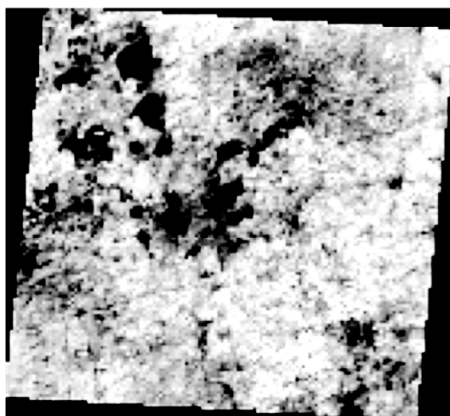
图3 样本 IC_1 绝对值与水稻实际丰度之间的散点图
Fig.3 Scatter plot between absolute value of IC_1 and actual rice abundance for 100 samples



a. 水稻丰度准真值图像



b. 本文方法得到的水稻丰度图像



c. 线性拉伸方法得到的水稻丰度图像

图4 3种水稻丰度图像的比较

Fig.4 Comparison among three rice abundance maps

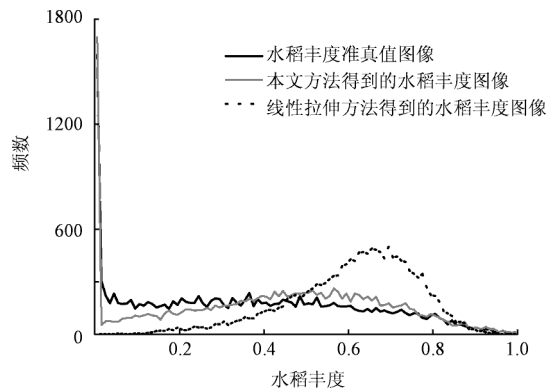


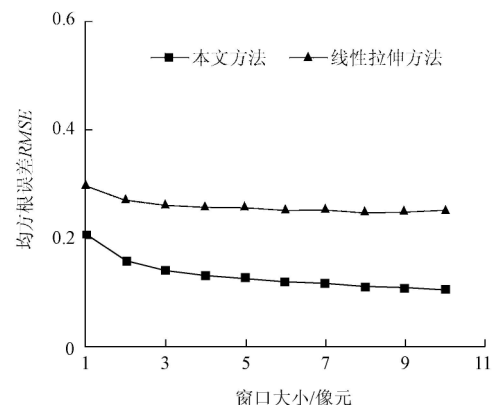
图5 3种水稻丰度图像的直方图对比

Fig.5 Histogram of three rice abundance maps

线性拉伸方法得到的水稻丰度图像的直方图接近高斯分布, 绝大部分像元丰度在 0.4~0.8 之间, 丰度低于 0.1 的像元很少。而水稻丰度准真值图像和本文方法得到的水稻丰度图像都是非高斯分布, 后两者的直方图非常接近。水稻丰度在接近 0 处有高峰, 这主要是研究区中如城区、水域、荷塘等没有水稻分布的像元。除此之外, 在丰度 0.02~0.8 之间大致呈均匀分布, 丰度大于 0.8 的像元逐渐减少。

3.3 精度分析

图 6 显示了本文方法与线性拉伸方法在不同窗口下的均方根误差 $RMSE$ 、偏差 $Bias$ 、决定系数 R^2 。由图 6 可以看出, 2 种方法的 $RMSE$ 都随窗口增大而减小, 但本文方法的 $RMSE$ 在各种窗口下都比线性拉伸方法小。当窗口增大到 10 像元×10 像元时, 本文方法的 $RMSE$ 下降到 0.10, 比窗口为 1 像元×1 像元时降低了 0.1; 线性拉伸方法的 $RMSE$ 下降到 0.25, 比窗口为 1 像元×1 像元时降低了 0.05。2 种方法的偏差都比较稳定, 在不同窗口下几乎不变。本文方法的 $Bias$ 在 0.053~0.055 之间, 而线性拉伸方法的 $Bias$ 在 0.25~0.29 之间。2 种方法的 R^2 都呈现随窗口变大而增大的趋势, 本文方法的 R^2 在各窗口下都比线性拉伸方法的高, 但两者的变化趋势较为一致。在窗口为 1 像元×1 像元时, 本文方法的 R^2 为 0.54, 而线性拉伸方法为 0.48; 在窗口为 10 像元×10 像元时, 本文方法的 R^2 为 0.74, 而线性拉伸方法为 0.67。



