

遥感时间序列数据滤波重建算法发展综述

李 儒^{1,2}, 张 霞¹, 刘 波^{1,2}, 张 兵³

1. 中国科学院 遥感应用研究所 遥感科学国家重点实验室, 北京 100101;

2. 中国科学院 研究生院, 北京 100039;

3. 中国科学院 对地观测与数字地球科学中心, 北京 100080

摘 要: 遥感时间序列数据 (MODIS, NOAA/AVHRR, SPOT/VEGETATION 等) 在植被生长监测、物候信息提取、土地利用类型监测等诸多领域得到了广泛应用, 是生产研究的重要数据源之一。由于传感器、云层大气等影响, 遥感时间序列数据存在着严重的噪声, 应用前必须进行序列滤波重建工作。综述现有各类滤波重建方法, 对研究中广为采用的 3 类主要方法 (基于最小二乘的非对称高斯函数拟合、Savitzky-Golay 滤波、基于离散傅里叶的系列分析方法) 集中阐述其理论基础、应用步骤和优缺点。总结当前遥感时间序列滤波重建方法需要进一步改进之处。

关键词: 遥感时间序列, 滤波, 重建

中图分类号: TP751.1 **文献标识码:** A

1 引言

遥感时间序列数据 (MODIS, NOAA/AVHRR, SPOT/VGT 时间序列等) 应用广泛, 如植被生长监测 (张霞, 2006; 范锦龙 & 吴炳方, 2004), 土地覆盖类型解译 (Bethany *et al.*, 2007), 物候特征识别与信息提取 (Xin *et al.*, 2002), 甚至应用于植物物候变化对 GPP 和 NPP 的影响研究 (White *et al.*, 1999)。因此它是许多研究、生产必要数据源之一。

由于卫星传感器自身性能、云层、大气条件等影响, 时间序列存在严重的噪声, 限制了它的深入应用 (Gutman *et al.*, 1995; Cihlar *et al.*, 1997)。因此在应用前必须正确、有效的去除噪声、重建时间序列 (林忠辉等, 2006)。为了有效的解决这个问题, 研究者发展了许多方法 (顾娟等, 2006)。

2 遥感时间序列数据滤波重建算法归纳

2.1 最大值合成法 (MVC)

时间序列首先需要去云处理和缺损值拟合, 最

大值合成法 (MVC) (Holben, 1986) 是在生成时间序列过程中解决这一问题的主要方法。它遵循基本假设, 即云层的位置是不断变化的。因此可以理解在一段时间内, 任意一处有可能存在一个没有云层覆盖的晴天, 反映在图像上就是该位置在这一时间段内的图像 DN 值 (或植被指数) 具有某种特性, 如在 NDVI 图像中即表现为相对较大值。按照事先规定的时间间隔, 将间隔内序列中各点按值大小排序, 新图像某点的值为序列中该点最大值。该方法在生产时间序列初级产品去除云噪声和拟合缺损点时很有效, 但实际应用中用户得到的是合成数据, 再次使用最大值合成法欠妥。

2.2 最佳指数斜率提取法 (BISE)

最佳指数斜率提取法 (BISE) 认为时间序列中的突变 (突增突降) 可能是有云或无云状态转变, 或是传感器视场角发生变换造成的, 因此可能会出现虚假最大值的情况, 故采用滑动合成时段避免上述问题、识别并替代序列中的噪声。这一思想有 2 种实现, 一是 Viovy 等 (1992) 提出的原始算法, 另外是 Lovell & Graetz (2001) 提出的改进算法。Viovy 等分别以 10d 为间隔, 尝试选择滑动窗口为 10, 20, 30, 40d, 对比结果发现 30d 比较理想。但不论是原算

法还是改进算法,需要根据不同气候条件和地域植被特征,依经验尝试选择参数,这一定程度限制了该方法的应用。

2.3 时间窗口的线性内插算法(TWO)

时间窗口的线性内插算法(TWO)是 Park & Tateishi(1998)提出用来重建 NDVI 时间序列的修正算法。该算法可以在无辅助数据的情况下,仅利用时间序列本身检测并消除云和大气影响。它基于3个假设:(1)真实的 NDVI 不低于观测值;(2)在一个给定窗口内给定像元的真实 NDVI 值表现单调性(单调递增或者递减);(3)在一定时期内,给定像元的真实 NDVI 不低于同期观测 NDVI 的最大值。该方法的核心问题是时间窗口大小的选择。Park & Tateishi 在其实验中得出的结论显示:当使用 10 d 合成 NDVI 数据时,4—5d 的时间窗口相对可行。显然这也是依靠经验选择关键参数,使得算法结果带有极大的不确定性。

