

# 藏北地区三种时序 NDVI 重建方法与应用分析

宋春桥<sup>1,2</sup>, 游松财<sup>3</sup>, 柯灵红<sup>1,2</sup>, 刘高焕<sup>1</sup>

(1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 北京 100101;

2. 中国科学院研究生院, 北京 100049; 3. 中国农业科学院农业环境与可持续发展研究所, 北京 100081)

**摘要:** 遥感植被指数时间序列数据集, 已广泛应用于陆地生态环境变化监测与模拟、植被覆盖动态变化分析、植被物候特征识别与信息提取等多方面的研究。但其因受遥感器采集与传输过程、大气条件、地面状况等诸多因素的影响, 时序 NDVI 数据包含各种噪声, 因此研究者们发展了一系列时间序列曲线重建方法。本文对近年来提出或改进的重建算法原理、优缺点进行阐述; 然后, 选择当前最为常用的 3 种方法, 即非对称高斯函数(AG)拟合、双 Logistic 曲线(D-L)拟合和 Savitzky - Golay(S-G)滤波法, 以藏北地区不同土地覆被类型样点像元 NDVI 时间序列为实例, 对算法的去噪效果、保真性能、生长峰值及细节处理效果等方面进行比较研究。结果表明, AG 与 D-L 拟合两种算法具有较好的一致性, 但对生长峰值模拟有所差异; 3 种方法对荒漠、荒漠草原、草原、灌丛、作物用地及林地等不同覆被类型各具优势, 表现出区域、覆被类型和应用目的差异性。最后, 基于 AG 拟合算法对整个藏北地区 2007—2009 年 MODIS NDVI 时间序列进行重建, 处理后像元 NDVI 空间格局异质性减弱。

**关键词:** MODIS NDVI; 时间序列重建; 藏北; 非对称高斯函数拟合; 双 Logistic 曲线拟合; Savitzky - Golay 滤波

**DOI:** 10.3724/SP.J.1047.2011.00133

## 1 引言

卫星传感器所获取的植被指数时间序列数据, 能够精确地反映陆地生态系统植被的生长状态和季相、年际变化特征。其已广泛应用于全球与区域生态环境变化监测与模拟、植被覆盖动态变化研究、植被物候特征识别与信息提取等许多方面<sup>[1-5]</sup>。目前, 常用的 NDVI 时间序列主要来自 NOAA AVHRR、SPOT VEGETATION 和 MODIS TERRA/AQUA 等传感器, 同时, ENVISAT/MERIS NDVI 时间序列数据集也逐渐用于研究。

NDVI 反映植被生长过程的时间序列曲线理论上是连续平滑的, 但由于传感器本身性能、数据传输过程失误, 太阳光照角度、观测视角、地物双向性反射以及云、大气气溶胶等观测条件因时间而异。此外, 因地表水、冰雪等随机干扰因素等, 因此, 观测值包含很多不可预测的噪声, 得到的 NDVI 时间序列呈锯齿状的不规则波动, 反映季节变化趋势不明显, 使陆地生态系统监测、植被覆盖变化分析和信息提取等研究结论引进了诸多干扰<sup>[6-7]</sup>。尽管常用的 NDVI 数据集广泛以最大值合成法(MVC)来

降低这些噪声, 但是一些假高值、合成期内连续的 NDVI 低值等噪声残差仍然存在, 为此发展了一系列 NDVI 时间序列重建的方法。

当前 NDVI 时间序列曲线重建的主要方法包括: 阈值去噪法、基于滤波函数的拟合法及曲线拟合法<sup>[8]</sup>。3 种方法既可以基于原始的 NDVI 时间序列曲线重建, 也可以对 MVC 处理后的数据集进行优化。阈值去噪法包括 Viovy 等提出的最佳指数斜率提取法(BISE)和 Lovell 等在此基础上提出的改进算法<sup>[9-10]</sup>; 基于滤波的拟合法主要包括 Savitzky - Golay 滤波法<sup>[11]</sup>、均值迭代滤波法(MVI)<sup>[12]</sup>、傅立叶变换<sup>[13]</sup>和时间序列谐波分析法(HANTS)<sup>[14]</sup>; 曲线拟合法以双 Logistic 函数拟合法<sup>[15]</sup>和非对称性高斯函数拟合法<sup>[16]</sup>为代表。以上基本涵盖了遥感 NDVI 时间序列曲线重建的主要方法。目前, 各种重建方法分别被国内外学者应用于全球不同区域的研究, 尚没有达成共识。

研究者在不断改进或者提出新的算法, 希望能够克服现有方法本身的适应性或缺陷, 有效滤除噪声、反映植被真实的生长状况和物候规律等。同时, 部分国内外学者选择多种或两种不同的重建算

收稿日期: 2010-06-29; 修回日期: 2010-10-22.

基金项目: 藏北高原地区土壤水分与土壤温度时空变化模拟分析(No. 40971132)。

作者简介: 宋春桥(1986-), 男, 湖南衡阳人, 硕士研究生, 研究方向: 遥感与 GIS 应用。chunqiao\_song@163.com

法进行了比较研究,分析各自的优劣<sup>[17-18]</sup>。如 Sakamot 等以日本为研究区,对小波变换和傅立叶变换方法进行了比较分析,结果表明,小波变换总体上要优于傅立叶变换<sup>[19]</sup>。Jennifer 选择加拿大阿尔伯塔中西部的洛矶山脉中的一个典型区域作为示范区,比较两种曲线拟合方法与四种滤波方法对 NDVI 时间序列重建效果,认为两种曲线拟合方法总体上优于 4 种滤波方法<sup>[20]</sup>。Beck 等在北半球高纬度的芬兰、瑞典和挪威的交界区域对双 Logistic 函数拟合法、非对称性高斯函数拟合法和傅立叶变换法进行了比较研究,结果认为,双 Logistic 曲线拟合法和非对称性高斯函数拟合法的处理效果较相似,二者都优于傅立叶变换<sup>[15]</sup>。这些比较研究大多针对国外某些区域和植被覆盖特征来分析方法的优劣,并且没有深入探究各种重建方法,对不同 NDVI 生长曲线特征以及重建后数据的应用目的(如年际 NDVI<sub>max</sub> 变化、生长季内 NDVI 累积值反应生物量、NDVI 生长季曲线提取物候信息等)的处理是否存在差异,因此,其研究结论对于其他研究区域或研究目的的重建方法选取需要进一步分析考虑。本人认为,这些重建方法基本不存在某种明显优于其他的算法,而是因不同的研究区域植被