a. 均方根误差 $RMSE$

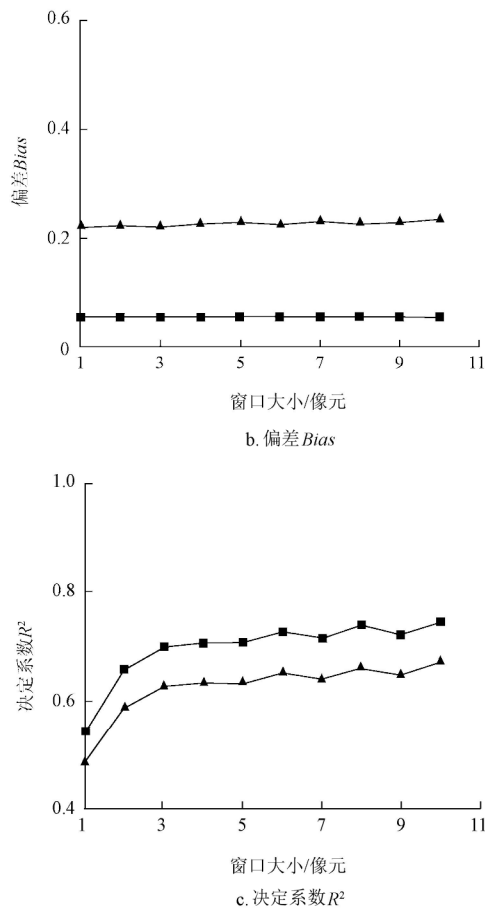


图6 本文方法与线性拉伸方法在不同窗口下的均方根误差 $RMSE$ 、偏差 $Bias$ 、决定系数 R^2

Fig.6 Root mean square error (top), $Bias$ (middle), and R^2 (bottom) for rice abundance derived from proposed method and linear scaling method at different window size

4 讨论

与线性拉伸方法相比,本文方法仅需少量先验知识(水稻的生长曲线及样本的水稻面积比例),就可以提高水稻面积的提取精度。由于农作物生长曲线可以通过先验知识建立,如已有的光谱库、纯净像元生长曲线;而样本农作物实际比例可以从野外调查数据、中/高分辨率遥感影像识别结果、土地利用调查数据、航拍图像等数据中得到,因此,本文方法可以很方便地应用于农作物面积提取。

然而,中国农作物种植结构复杂,以水稻为例,不仅有早稻、中稻和晚稻之分,有的水稻田还与其他作物和果蔬相间,这给农作物面积提取带来了很大的困难。本文提出的方法是否在各种条件下都能得到稳健的结果还有待进一步的试验佐证。

5 结论

该文提出了一种新的独立成分端元丰度校正方法,通过加入先验知识结合回归分析改进了端元丰度校正过程。将本文方法应用于江苏省兴化市的水稻面积提取,并同线性拉伸方法的提取结果进行比较。结果分析与精

度评价表明:本文方法提取结果的均方根误差 $RMSE$ 和偏差 $Bias$ 在各种空间尺度下均小于线性拉伸方法,而决定系数 R^2 在各种空间尺度下均高于线性拉伸方法。在 10 像元 \times 10 像元尺度下,该文提取结果的均方根误差 $RMSE$ 为 0.10, 偏差 $Bias$ 为 0.053, 决定系数 R^2 为 0.74。由此说明本文提出的基于回归分析的独立成分端元丰度校正方法在农作物面积提取中具有更好的应用能力。

[参考文献]

- [1] Faostat. FAO Statistical databases[Z]. Rome: 2011.
- [2] Zou Jianwen, Huang Yao, Zheng Xuhua, et al. Quantifying direct N_2O emissions in paddy fields during rice growing season in mainland China: Dependence on water regime[J]. Atmospheric Environment, 2007, 41(37): 8030–8042.
- [3] Ozdogan M. The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis[J]. Remote Sensing of Environment, 2010, 114(6): 1190–1204.
- [4] Xiao Xiangming, Boles Stephen, Liu Jiyuan, et al. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images[J]. Remote Sensing of Environment, 2005, 95(4): 480–492.
- [5] 黄青, 唐华俊, 周清波, 等. 东北地区主要作物种植结构遥感提取及长势监测[J]. 农业工程学报, 2010, 26(9): 218–223.
- [6] 林文鹏, 王长耀, 储德平, 等. 基于光谱特征分析的主要秋季作物类型提取研究[J]. 农业工程学报, 2006, 22(9): 128–132.
- [7] 彭代亮. 基于统计与 MODIS 数据的水稻遥感估产方法研究[D]. 杭州: 浙江大学, 2009.
- [8] Peng Dailiang. The Study on the Method of Rice Yield Estimation Using Statistical and MODIS data[D]. Hangzhou: Zhejiang University, 2009. (in Chinese with English abstract)
- [9] Galford G L, Mustard J F, Melillo J, et al. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil[J]. Remote Sensing of Environment, 2008, 112(2): 576–587.
- [10] Lobell D B, Asner G P. Cropland distributions from temporal unmixing of MODIS data[J]. Remote Sensing of Environment, 2004, 93(3): 412–422. (in Chinese with English abstract)
- [11] 刘姣娣, 曹卫彬, 刘学, 等. 棉花遥感识别的混合像元分解[J]. 农业工程学报, 2011, 27(6): 182–186.
- [12] Liu Jiaodi, Cao Weibin, Liu Xue, et al. Decomposition of mixed pixel for cotton identification using remote sensing data[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(6): 182–186. (in Chinese with English abstract)