2.4 基于非对称的高斯函数拟合法(AGFF)

基于非对称的高斯函数拟合法(AGFF)是由 Per 等在 2002 年提出的,这种方法首先在时间维选择一最大或最小值区间作为局部拟合区间,使用局部拟合模型,即高斯形拟合函数对这一区间数据进行拟合,最后通过全局拟合模型,将局部拟合结果合并。该方法采用分段拟合思想,保证了拟合结果更符合当前时段的真实特性,较少受外界或全局量干扰,仅从这点而言,应用在长时间范围内季相信息的提取优势明显。

2.5 Savitzky-Golay 滤波(S-GF)

Savitzky-Golay 滤波(S-GF)获得相对较多关注(Savitzky & Golay, 1964; 辜智慧, 2003; Chen *et al.*, 2004),许多学者对它进行改进或加入辅助数据以提高其运行精度。这种方法重建的 NDVI 时间序列能够清晰描述序列的长期变化趋势以及局部的突变信息,且不受数据时间、空间尺度和传感器的限制。对该方法产生影响最大的因素是滑动窗口大小的选择。由于参数依靠经验获得,该方法存在人为扰动的问题。

2.6 基于离散傅里叶分析的系列算法

基于离散傅里叶分析的系列算法(Sellers

et al., 1996; Immerzeel *et al.*, 2005)也称谐波分析法(Harmonic Analysis Algorithm, HAA)。将每个独立数据点的时间曲线表达为一系列余(正)弦波加性项的和,每一余弦波由唯一振幅和相位确定。将这些连续的谐波项叠加起来生成一条复杂的曲线,每一谐波项在原始时间序列曲线的总方差中占有一定百分比。由于一些输入参数不能够通过一定客观准则确定,通过经验比较获得,这使该方法处于不利地位。但尽管如此,其应用仍然十分广泛,是遥感时间序列重建的主要算法之一。

2.7 局部最大值拟合法(LMF)

局部最大值拟合法(LMF)(Nagatani *et al.*, 2002; Sajia *et al.*, 2007)先进之处是其在傅里叶曲线拟合(重建)前先做数据去噪,用去噪结果作为傅里叶拟合算法输入集。该方法分为2步:数据去噪和傅里叶曲线拟合(重建)。在2个相邻的等大小(指数据点个数)区间 A, B 中分别选出其中最大值,则两个区间共同拥有的数据点 t_{middle} 的值即为 $\min(\max(A), \max(B))$, 该值作为 t_{middle} 滤波结果输出值。该方法把数据滤波与曲线拟合分开处理,如果此策略合理有效,与傅里叶方法相比,将节省大量时间消耗,非常具有现实意义,因为在常规的傅里叶曲线拟合中大半时间和精力耗费在为数据去噪进行的迭代运算中。据称该方法可以有效的去除云、雾和其他一些大气影响,在作物种植模式识别应用中效果良好。

以上基本涵盖了当前遥感时间序列数据滤波重建的绝大部分算法。除了最大值合成法在生产合成时间序列数据初级产品时得到大家基本认可外,其他方法学界尚没有达成共识,研究者在努力改进或者研究新的算法,希望新方法可以克服现有方法固有的缺陷,有效滤去噪声、反映植被真实的生长特性和物候规律等。

3 三类主要滤波重建算法介绍及评述

3.1 基于非对称高斯函数拟合

基于非对称高斯函数拟合方法(Per *et al.*, 2002)使用分段高斯函数(曲线)组合来模拟植被季相生长(物候)规律,一个组合代表一次植被盛衰过

程,最后通过平滑连接各高斯拟合曲线,实现时间序列重建。其过程大致分为:区间提取、局部拟合(local model function)和整体连接(global model function)。

3.1.1 区间提取

由于该方法是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,每个区间只包含1个极值,分别对该区间进行高斯拟合,故区间选择是关键一步,影响后期全局拟合(曲线连接)效果。