覆盖类型特征和应用目的而异。不同的重建方法在 NDVI 时间序列曲线的上包络线、存在连续的噪声点的曲线、不存在明显季节性或者非单峰(一年两熟或多熟植被)的时间序列以及在返青期—生长期—休眠期的过渡等方面的拟合都各具优势。因此,本文以藏北草地几种主要的植被覆盖类型为研究对象,探究非对称性高斯函数拟合法、双 Logistic 函数拟合法和 Savitzky - Golay 滤波法三种主流方法对不同特征的 NDVI 时间曲线的重建效果,比较分析各方法在 NDVI 时间曲线不同时段的处理能力;同时,本文也为其他研究在选择 NDVI 重建方法过程中提供一种分析思路。

2 研究区域与数据源分析

(1)研究区域

本文以藏北地区植被 NDVI 时间序列重建作为研究对象,评价 3 种算法对于该区域不同类型植被覆盖的应用特点。藏北地区位于青藏高原腹地,素有“江河源”、“生态源”、“中华水塔”之美誉<sup>[21]</sup>,是我国长江、怒江、澜沧江等主要江河的发源地(图 1 为藏北地理位置与土地覆盖)。其幅员辽阔,南北

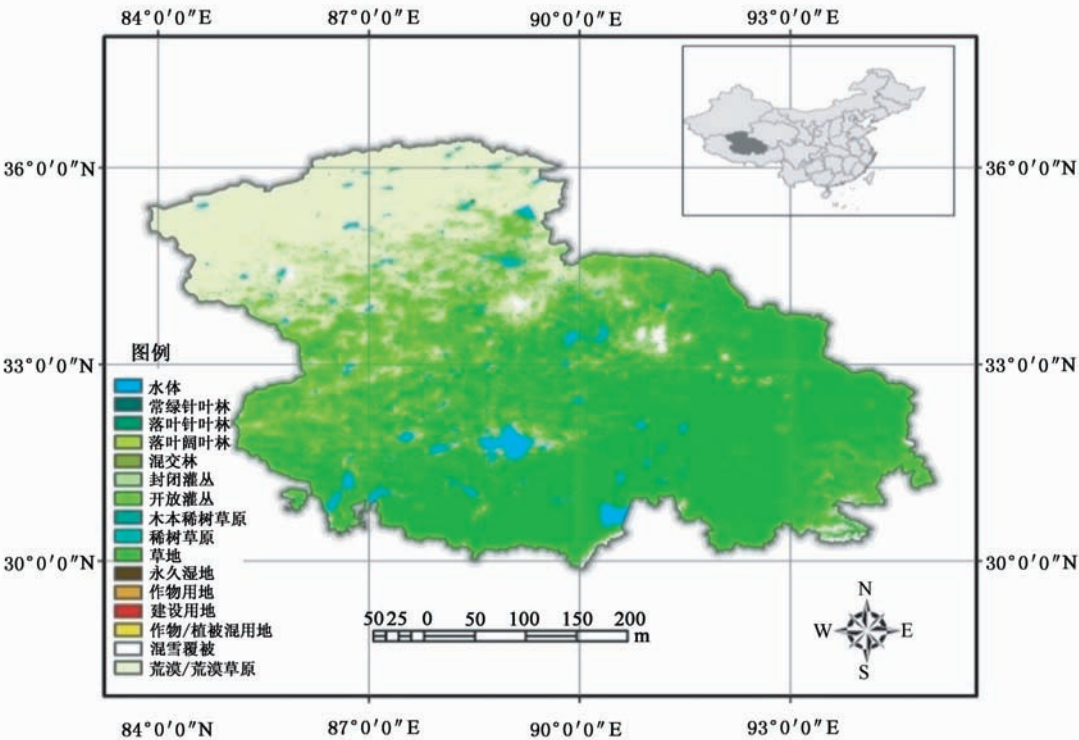


图 1 藏北地理位置与土地覆盖类型图

Fig. 1 Map of North Tibet, showing the location and land cover type of the study area

纵跨 6 个纬度带(30°27′25″~35°39′13″N),东西横越 11 个经度带(83°41′14″~95°10′46″E)。由于疆域辽阔,加之多变的垂直带地形影响,气候及水热条件差异较大,所以,从东南到西北,植被类型分布与生物气候带相适应,从东南向西北依次出现亚高山疏林—灌丛草甸、草甸、草原、荒漠草原四大基带及相邻的过渡亚带<sup>[22]</sup>。此外,星罗棋布的湖泊、冰盖雪原,以及少量耕地、城镇村落点缀其中。藏北地区,不仅有多种类型的植被覆盖,并且其对中国西部地区高寒、干旱、低植被盖度等特征具有代表性,对今后中国西部地区陆地生态系统方面研究中 NDVI 时间序列重建方法的选择提供依据。

(2)数据分析

研究使用的实例数据选择 EOS/Terra 卫星的 MODIS 产品之一 MOD13Q1 V004,包括 MVC 方法 16 天合成的 250m 分辨率 NDVI 及其质量控制

数据,该产品经过几何校正和大气校正。该时间序列数据集从美国 NASA LP DAAC 工作组(LAND PROCESSES DISTRIBUTED ACTIVE ARCHIVE CENTER)网站下载(<https://lpdaac.usgs.gov/lpdaac/products>),时间覆盖 2007 年 1 月—2009 年 12 月 3 年时间共 69 期影像。MOD13Q1 NDVI 原始数据集 DN 值域范围为-3000~10 000 的,其中,DN 为-3000 表示该像元为填充值,从 DN 值转化成正常 NDVI 值域的关系式为:

$$NDVI = 0.0001 \times DN$$

(1)

研究中参考的藏北地区土地覆盖类型数据选用 MODIS 土地覆盖分类产品 MCD12Q1,时间分辨率基本是 1 年,根据 Terra 和 Aqua 卫星一年的观测数据利用监督分类方法制作出的土地覆盖图。它包含 5 种土地覆盖分类方案,本文选择国际地圈生物圈计划(IGBP)定义的分类方案数据(分类类别及代号如表 1 所示)。

表 1 MCD12Q1 产品 IGBP 土地分类方案及其代号

Tab. 1 IGBP land cover classification scheme and codes in MCD12Q1 product

代码	0	1	2	3	4	5	6	7	8
类别	水体	常绿针叶林	常绿阔叶林	落叶针叶林	落叶阔叶林	混交林	封闭灌丛	开放灌丛	木本稀树草原
代码	9	10	11	12	13	14	15	16	254
类别	稀树草原	草原	永久湿地	农田	城镇与建设用地	农田/自然植被混合	冰雪覆被	荒漠及荒漠草原	未分类

研究区域范围内主要的土地覆被类型包括:水体、冰雪、耕地、林地、灌丛、草原、荒漠及荒漠草原。其中,水体、永久湿地、建设用地、和冰雪覆被不在研究范围之内,并将类型 1、2、3、4、5 合并为林地,6 与 7 类合并为灌丛区,8 与 9 地类合并为稀树草原,12 与 14 类合并为作物用地。鉴于荒漠与荒漠草原的覆被 NDVI 时间曲线特征的差异,而在 IGBP 分类法中未将二者进行区分,因此,参考草地类型图(中国科学院地理科学与资源研究所资源环境数据中心提供)及实地验证,区分两种类型的典型样区,选取一定数量代表像元。图 2 是选取藏北地区合并后的 6 类植被覆盖中每种类型的典型像元的原始 NDVI 时间曲线,时间序列中存在明显的突降和突增的 NDVI 噪声点。