- [11] Rembold F, Maselli F. Estimating inter-annual crop area variation using multi-resolution satellite sensor images[J]. Photogrammetric Engineering and Remote Sensing, 2006, 72(1): 55—62.
- [12] Verbeiren S, Eerens H, Piccard I, et al. Sub-pixel classification of SPOT-VEGETATION time series for the assessment of regional crop areas in Belgium[J]. International Journal of Applied Earth Observation and Geoinformation, 2008, 10(4): 486—497.
- [13] 赵慧洁, 李娜, 贾国瑞, 等. 改进独立成分分析在高光谱图像分类中的应用[J]. 北京航空航天大学学报, 2006, 32(11): 1333—1336.
Zhao Huijie, Li Na, Jia Guorui, et al. Improved independent component analysis applied to classification hyperspectral imagery[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(11): 1333—1336. (in Chinese with English abstract)
- [14] 罗文斐, 钟亮, 张兵, 等. 高光谱遥感图像光谱解混的独立成分分析技术[J]. 光谱学与光谱分析, 2010, 30(6): 1628—1633.
Luo Wenfei, Zhong Liang, Zhang Bing, et al. Independent component analysis for spectral unmixing in hyperspectral remote sensing image[J]. Spectroscopy and Spectral Analysis, 2010, 30(6): 1628—1633. (in Chinese with English abstract)
- [15] Wang Jing, Chang Chein-I. Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(9): 2601—2616.
- [16] 杨竹青, 李勇, 胡德文. 独立成分分析方法综述[J]. 自动化学报, 2002, 28(5): 762—772.
Yang Zhuqing, Li Yong, Hu Dewen. Independent component analysis: A survey[J]. Acta Automatica Sinica, 2002, 28(5): 762—772. (in Chinese with English abstract)
- [17] 钟家强, 王润生. 基于独立成分分析的多时相遥感图像变化检测[J]. 电子与信息学报, 2006, 28(6): 994—998.
Zhong Jiaqiang, Wang Runsheng. Multitemporal remote sensing images change detection based on ICA[J]. Journal of Electronics and Information Technology, 2006, 28(6): 994—998. (in Chinese with English abstract)
- [18] De Lathauwer L, De Moor B, Vandewalle J. An introduction to independent component analysis[J]. Journal of Chemometrics, 2000, 14(3): 123—149.
- [19] Hyvärinen A, Oja E. Independent component analysis: algorithms and applications[J]. Neural Networks, 2000, 13(4/5): 411—430.
- [20] 王小敏. 独立分量分析及其在遥感动态监测中的应用研究[D]. 南京: 南京理工大学, 2008.
Wang Xiaomin. Independent Component Analysis and its Application in Remote Sensing Dynamic Monitoring[D]. Nanjing: Nanjing University of Science and Technology, 2008. (in Chinese with English abstract)
- [21] Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis[J]. IEEE Transactions on Neural Networks, 1999, 10(3): 1045—9227.
- [22] 郭武, 张鹏, 王润生. 独立分量分析及其在图像处理中的应用现状[J]. 计算机工程与应用, 2008, 44(23): 172—177.
Guo Wu, Zhang Peng, Wang Runsheng. Independent component analysis and its applications in image processing[J]. Computer Engineering and Applications, 2008, 44(23): 172—177. (in Chinese with English abstract)

Endmember abundance calibration method for paddy rice area extraction from MODIS data based on independent component analysis

Liu Jianhong, Zhu Wenquan^{*}, Sun Guannan, Zhang Junzhe, Jiang Nan

(State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Resources Science and Technology, Beijing Normal University, Beijing 100875, China)

Abstract: There is a large discrepancy between the actual abundance of land cover and the result derived from the endmember abundance calibration of Independent Component Analysis (ICA). In order to solve this problem, this paper proposed a new method for the endmember abundance calibration of ICA by combining regression analysis. The new method includes 3 steps: Firstly, decomposing the remote sensing time-series data to obtain the independent component of the object feature; Secondly, selecting a certain amount of samples with actual object feature abundance and then building the relationship between the actual abundance and derived independent component using regression analysis; Finally, using regression relationship to derive the abundance of object feature for each pixel. Based on the MODIS time-series data, the new method and the linear scaling method were applied in Xinghua county, Jiangsu province of China for the mapping of rice abundance. The results derived from these two methods were then compared with the actual rice abundance map of the study area. Results showed that the Root Mean Square Error (RMSE) and Bias of the rice map derived from the new method was all smaller than that by linear scaling method, where the determination coefficient (R^2) was all higher than that by linear scaling method at different spatial scales. The new method can enhance the application of ICA model in crop acreage mapping and provide a basis for large-scale crop identification and area extraction.

Key words: remote sensing, time series analysis, independent component analysis (ICA), growth cycle, regression, paddy rice