单纯从数据序列中选取这样的区间本身不是难点,但由于噪声影响,使得该区间内数据跳动很大。因此区间提取前,必须进行数据平滑处理。Per & Lars (2002)通过滑动均值窗口来实现数据平滑提取时间窗口。该方法区分左右半边窗口(n_L , n_R),亦即滑动窗口的大小为($n_L + n_R$)。窗口中心点左右分别选用不同数目的邻居点参与数据平滑,体现了对前后数据不同的信赖程度,近似可理解为一种加权方式。显然这两个值控制着数据平滑效果,该值越大,结果越平滑,但损失的信息也就相对越多。

3.1.2 局部拟合

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

裁切函数(cutoff function) $\alpha(t)$, $\beta(t)$ 是分别定义在($t_L + t_C$)/2 和($t_C + t_R$)/2 的函数。

通过全局拟合函数 $F(t)$ 将局部拟合曲线合并是该方法的关键之一。Per 等(2002)强调,采取这样分段拟合的策略避免全局数据对局部拟合的干扰,各个峰值相对独立使该方法在拟合复杂情况下的时间序列曲线具有良好的适用性和灵活性,拟合后的曲线更接近真实情况,但该方法存在区间提取问题。使用滑动均值窗口实现数据平滑效果暂且不论,仅对序列中数值的干扰,就足以对后期拟合产生相当影响。而且对于某些强噪声,算法也很容易将其误判为提取目标而获得虚假区间。Per 等(2002)指出,有望通过傅里叶序列拟合确定平均季节数目来解决该问题。

3.2 Savitzky-Golay 滤波

Savitzky-Golay 滤波基本公式:

$$Y_j^* = \frac{\sum_{i=-m}^{i=m} C_i Y_{j+i}}{N}$$

定义:高斯函数 $f(t)$

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

其中:

$$g(t; a_1, \dots, a_5) = \begin{cases} \exp\left[-\left(\frac{t - a_1}{a_2}\right)^{a_3}\right], & \text{当 } t > a_1 \text{ 时} \\ \exp\left[-\left(\frac{a_1 - t}{a_4}\right)^{a_5}\right], & \text{当 } t < a_1 \text{ 时} \end{cases}$$

线性参数 c_1, c_2 决定曲线的基准和振幅; a_1 是与时间 t 相独立的参数,决定曲线最大值或最小值的位置; a_2, a_3, a_4, a_5 分别定义左右半边曲线的宽度和峭度(与曲线峰挺拔程度有关的量)。这4个参数控制拟合结果尽可能适应非对称条件下时间序列曲线拟合,其效果优于傅里叶拟合中使用绝对对称的sine和cosine曲线。

3.1.3 整体连接

定义全局拟合函数:使用 $f_L(t)$, $f_C(t)$, $f_R(t)$ 分别表示一个极大值区间的局部拟合函数,及其左右两边两个极小值区间的左极小值拟合函数和右极小值拟合函数,整个区间用 $[t_L, t_R]$ 表示,则全局拟合函数:

式中 Y 是指NDVI的原始值, Y^* 是NDVI的拟合值, C_i 是第 i 个NDVI值滤波时的系数, N 是指卷积数目,也等于滑动数组的宽度($2m + 1$)。系数 j 是指原始NDVI数组的系数。滑动数组包含有($2m + 1$)个点。该方法本质上是一种平滑滤波,因此2个参数控制滤波效果,一个是 m ,即滤波窗口大小;二是平滑多项式的次数,较低的次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但又过拟合,造成新噪声出现。由于使用窗口滑动进行数据平滑,会使序列前 m 个和后 m 个数据无法参与运算,如果将数据系列首尾相接成环,即可解决此问题。在文献(辜智慧,2003; Chen *et al.*, 2004)中使用了这样的策略。

需要特别指出的是文献(辜智慧,2003; Chen *et al.*, 2004)使用该算法时引入了“拟合影响因子”,通过比较前后相邻拟合计算得到的该因子,确定拟合是否达到最佳而停止。这比傅里叶系列方法中通过经验确定拟合终止条件要客观先进得多。“拟合影响因子 F_k ”定义为:第 k 次拟合的影响因