3 3 种时序 NDVI 的重建方法

本文采用非对称性高斯函数拟合法、双 Logistic 函数拟合法及 Savitzky-Golay 滤波法对 NDVI

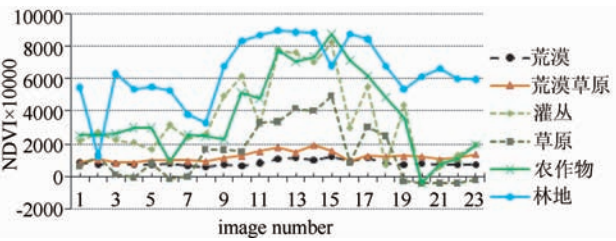


图 2 藏北地区 2008 年各类型覆被代表像元原始 NDVI 时间序列曲线

Fig. 2 Original NDVI time-series curves of representative pixels for each land cover type of North Tibet in 2008 (注:image number 指 MODIS 数据集 16 天合成期号,即每 16 天数据最大值合成法处理为一期,全年共 23 期,下同。)

进行重建,并采用 Timesat 软件包以便于获取季节性植被生长规律信息等<sup>[23-24]</sup>。该软件平台以 Savitzky-Golay 滤波法、非对称高斯函数拟合法及双 Logistic 曲线拟合法为内核,能对 NASS/NOAA PAL、TERRA or AQUA/MODIS 等 NDVI 时间序列数据进行有效地处理提取信息。同时,该软件提供可视化界面也避免了数据处理过程研究人员编

程实现每一种算法并调试各种繁多敏感的参数,而只需要对有效值域范围、滑动窗口大小及迭代次数等几个关键参数进行设置。

(1)非对称性高斯函数拟合法

非对称高斯函数拟合方法是一个从局部拟合到整体拟合的方法过程,使用分段高斯函数来模拟植被生长过程,最后通过平滑连接各高斯拟合曲线实现时间序列重建<sup>[16]</sup>。其过程大致可以分为区间提取、局部拟合和整体连接 3 步骤<sup>[23]</sup>：

首先,利用一个滑动窗口获取 NDVI 时序数据的谷值和峰值；

然后,对位于谷值和峰值之间的 NDVI 时序数据进行两次局部拟合,使得最优化拟合函数较好地描述 NDVI 数据的上包络曲线,局部拟合公式为：

$$f(t) \equiv f(t; c_1, c_2, a_1, \dots, a_5) = c_1 + c_2 g(t; a_1, \dots, a_5)$$
  
(2)

其中,

$$g(t; a_1, \dots, a_5) = \begin{cases} \exp\left[-\left(\frac{t-a_1}{a_2}\right)^a\right], & \text{if } t > a_1 \\ \exp\left[-\left(\frac{a_1-t}{a_4}\right)^{a_5}\right], & \text{if } a_1 > t \end{cases}$$
  
(3)

式中,  $c_1$  和  $c_2$  控制曲线的基准和幅度;  $a_1$  决定峰值和谷值的位置;  $a_2, a_3, a_4, a_5$  控制曲线左、右半部分的宽度和陡峭度。

最后,利用各局部拟合函数构建整体拟合函数,从而较好地描述整个植被生长期的 NDVI 变化过程。整体拟合函数为：

$$F(t) = \begin{cases} \alpha(t)f_L(t) + (1-\alpha(t))f_C(t), & t_L < t < t_C \\ \beta(t)f_C(t) + (1-\beta(t))f_R(t), & t_C < t < t_R \end{cases}$$
  
(4)

式中,  $[t_L, t_R]$  区间是 NDVI 整个变化区间,  $f_L(t)$ ,  $f_C(t)$ ,  $f_R(t)$  分别代表  $[t_L, t_R]$  区间内左边谷值、中间峰值及右边谷值所对应的局部函数,  $\alpha(t)$  和  $\beta(t)$  为位于  $[0, 1]$  之间的剪切系数。

(2)双 Logistic 函数拟合法

双 Logistic 曲线拟合法是由 Beck 等于 2006 年提出的一种新的算法,与非对称高斯函数拟合法类似,双逻辑函数滤波也是一种半局部拟合方法<sup>[24]</sup>。首先将整个时间序列中时间点对应的值按极大或极小值分成多个区间,分别对该区间进行双 Logistic 函数局部拟合,其局部拟合方式与非对称高斯拟合方法类似,如式(2),采用的双 Logistic 函数如公

式(5)所示<sup>[18]</sup>：

$$g(t; a_1, \dots, a_4) = \frac{1}{1 + \exp\left(\frac{a_1 - t}{a_2}\right)} - \frac{1}{1 + \exp\left(\frac{a_3 - t}{a_4}\right)}$$
  
(5)

式中,参数  $a_1, a_2, a_3, a_4, a_5$  的含义与公式(3)中相同。

然后,基于公式(4)的整体拟合函数将各局部拟合函数的特征加以综合,重建新的 NDVI 时间序列曲线。

(3)Savitzky - Golay 滤波法

Savitzky - Golay 滤波法最初由 Savitzky 和 Golay 于 1960s 提出,是一种移动窗口的加权平均算法,但是,其加权系数不是简单的常数窗口,而是通过在滑动窗口内对给定高阶多项式的最小二乘拟合得出。NDVI 时间序列数据集的 Savitzky - Golay 滤波过程可以用下式描述<sup>[10]</sup>：

$$\bar{Y} = \frac{\sum_{i=-m}^m C_i \times Y_{j+1}}{N}$$
  
(6)

其中:  $\bar{Y}$  为合成序列数据,  $Y_{j+1}$  代表原始序列数据,  $C_i$  为滤波系数,  $N$  为滑动窗口所包括的数据点  $(2m+1)$ 。

在 Savitzky-Golay 滤波方法重建 NDVI 序列数据的应用中,需要确定滑动窗口的大小,以及平滑多项式的阶数,以保证 NDVI 数据拟合的准确性。如滑动窗口设置偏小,容易产生大量冗余数据,不易获取数据集的长期趋势;反之,则可能遗漏一些细节信息;同样,较低的阶数可以得到更平滑的结果,但会保留异常值,高的次数可以去掉异常值却过度拟合,产生新的噪声。

上述 3 种方法对 2007 - 2009 年共 69 期 NDVI 时间序列重建效果对比分析处理流程如图 3 所示。首先,通过 MRT 软件对从 NASA 网站上获取的 HDF 格式 MOD13 Q1 产品进行转换,并区域拼接、投影转换等预处理;然后,基于 Timesat 2.3 平台的原始 NDVI 时间序列进行平滑去噪,其关键步骤为设置处理参数,包括 NDVI 有效值域 (range, -0.05 - 1.00)、噪声去除阈值 (spike, 2)、滑动窗口大小 (w-size, 5 - 6 - 7)、拟合峰值参数 (altitude, 5)、迭代次数 (fitting steps, 3) 等,基于 Savitzky-Golay 滤波法、非对称高斯函数拟合法及双 Logistic 曲线拟合法产生拟合后的 NDVI 时间序列数据集。最后,以处理后的 2008 年 MODIS 土地覆盖类型数据,各类型随机选取 30 - 50 个不等的样点,分析 3 种方法对不同类型覆被的重建特征。



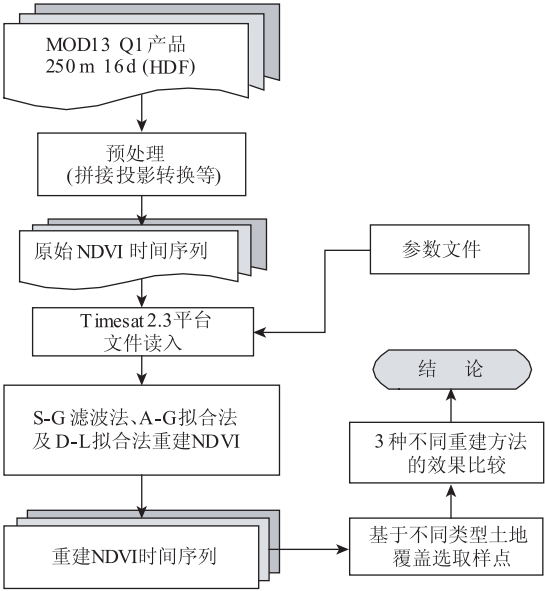


图 3 3 种时序 NDVI 重建方法比较研究处理流程图  
Fig. 3 Flowchart of main procedures involved in comparison of time-series NDVI reconstruction techniques

## 4 重建算法的应用与结果分析

### 4.1 藏北地区不同覆被时序 NDVI 曲线特征重建结果比较

在分析大量像元的 NDVI 时间序列曲线噪声特征及重建效果的基础上,分别选取能充分反映荒漠、荒漠草原、草原、灌丛、作物及林地六种覆被类型不同特征的典型像元,基于 3 种主流 NDVI 时间序列重建方法拟合曲线,如图 4(a - f)所示。荒漠及荒漠草原主要分布在尼玛和双湖特别行政区北部,可可西里、昆仑山与冈底斯山脉之间海拔 5000m 左右的河谷、高原湖盆及山坡地。气候严寒干旱、风大,植被组成单调。在荒漠覆被区域,NDVI 时间序列没有明显的生长季曲线,在 6 - 9 月稍有植被活动迹象,NDVI 值相对冷季有所抬升,并伴随连续的噪声点。一年中遭遇冰雪覆盖的时间长,因此,曲线不规则波动。图 4(a)表明,以局部拟合

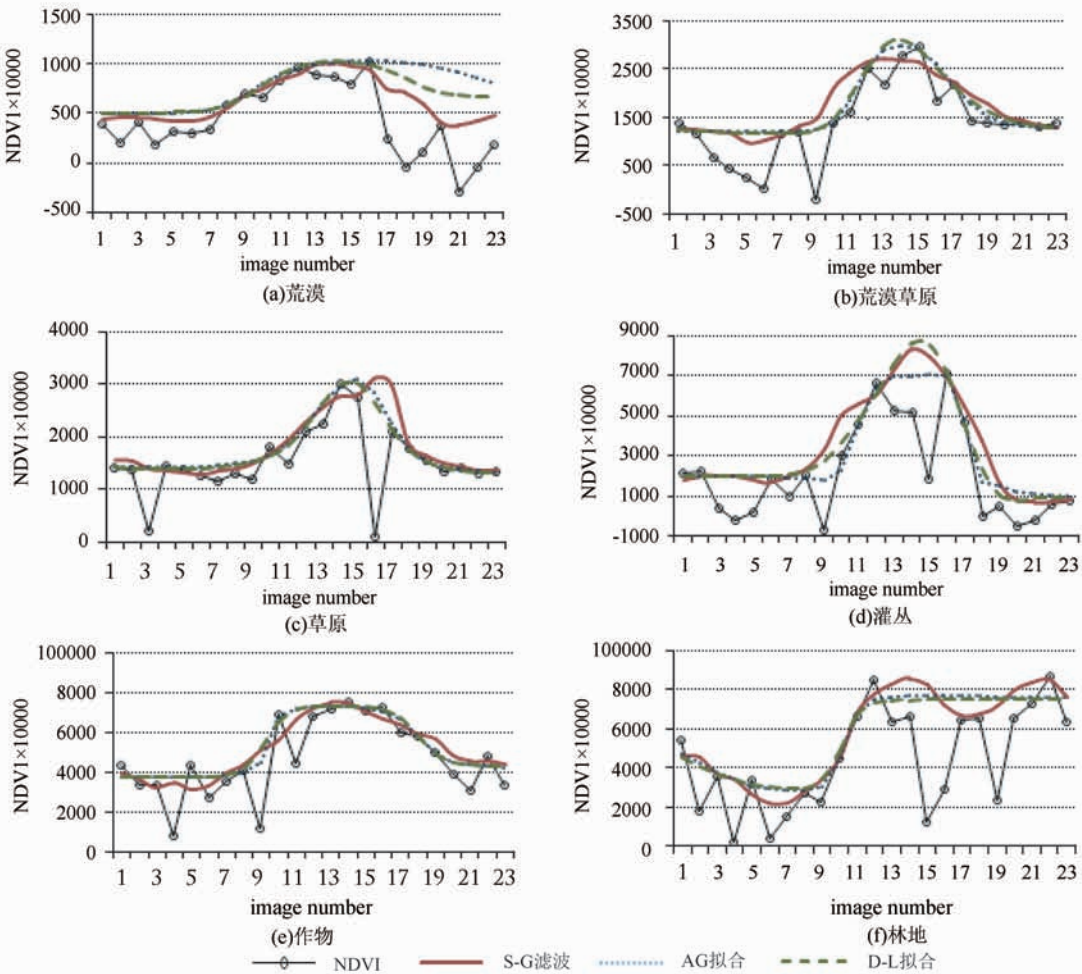


图 4 不同土地覆被类型典型像元 NDVI 时间序列重建前后曲线对比

Fig. 4 Comparison of original and reconstructed MODIS NDVI time-series in 2008

为主的 S-G 滤波法可以更加精确地描述 NDVI 时序数据的细微变化,而 AG 和 D-L 拟合方法以曲线上包络线吻合为主要特征,在消除地面及大气状况造成的不规则噪声波动的同时也使重建 NDVI 偏离真实值。对于 NDVI 的振幅( $NDVI_{max}$  与  $NDVI_{min}$ 之差)较小的时间序列曲线,如图 4(a),D-L 拟合方法重建曲线基本不能拟合 NDVI 的时间变化曲线特征,近似为直线。