子 F_k :

$$F_k = \frac{\sum_{j=1}^L (N_j^{k+1} - N_j^0)^2}{L}$$

N_j^{k+1} 是指 NDVI 经过 k 次拟合后得到的数列, N_j^0 是原始的 NDVI 值, L 是经过“输入值有效性”判别后被认定为有效输入值参加下步拟合的点的个数。当 F_k 符合如下条件时, 拟合终止: $F_{k-1} \geq F_k \leq F_{k+1}$ 。显然, F_k 是对观测值与真值间接近程度的刻画, 当它达到局部最小时, 标志这一序列点的拟合值与真值达到了最接近的程度。

S-GF 方法只需确定 2 组参数(滑动窗口尺寸 m 和多项式次数), 即可获得较高质量的 NDVI 数据, 并且参数敏感度低于 BISE。辜智慧(2003)给出了当研究长期变化趋势时的最优滤波参数为(7, 2), 拟合循环中的最优滤波参数为(3, 4)。该组参数的选定还需要结合研究区域的实际情况确定。据称该方法还可以有效的去除多时相 NDVI 遥感数据中由于云、气溶胶等大气影响造成的噪声, 充分利用对应的云状态数据, 可用于不同时间尺度、空间尺度和传感器的数据以获取较高质量的时序数据, 理论简单并易于实现。但该方法在 2 种情况下会失效: 一种是云和大气条件对 NDVI 的影响过于频繁或正好在植物生长高峰; 另一种是由于收割造成 NDVI 正常低值可能会被该方法加以提高。另外, 该方法只能对等间距的时间序列数据滤波重建, 这是由其方法本身造成的。

3.3 傅里叶拟合系列方法

傅里叶方法在时间序列曲线拟合(重建)中有丰富的应用, 以它为基础发展出了很多方法, 如傅里叶拟合(Olsson & Eklundh, 1994; Menenti *et al.*, 1993; Andres *et al.*, 1994; Moody & Johnson, 2001)、Sellers 算法(Sellers *et al.*, 1996; 林忠辉等, 2006)和 HANTS 算法(Immerzeel *et al.*, 2005; Jakubauskas *et al.*, 2001, 2002; Roerink *et al.*, 2000; 王丹等, 2005)等。这些算法使用最小二乘法拟合傅里叶系数矩阵, 通过迭代保留或剔除“假点”, 实现滤波最终获得拟合曲线达到序列重建目的。该算法中引入最小二乘拟合求出合适的傅里叶曲线系数阵 C , C 为 $(2 \times m - 1, 1)$ 的矢量。谐波波数(频数 m)是该方法重要参数, m 的大小决定拟合曲线的形状: m

大, 拟合曲线不够平滑, 但保留细节多, 同时可能造成拟合结果偏离实际情况, 达不到滤波效果; m 值小, 则产生的拟合曲线平滑, 描述细节少。

Seller 算法是在傅里叶原理基础上发展起来的。该算法引入权重, 确立了相应的定权准则使傅里叶分析方法更加合理, 因为去除云和大气影响主要途径即是比较 NDVI 观测值和拟合值, 通过一定的准则, 对不同值赋高低权重, 根据权重来选择参与拟合的点。权重可以理解为待拟合点对曲线的贡献率, 权重越大在拟合过程中贡献越大。不同学者的权重制定准则不同, 运行结果自然有优有劣。Sellers 等(1996)算法引入的定权准则:

$$W_i = \begin{cases} 0 & \text{当 } U_i \leq -k \text{ 时} \\ [1 + (U_i + r)/k]^4 & \text{当 } -k < U_i < -r \text{ 时} \\ 1 & \text{当 } -r \leq U_i \leq r \text{ 时} \\ [1 + (U_i - r)/k]^2 & \text{当 } U_i > r \text{ 时} \end{cases}$$

$$\text{同时应有: } \begin{cases} 0 \leq W_{i=1} \leq 1 \\ 0 \leq W_{i=n} \leq 1 \end{cases}$$

式中, $U = (Y - \hat{Y})/M$, M 是真值与拟合值之差绝对值的中数, 是根据观测值与估值间距离赋权, 这一方法通常情况是适用的, 但在拟合迭代过程中可能会出现某些噪声值更接近拟合值, 造成非噪声点估计与真值之差反而变大。这种情况下, 如果按照距离定权拟合就可能把非噪声点剔除而继续保留噪声参与曲线拟合。如果去除噪声点后, 非噪声点观测值与拟合值的差值马上变小。林忠辉等(2006)在研究中国华北 10d NDVI 时间序列时, 指出 Sellers 定权准则的其他缺陷。该准则主要适用于热带阔叶林, 对短时间间隔合成的时间序列效果欠佳。这些问题也表现在 Sellers 等定权未能很好考虑地表植被的动态变化过程, 如对动态异常高值仍赋高权重使冬季 NDVI 值整体高估等。