荒漠草原具有植被生长曲线特征,但全年 NDVI 普遍偏低,NDVI 最大值通常不超过 0.3,生长季大约  $120 \pm 10d$ ,与荒漠覆被类似,区域受地面冰雪、冻融影响较大。从图 4(b)不难发现,S-G 滤波方法重建由于对地面冰雪引入的噪声点细节拟合,造成生长季曲线的左半部分“左移”,从而使 NDVI 时间序列提取物候期引入误差,如获取的返青期信息较真实情况提前。AG 拟合与 D-L 拟合方法相似,对 NDVI 曲线的上包络线拟合较好,在对  $NDVI_{max}$  的拟合方面,AG 拟合方法更贴近真实值,而 D-L 拟合将使生长季峰值偏大。

草原是藏北地区的主要覆被类型,在中部地区及东部河谷、浅山阳坡地均有分布。植被覆盖以高山嵩草、紫花针茅、莎草、披碱草、禾草为主,具有明显的生长季特征,生长季内 NDVI 在 0.1~0.4 范围变化,受冰雪覆盖噪声影响相对较少。3 种方法对 NDVI 时间序列的重建效果都较好,只是出现强的突降或陡增点,采用 S-G 滤波法拟合曲线略微偏离真实值。

灌丛主要分布在森林带以西或森林线以上,NDVI 时序曲线形态与草原覆被类似,振幅相对更大,生长峰值 NDVI 可达 0.7 左右,由于大多灌丛覆被分布在高山森林线以上,在冬季受降雪影响较草原覆被大。3 种方法的拟合程度类似,而在返青早期和进入休眠期阶段,S-G 滤波方法的拟合效果稍欠缺。当生长峰期出现强噪声或持续性噪声,三种方法表现各异:D-L 拟合通常过度拟合 NDVI 峰值,S-G 滤波受噪声影响过于“保守”,造成  $NDVI_{max}$  值偏低,而 AG 拟合基于对上包络线的拟合,基本保持 NDVI 的真实值。

藏北地区作物用地包括耕地和旱地,分布范围极少。作物 NDVI 时序曲线的振幅大,生长期在 0.2~0.8 的变化范围,具有明显的生长季特征,但作物没有呈现其他植被生长曲线明显的对称性,因此 S-G 滤波法在拟合生长季返青和进入休眠期的时间信息将引入新的噪声(见图 4(e))。然而,S-G

滤波的局部拟合特点,在满足上包络线拟合和保持原始 NDVI 真实值之间具有“高低填平”的效果,对于将全年 NDVI 平均值或累加值作为研究对象或者模型(如作物估产模型)输入参数时,具有较优的效果。而 AG 拟合与 D-L 拟合在拟合生长季曲线和上包络线吻合度较好,前者更甚之。

藏北地区林地主要分布在东南边缘的高山峡谷,气候温暖湿润,少量的常绿针叶林和落叶针阔叶林条带分布。全年覆被 NDVI 值高,部分像元在 5-10 月时段内呈微弱的先增后降趋势,生长季不明显。但是由于森林线以上的冰雪覆盖可能由于温度降低影响林地的 NDVI 值陡降,另外云量相对增多,因此,像元 NDVI 时间序列曲线锯齿状波段显著。S-G 滤波法在拟合林地时序 NDVI 数据其优势明显,AG 与 D-L 拟合法可能在整个年份或某时段无法拟合,如图 4(f)所示。

#### 4.2 3 种重建算法的保真性特征比较

以上对不同覆被类型典型噪声样点分析了 S-G 滤波、AG 拟合及 D-L 拟合 3 种算法的拟合重建特征。同时,针对每种覆被类型的 NDVI 数据,随机选取 50 个左右受噪声影响的像元样点,探索 3 种算法重建时在保持原始 NDVI 时间序列曲线整体特征,以及处理未受噪声“污染”的阶段点真实值(以下简称保真性)的水平。

相关系数表示两个样本数组值间的相关程度,可以用于反映拟合后的 NDVI 时间序列曲线保持原始植被生长季曲线特征的能力。从图 5(a)可以看出,荒漠、荒漠草原、草原、灌丛、作物及林地 6 种覆被,由于荒漠及林地(特别常绿林)植被生长季曲线不明显,且受冰雪等噪声影响较大,因此,3 种算法拟合前后曲线的相关程度相比其他类型要低,其中,荒漠草原与草原其重建前后时间序列曲线相关性最强。具体分析不同类型 NDVI 的 3 种算法重建效果,对于荒漠类型,由于 S-G 滤波方法较强的细节拟合能力,比较适合生长季不明显且值偏低的植被 NDVI 重建,其相关系数较其他两种方法高。同样对于全年 NDVI 值均较大不存在典型的生长季曲线的林地 NDVI 时间序列曲线重建,S-G 滤波方法具有优势。从图中显示,荒漠草原等其他四种类型覆被 NDVI 的重建,S-G 滤波方法效果最差,AG 与 D-L 拟合方法比较接近,且 AG 拟合方法更优,这是因为 S-G 滤波方法受强噪声影响比较大而

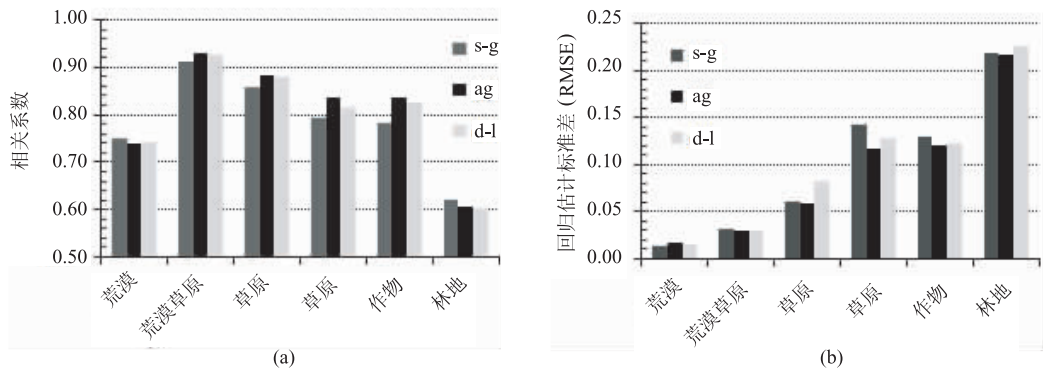


图 5 3 种算法对不同覆被类型 NDVI 重建保真性能力比较

Fig. 5 Comparison of the ability on keeping the main characters by NDVI against different land cover types among the three algorithms

偏离原始曲线。

此外,利用回归估计标准差(RMSE, 计算如式 7)描述重建后的 NDVI 年内时间序列与原始值之间的平均差异程度,反映重建前后 NDVI 值的代表性强弱,其值越小,拟合值的代表性越强。图 5(b)显示,6 种类型植被荒漠、荒漠草原的 RMSE 值偏低,而林地 RMSE 值最高,这一方面受重建算法对时间序列曲线的拟合能力影响,另一方面也与该植被类型本身 NDVI 值大小有关(如林、灌覆被 NDVI 整体偏大,其标准差绝对值相对较大)。对各种类型植被不同方法重建效果分析发现,除荒漠和林地类型 S-G 滤波方法的重建前后差异程度相对较小,荒漠草原等 4 种类型均推 AG 拟合方法最优,D-L 拟合次之。该统计结果与基于典型样点分析 3 种算法对不同覆被类型 NDVI 时间序列重建效果的比较相吻合。

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (NDVI_{p_i} - NDVI_{o_i})^2}{N}} \quad (7)$$

#### 4.3 非对称高斯拟合重建藏北草地 NDVI 结果分析

藏北地区从东南到西北依次以草甸—灌丛、草原、荒漠草原和荒漠为基质,少量亚高山林地、沼泽湿地、城镇建筑用地、耕地或旱地等覆被呈条带或星点状分布。依据上述 3 种 NDVI 时间序列重建方法特征分析,S-G 滤波方法对于生长季曲线特征不是非常明显的覆被,如荒漠及东南部的林区,其拟合细节效果较好,但是,其容易受较强噪声的影响而偏离真实值。对于荒漠草原、草原及草甸、灌丛覆被的 NDVI 时间序列重建,AG 拟合与 D-L 拟

合对生长季开始、结束时段及  $NDVI_{max}$  重建将得到更好的结果,且 AG 拟合方法更合适。藏北地区大部分覆被是草原、草甸以及荒漠草原等,而荒漠与林地的植被覆盖分析通常不是研究的重点,因此,选择 AG 拟合法对整个藏北地区 2007—2009 年共 69 期 NDVI 时间序列曲线拟合。重建前后多数像元的 NDVI 值发生不同程度的改变,有效地消除了由于大气条件、地面冰雪等降低的 NDVI 值或假高值。表 2 为 AG 拟合法重建前后藏北地区 2008 年 209 合成期 NDVI 分级图像像元数目对照,能简单地反映去噪之后各级别类像元变化及 NDVI 总体转移状况。NDVI(-1~1)分为  $\leq 0$ 、 $0 \sim 0.15$ 、 $0.15 \sim 0.30$ 、 $0.30 \sim 0.45$ 、 $0.45 \sim 0.60$ 、 $> 0.60$  共 6 个级别,前两级的像元数目在时间序列重建后明显减少,特别是  $0 \sim 0.15$  范围类的像元数目,近 20% 的像元发生所属级别变化,主要是由高寒环境地面冰雪覆盖造成 NDVI 值偏低被重建纠正,也不排除消除了大气云掩引入的噪声;而  $> 0.15$  的 4 个级别内的像元数目均呈增长状况。从重建前后各级像元数目的动态变化矩阵发现:这种变化包括从低值级别转移,同时该级内像元也转化为更高 NDVI 值范围内;尽管 NDVI 值在  $> 0.60$  的范围内像元数目增多,还是存在部分像元的 NDVI 值降低的情况,主要因为重建处理消除了“假”高值噪声。对整个研究区域像元 NDVI 的平均值统计发现,重建后的值相对于初始值提高 0.02,分析其原因主要在于影像中大部分噪声像元是受云遮盖或临时性降雪或冰雹造成地表实际 NDVI 值降低(在图 3, 4 中 NDVI 时间序列曲线中也能反映),通过重建其真实值得到有效的改善。



表 2 藏北地区 2008 年 209 合成期 MODIS NDVI 重建前后分级像元数量

Tab. 2 Comparison of the number of pixels in each classified NDVI group between raw and reconstructed images

分级	$\leq 0$	0~0.15	0.15~0.30	0.30~0.45	0.45~0.60	$>0.60$	均值
原始数据(raw data)	88928	2311697	2688442	916822	597332	444972	0.22
重建数据(AG fitting)	45564	1895318	2780662	1040766	634433	615302	0.24

NDVI 时间序列的重建使像元值发生不同程度的改变,但是通过连续色阶来反映整个藏北地区处理前后的变化并不显著(人视觉识别的色差的能力限制)。因此,选取藏北 NDVI 值整体偏高的东南

部区域(31.1°~32.6°N, 91.5°~94.0°E)重建前后对照。此区域土地覆被垂直带谱明显,呈现亚高山针阔叶林-灌丛-草甸-苔原矮草-冰雪覆被的梯度变化,能较好地展现 NDVI 的重建效果。图 6 中包括