傅里叶谐波分析发展的另一种算法: HANTS (Harmonic Analysis of Time Series) 不像快速傅里叶变换算法(FFT)那样严格要求图像等时间间隔, 可以处理不等时间间隔的图像序列(Roerink *et al.*, 2000)。现在大家通常使用的 HANTS 分析是由 National Aerospace Laboratory 编程实现的封装 HANTS 软件包。该软件运行总共需要输入 8 个参

数,其中重要的是频数和拟合允许误差(FET)。但遗憾的是这两个参数要靠经验、不断的尝试才能获得最佳值,这也直接引入了人为干扰因素,是目前基于傅里叶分析方法的通病之一。同时,正如前所述,如何定权关系到该类方法性能好坏,不同的地域通常在定权上会有相应的侧重,HANTS软件的定权准则是集成在软件里面的,属于黑箱操作,无法根据现实条件及时调整,从而影响其具体应用效果。

基于傅里叶分析的方法主要优势在于它既很好的去除噪声,又能够利用代表不同植被不同生长周期的谐波,最大可能的反映植被生长曲线内部蕴含的周期性规律,利用遥感图像存在的时间性和空间性特点将其空间上分布规律和时间上的变化规律联系起来(张霞,2006)。这是其他方法所不可比拟的。

4 方法简评和展望

从遥感时间序列数据被关注以来,尽管有很多时间序列滤波重建算法提出,但并没有形成学者一致认同的方法。已有方法会有不能避免的理论或技术缺陷,主要表现在:

(1)现有方法为了使自己得以应用而附加假设或者限制条件、排除适用,表现在数据预处理中,通常通过设置阈值提前剔除数据,这对序列重建到底产生多大的影响没有严密的论证。

(2)各类方法中的参数大多要依经验选取,如谐波分析中的频数、迭代拟合终止条件的选择(Roerink *et al.*, 2000;张霞,2006)、S-GF滤波(辜智慧,2003;Chen *et al.*, 2004)、LMF(Sajia *et al.*, 2007)、非对称高斯法(Per *et al.*, 2002)中滑动窗口大小的选择等。这些是应用中的关键参数,选择好坏直接影响算法应用的成功与否,人工选择难免引入个人主观干扰。同时由于是经验选择,整个测区使用同一参数值也多与实际不符,尤其是像中国这样土地利用类型复杂的地域。

(3)各类方法中有尚未很好解决的技术难点,该问题影响算法应用。如在谐波分析中定权准则的确定,传统的按距离定权是否合理、滤波中的抛点方法是否得当;非高斯对称法中区间提取时数据平滑对序列拟合产生的影响也尚未有评估,这样会

产生结果可靠性问题。

研究者往往有选择性地对不同应用开发适用于该领域的算法,如谐波分析因为其自身特点在植被生长季相特征、物候信息提取等领域使用频繁,Savitzky-Golay滤波在植被类型提取中表现出色,非对称高斯拟合在长期季相信息提取中效果好。对现有算法改进和应用在不断完善,如局部最大值拟合法的提出。如何妥善解决以上问题,是今后遥感时间序列数据滤波重建研究努力的方向。目前,根据不同的应用领域选择合适的方法,通过实验选择参数组合使用各类方法也是应用趋势之一。

REFERENCES

- Andres L, Salas W and Skole D. 1994. Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. *International Journal of Remote Sensing*, **15**:1115—1121
- Bethany A B, Robert W J, John F H, *et al.* 2007. A curve fitting procedure to derive inter-annual phenologies from time-series of noisy satellite NDVI data. *Remote Sensing of Environment*, **106**: 137—145
- Chen J, Per J, Masayuki T, *et al.* 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the savitzky-golay filter. *Remote Sensing of Environment*, **91**: 332—344
- Cihlar J, Ly H, Li Z Q, *et al.* 1997. Multitemporal, multichannel AVHRR data sets for land biosphere studies-artifacts and corrections. *Remote Sensing of Environment*, **60**:35—57
- Fan J L and Wu B F. 2004. A methodology for retrieving cropping index from NDVI profile. *Journal of Remote Sensing*, **8**(6): 628—636
- Gu J, Li X and Huang C L. 2006. Research on the reconstructing of time-series NDVI data. *Remote Sensing Technology and Application*, **21**(4): 391—395
- Gu Z. 2003. A Study of Calculating Multiple Cropping Index of Crop in China Using SPOT/VGT Multi-temporal NDVI Data. Beijing: Beijing Normal University
- Gutman G and Ignatov A. 1995. Global land monitoring from AVHRR: potential and limitations. *International Journal of Remote Sensing*, **16**: 2301—2309
- Holben B N. 1986. Characteristics of maximum-value composite images for temporal AVHRR data. *International Journal of Remote Sensing*, **7**: 1435—1445
- Immerzeel W W, Quiroz R A and Jong S M. 2005. Understanding precipitation patterns and land use interaction in tibet using harmonic analysis of SPOT VGT-S10 NDVI time series. *International Journal of Remote Sensing*, **26**(11):2281—2296

- Jakubauskas M E, Legates D R and Kastens J H. 2001. Harmonic analysis of time-series AVHRR NDVI data. *Photogrammetric Engineering & Remote Sensing*, **67**(4):461—470
- Jakubauskas M E, Legates D R and Kastens J H. 2002. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Computers and Electronics*, **37**:127—139
- Lin Z H and Mo X G. 2006. Phenologies from harmonics analysis of AVHRR NDVI time series. *Transaction of GSAE*, **22**(12):138—144
- Lovel J L and Graetz R D. 2001. Filtering pathfinder AVHRR land NDVI data for Australia. *International Journal of Remote Sensing*, **22**(13):2649—2654
- Menenti M, Azzali S, Verhoef W and van S R. 1993. Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. *Adv. Space Res.*, **13**:233—237
- Moody A and Johnson D. 2001. Land-surface phenologies from AVHRR using the discrete Fourier transform. *Remote Sensing of Environment*, **75**:305—323
- Nagatani I, Saito G, Toritani H and Sawada H. 2002. Agricultural map of Asian Region using time series AVHRR NDVI data. Proceedings Online of the 23rd Asian Conference on Remote Sensing, <http://www.gisdevelopment.net/aars/acrs/2002/pos2/184.pdf>
- Olsson L and Eklundh L. 1994. Fourier series for analysis of temporal sequences of satellite of sensor imagery. *International Journal of Remote Sensing*, **15**(18):3735—3741
- Park J and Tateishi R. 1998. Correction of time series NDVI by the method of temporal window operation (TWO). Proceedings of the 1998 Asian Conference on Remote Sensing. <http://www.gisdevelopment.net/aars/acrs/1998/ps2/ps2004.shtml>
- Per J and Lars E. 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geosciences and Remote sensing*, **40**(8):1824—1832
- Roerink G J, Menen M and Verhoef W. 2000. Reconstructing cloudfree NDVI composites using fourier analysis of time-series. *International Journal of Remote Sensing*, **21**(9):1911—1917
- Sajia A, Ipshita S, Kazi G R, *et al.* 2007. Adapting the LMF temporal splining procedure from serial to MPI/linux clusters. *Journal of Computer Science*, **3**(3):130—133
- Savitzky A and Golay M J E. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**:1627—1639
- Sellers P J, Los S O, Tucker C J, *et al.* 1996. A reversed land surface parameterization (SiB2) for atmospheric GCMs, Part 2: the generation of global fields of terrestrial biophysical parameters from satellite data. *Journal of Climate*, **9**:706—737
- Viovy N, Arino O and Belward A S. 1992. The best index slope extraction (BISE): a method for reducing noise in NDVI time series. *International Journal of Remote Sensing*, **13**:1585—1590
- Wang D, Jiang X G, Tang L L and Xi X H. 2005. The application of time-series fourier analysis to reconstructing cloud-free NDVI images. *Remote Sensing for Land & Resources*, **64**:29—32
- White M Z, Running S W and Thornton P E. 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. *International Journal of Biometeorology*, **42**(3):139—145
- Xin J F and Louise L. 2002. Mapping crop key phenological stages in the north China plain using NOAA time series images. *International Journal of Applied Earth Observation and Geoinformation*, **4**:109—117
- Yves J, José A and Sobrino W V. 2006. Changes in land surface temperatures and NDVI values over Europe between 1982 and 1999. *Remote Sensing of Environment*, **103**:43—55
- Zhang X. 2006. Study on Characteristics of Temporal Profile of Spectral Index and Its Application in Cropping System Information Mapping. Beijing: Chinese Academy of Sciences