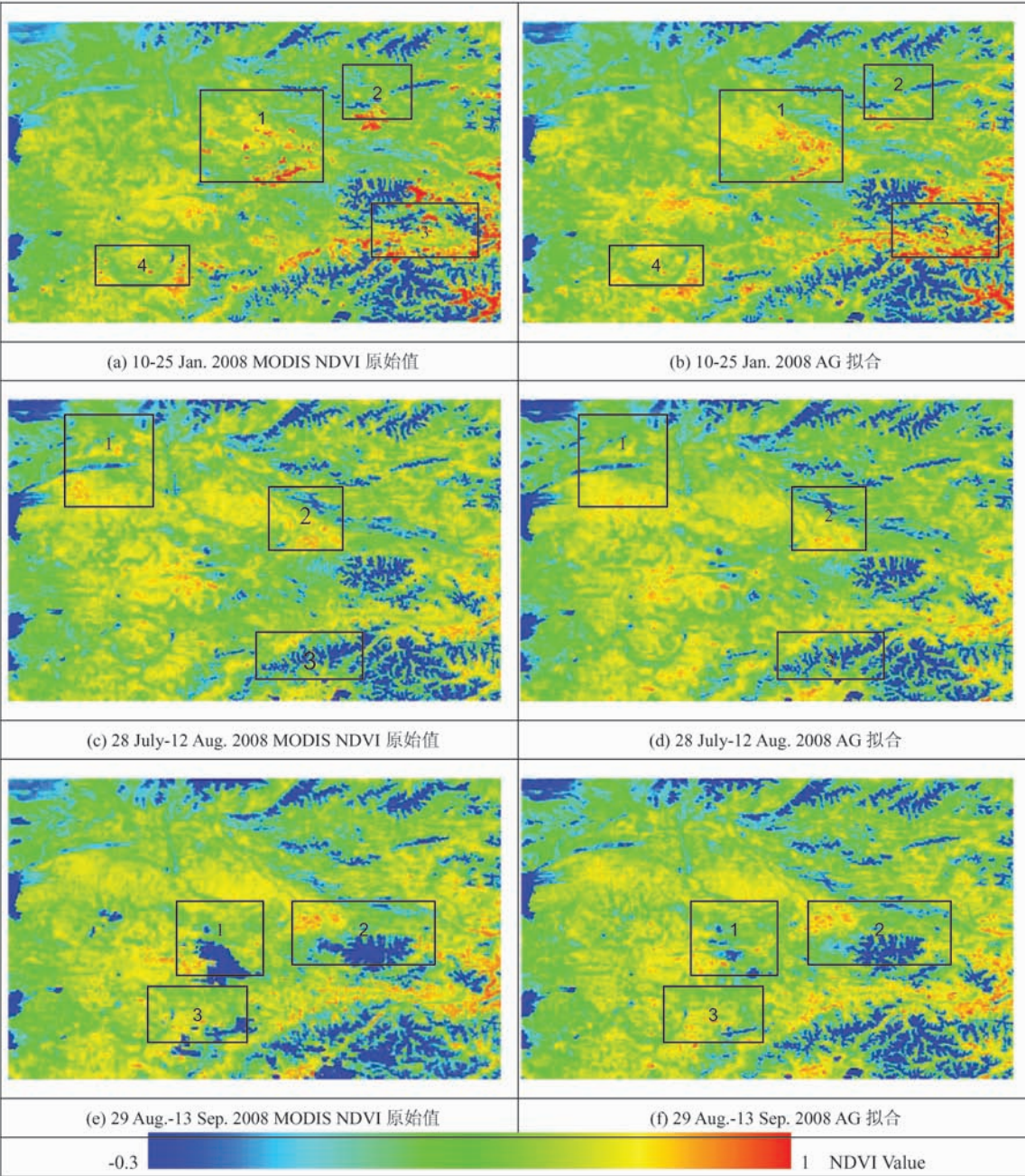


图 6 藏北地区 2008 年中 3 个合成期的 MODIS NDVI 影像 AG 拟合法重建前后比较

Fig. 6 NDVI images of part of North Tibet in three composition periods; before and after the application of AG method



2008 年 6 月 10 - 25 日、7 月 28 日 - 8 月 12 日、8 月 29 日 - 9 月 13 日 3 个合成期的 MODIS NDVI 重建前后对比,分别反映植被返青不久、生长季顶峰、植被进入休眠期前 3 个关键物候期的 NDVI 特征。总体上,NDVI 时间序列重建最大程度地去除地面状况、大气条件造成 NDVI 突降和假高情况,从而使区域内 NDVI 的异质性减弱,在图 6(a)、图 6(b) 中的 1、2、4 标识,图 6(c)、图 6(d) 中的 1、2 标识,图 6(e)、图 6(f) 中的 1、2、3 标识均能得到反映。同时,重建处理消除临时性地面冰雪覆盖造成的 NDVI 突降,使该区域的山脉走向(山谷线)更加明晰(通过林地或灌丛的高 NDVI 值,红橙色表示),这在图 6(a)、图 6(b) 中的 3 标识,图 6(c)、图 6(d) 中的 3 标识均能得到体现。

## 5 结论与讨论

该研究表明,非对称高斯函数拟合与双 Logistic 曲线拟合方法,均是一个从半局部拟合到整体拟合的过程,在对不同类型植被覆盖 NDVI 时间序列曲线重建结果基本一致,其优势在于对时间序列曲线的上包络线拟合效果好,并且在生长季返青期和进入休眠期的过渡时段处理比较精确。两种方法在生长季峰值拟合存在差异,双 Logistic 曲线拟合过度强调  $NDVI_{max}$  而使时间序列曲线整体少量抬升。Savitzky - Golay 滤波法在荒漠和林地覆被 NDVI 时间序列曲线重建具有优势,主要因为这两种覆被时序曲线生长季特征不明显,基于 Savitzky - Golay 的滤波方法能更明晰地模拟局部细节信息,而前两种方法可能会由于曲线“振幅(amplitude)”偏小而出现整段或部分无法拟合的情况。3 种方法重建草原、灌丛覆被类型的 NDVI 时序曲线均具有较好的效果,相对而言,Savitzky - Golay 滤波对噪声比较敏感,特别是灌丛 NDVI 数据的处理,较强的噪声可能使其“过度”拟合而引入新的噪声。此外,农作物生长季曲线不具对称性,基于 Savitzky - Golay 滤波方法拟合在生长季开始和结束段拟合效果不佳,但是其“高低填平”特点使其时段内平均值和累计量结果较好,因此,在考虑不同的应用角度时,如 NDVI 峰值、生长季累积值、全年累积值、全年 NDVI 平均值等指标选取,以及用于反映植被生物量、作物估产或植被活动强度年际监测等,需要针对具体问题选择合适的方法。

非对称性高斯函数拟合法、双 Logistic 曲线拟合法和 Savitzky - Golay 滤波法作为当前 NDVI 时间序列数据重建的 3 种主要手段,其比较研究可以深入了解各方法在不同研究区域、植被覆盖类型、NDVI 时序曲线特征的适应性及其自身优缺点,便于其他研究人员根据研究区域和应用目的选择更适宜的方法,从而最大程度地减少遥感数据噪声引入造成研究结果的不确定性甚至错误结论。然而,本文研究也仅局限于高寒环境的藏北地区土地覆被 NDVI 时间序列数据作为实例进行比较分析,而研究结果表明不同重建方法具有区域和覆被类型差异性,因此,选择一个更大范围或者全球区域、覆被类型更为齐全的比较研究是非常必要的。另外,根据比较研究发现不同方法的优缺点及适应性,可以改进或提出更加全面的 NDVI 时间序列重建方法。

## 参考文献:

- [1] Zhou L, Tucker C J, Kaufmann R, *et al.* Variations in Northern Vegetation Activity Inferred from Satellite Data of Vegetation Index during 1981 to 1999[J]. *Journal of Geophysical Research*, 2001, 106(17): 20069 - 20083.
- [2] Yu F, Price K P, Ellis J, *et al.* Satellite Observations of the Seasonal Vegetation Growth in Central Asia, 1982 - 1990[J]. *Photogrammetric Engineering & Remote Sensing*, 2004, 70(4): 461 - 469.
- [3] Ackerman S A, Strabala K I, Menzel W P, Frey R A, Moeller C C, & Gumley L E. Discriminating Clear Sky from Clouds with MODIS[J]. *Journal of Geophysical Research*, 1998, 103(32), 141 - 157.
- [4] Julien Y, Sobrino J A, & Verhoef W. Changes in Land Surface Temperatures and NDVI Values over Europe between 1982 and 1999 [J]. *Remote Sensing of Environment*, 2006, 103(2006), 43 - 55.
- [5] 范锦龙, 吴炳方. 复种指数遥感监测方法[J]. *遥感学报*, 2004, 8(6): 628 - 636.
- [6] Cihlar J, Ly H, Li Z Q, *et al.* Multitemporal, Multi-channel AVHRR Data Sets for Land Biosphere Studies-artifacts and Corrections[J]. *Remote Sensing of Environment*, 1997, 60: 35 - 57.
- [7] 李杭燕, 颀耀文, 马明国. 时序 NDVI 数据集重建方法评价与实例研究[J]. *遥感技术与应用*, 2009, 24(5): 596 - 602.
- [8] 吴文斌, 杨鹏, 唐华俊, 等. 两种 NDVI 时间序列数据拟合法比较[J]. *农业工程学报*, 2009, 25(11): 183 -

- 188.
- [9] Viovy N, Arino O and Belward A S. The Best Index Slope Extraction (BISE): A Method for Reducing Noise in NDVI Time Series[J]. *International Journal of Remote Sensing*, 1992, 13: 1585 – 1590.
- [10] Lovel J L and Graetz R D. Filtering Pathfinder AVHRR Land NDVI Data for Australia[J]. *International Journal of Remote Sensing*, 2001, 22 (13): 2649 – 2654.
- [11] Jin Chen, Jönsson P, Tamura M, *et al.* A Simple Method for Reconstructing a High-quality NDVI Time-series Data Set Based on the Savitzky-Golay Filter [J]. *Remote Sensing of Environment*, 2004, 91: 332 – 344.
- [12] Ma Mingguo, Veroustraete F. Reconstructing Pathfinder AVHRR Land NDVI Time-series Data for the Northwest of China[J]. *Advances in Space Research*, 2006, 37: 835 – 840.
- [13] 闫慧敏, 曹明奎, 刘纪远, 等. 基于多时相遥感信息的中国农业种植制度空间格局研究[J]. *农业工程学报*, 2005, 21(4): 85 – 90.
- [14] Roerink G, Menenti M, Verhoef W. Reconstructing Cloudfree NDVI Composites Using Fourier Analysis of Time Series[J]. *International Journal of Remote Sensing*, 2000, 21(9): 1911 – 1917.
- [15] Beck P, Atzberger C, Høgda K, *et al.* Improved Monitoring of Vegetation Dynamics at Very High Latitudes: A New Method Using MODIS NDVI[J]. *Remote Sensing of Environment*, 100(3): 321 – 334.
- [16] Jönsson P, Eklundh L. Seasonality Extraction by Function Fitting to Time-series of Satellite Sensor Data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2002, 40(8): 1824 – 1932.
- [17] Gu Juan, Li Xin, Huang Chunlin, Okin G S. A Simplified Data Assimilation Method for Reconstructing Time-series MODIS NDVI Data [J]. *Advances in Space Research*, 2009, 44: 501 – 509.
- [18] Julien Y, Sobrino J A. Comparison of Cloud-reconstruction Methods for Time Series of Composite NDVI Data[J]. *Remote Sensing of Environment*, 2010, 114: 618 – 625.
- [19] Sakamoto T, Yokozawa M, Toritani H, *et al.* A Crop Phenology Detection Method Using Time-series MODIS Data [J]. *Remote Sensing of Environment*, 2005, 96(3/4): 366 – 374.
- [20] Hird J N, McDermid G J. Noise Reduction of NDVI Time Series: An Empirical Comparison of Selected Techniques [J]. *Remote Sensing of Environment*, 2009, 113: 248 – 258.
- [21] 甘肃草原生态研究所草地资源室和西藏自治区那曲地区畜牧局. 西藏那曲地区草地畜牧业资源[M]. 兰州: 甘肃科学技术出版社, 1991, 12 – 68.
- [22] 李辉霞. “3S”技术在草地退化动态监测中的应用——以西藏自治区那曲县为例[D]. 中国科学院成都山地灾害与环境研究所, 2004.
- [23] 李儒, 张霞, 刘波, 等. 遥感时间序列数据滤波重建算法发展综述[J]. *遥感学报*, 2009, 13(2): 335 – 341.
- [24] Jönsson P, Eklundh L. TIMESAT: A Program for Analyzing Time-series of Satellite Sensor Data [J]. *Computers & Geosciences*, 2004, 30: 833 – 845.

## Analysis on Three NDVI Time-series Reconstruction Methods and Their Applications in North Tibet

SONG Chunqiao<sup>1, 2</sup>, YOU Songcai<sup>3</sup>, KE Linghong<sup>1, 2</sup>, LIU Gaohuan<sup>1</sup>

(1. State Key Lab of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China;

3. Institute of Environment and Sustainable Development in Agriculture; Chinese Academy of Agricultural Sciences, Beijing 100081, China)

**Abstract:** Remote sensing vegetation index time-series datasets have been applied widely in various fields, such as monitoring and simulation of terrestrial ecological environment system on global and regional scales, analysis of vegetation cover dynamics, recognizing features and extracting information of vegetation phenology and so on. However, due to errors from information collection and transmission process of

sensors, different atmospheric conditions and random ground situations etc, various residual noise exists in NDVI time-series data. Therefore, a series of NDVI time-series data reconstruction methods have been developed to solve the problem. In this paper, various newly proposed and modified reconstruction methods are summarized and evaluated firstly. Then the three primary methods, i. e. asymmetric Guassian (AG) function fitting, double logistic (D-L) fitting method and Savitzky-Golay (S-G) filtration, are introduced in terms of basic principles and features. Based on the TIMESAT program, by taking some NDVI time-series of pixel samples of different land cover types as instance, the suitability and defects of the three approaches are analyzed and discussed. The result shows that the former two methods generally present similar performances, but on fittings of peak of these time-series curves there are minor differences. Besides, the three methods have inconsistency effects in different study areas, land cover and research purposes. In detail, the S-G method has a more excellent performance in reconstructing NDVI curves in alpine desert and forest-covered area with much noise, while the D-L and AG function fitting methods could conduct more accurate reconstructed time-series dataset in alpine shrub or marshy grassland regions, especially for fitting the start and end period of vegetation growing season. For crops-planted land cover, the double logistic fitting method can't fit the correct curve of NDVI time-series, and the asymmetric Guassian function fitting has the best performance. Finally, based on AG function fitting method, the NDVI time-series of 2007—2009 in North Tibet are reconstructed, and the processed NDVI values are clearly more homogeneous than those in raw data.

**Key words:** MODIS NDVI; time-series reconstruction; North Tibet; asymmetric Guassian function fitting; double logistic curve fitting; Savitzky-Golay filtration