附中文参考文献

- 范锦龙, 吴炳方. 2004. 复种指数遥感监测方法. *遥感学报*, **8**(6):628—636
- 顾娟, 李新, 黄春林. 2006. NDVI 时间序列数据集重建方法述评. *遥感技术与应用*, **21**(4):391—395
- 辜智慧. 2003. 中国农业复种指数的遥感估算方法研究. 北京: 北京师范大学
- 林忠辉, 莫兴国. 2006. NDVI 时间序列谐波分析与地表物候信息获取. *农业工程学报*, **22**(12):138—144
- 王丹, 姜小光, 唐伶俐, 习晓环. 2005. 利用时间序列傅里叶分析, 重构无云 NDVI 图像. *国土资源遥感*, **64**:29—32
- 张霞. 2006. 光谱指数时间谱特性研究及其在种植模式信息提取中的应用. 北京: 中国科学院遥感应用研究所

Review on methods of remote sensing time-series data reconstruction

LI Ru^{1,2}, ZHANG Xia¹, LIU Bo^{1,2}, ZHANG Bing³

1. Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China;
2. Graduate University of CAS, Chinese Academy of Sciences, Beijing 100039, China;
3. Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, Beijing 100080, China

Abstract: Remote sensing time-series have been applied successfully in various fields, such as vegetation change monitoring, phenological (seasonality) information extraction, land use dynamic classification etc. It is one of the most important data sources for kinds of research work & engineering project. However, due to the effect of sensor, cloud, and atmospheric conditions, there are serious residual noise in time-series data. Therefore, prior to further applications, it is need to filter residual noise to reconstruct the series. Many methods have been developed to solve this problem.

In this paper, the following methods are summarized first, including Maximum Value Composite (MVC), Best Index Slope Extraction Algorithm (BISE), Temporal Windows Operation (TWO), Asymmetric Gaussian Function Fitting Approach (AGFF), Savitzky-Golay Filtering (S-GF), Harmonic Analysis Algorithm (HAA), Local Maximum Fitting (LMF). MVC is more acceptable than other methods because it is useful when producing remote sensing time-series products. But the products are the primary ones and contain much residual noise. This method is helpless to reconstruct the series further. In fact, the data needed to reconstruct before applications of the products made according to MVC. BISE uses a sliding time window to capture local maxima. It requires the determination of the sliding period and a threshold for acceptable percentage increase. TWO can reduce the noise caused by cloud and atmosphere without auxiliary data. However, the requirement of the parameters from experience limits its application. AGFF and S-GF are two strategies developed in recent years. LMF, compared with HAA, first filters noise and then reconstructs the data processed.

Then three most frequently-used approaches, AGFF, S-GF and HAA, are introduced in detail in terms of the basic theories, application steps and advantages & disadvantages. AGFF employs more than two combinative Gauss-shape curves to fit the series. Every combination simulates a cycle of plant life. Finally, these combinative fitting curve are joined together to represent the reconstructed data. By this strategy, every combination is independent. AGFF which is more flexible than other methods could avoid curve distortion. But the parameter, the running window used to find a consistent set of maxima and minima, is difficult to retrieve. Thus it makes the further process less reliable and steadier. S-GF is a process of smoothing and filtering in essence. Two parameters control the result; one is the scale of the smoothing window, and the other is convoluting integer. The two parameters are all obtained depending on experience. Meanwhile, a new variable, fitting-effect index is introduced to control the iteration stop. It is more advanced than the other methods which terminate the iteration by a given threshold. S-GF has received more concern because its product can clearly reflect long time trendy and local change information. Meanwhile its products are less subject to special scale and remote sensor. HAA is a term standing for a methods set which use harmonic to fit data while simulate the seasonal regulation. These methods make use of sine or cosine to fit the data and simulate the seasonal regulation which is much concerned. Two main methods of HAA are Seller Algorithm and HANTS.

At last, some comments are given, discussing the defects of the approaches, and which aspects need to be improved and how to. For example, most of parameters needed in these reconstruction methods have to be decided by experience. Thus subjective errors from operator would affect the reliance and stability of products. So new strategy or improved ones need to solve this problem.

Key words: remote sensing time-series, filtering, reconstruction