

引用格式: Wu Chuanhu, Tao Yuxiang, Luo Xiaobo. Reconstruction of Chongqing's Long Time-series NDVI through an Improved S-G Filter based on Google Earth Engine[J]. Remote Sensing Technology and Application, 2021, 36(5): 1189-1198. [吴川虎, 陶于祥, 罗小波. 基于 Google Earth Engine 的重庆市植被指数长时间序列 S-G 滤波方法的改进与实现[J]. 遥感技术与应用, 2021, 36(5): 1189-1198.]  
DOI: 10.11873/j.issn.1004-0323.2021.5.1189

## 基于 Google Earth Engine 的重庆市植被指数长时间序列 S-G 滤波方法的改进与实现

吴川虎<sup>1,2</sup>, 陶于祥<sup>1,2</sup>, 罗小波<sup>1,2</sup>

(1. 重庆邮电大学计算机科学与技术学院, 重庆 400065;  
2. 重庆邮电大学空间大数据研究中心, 重庆 4000065)

**摘要:** 由于云、大气等因素, 归一化植被指数(Normalized Difference Vegetation Index, NDVI)时间序列数据存在一定的误差。Savitzky-Golay(S-G)滤波方法能够在一定程度上减小这种误差, 抑制突降的低质量像元值, 但对于低质量像元高值的抑制和高质量像元值的保护有所欠缺, 而且不能很好的运用于不同时间间隔的时间序列影像重建中。基于 Google Earth Engine(GEE)云平台, 综合利用空间插值和时间滤波以及像元质量分析对重庆 2014 年春至 2018 年冬 250 m 分辨率的 MOD13Q1 长时间序列数据集进行重建研究。同时使用皮尔逊相关系数(Pearson)、新提出的平滑度指数以及 NDVI 变化差值, 对样本点和整幅影像定量对比重建结果。研究表明: 相同参数下, 新方法的重建影像与原始影像的相关性高于 S-G 方法; 模拟噪声实验中, 其与两幅模拟噪声影像的相关性分别为 0.87 和 0.94, 而 SG 方法的相关性仅为 0.65 和 0.61。

**关键词:** S-G 滤波; 植被指数; 时间序列重建; Google Earth Engine(GEE)

**中图分类号:** TP75; TP79 **文献标志码:** A **文章编号:** 1004-0323(2021)05-1189-10

### 1 引言

来自多源遥感卫星数据的植被指数时间序列数据集在土地利用类型分类<sup>[1-2]</sup>、植被活动监测<sup>[3-4]</sup>、物候信息提取<sup>[5-6]</sup>等领域有着非常广泛的应用, 因此它已经成为了许多科学研究、工程项目的重要数据来源之一<sup>[7]</sup>。然而云、云阴影、大气溶胶、冰雪、视场、数据传输等都会污染遥感植被指数<sup>[8]</sup>。为了解决这些问题, 虽然研究者已经开发了许多处理方法, 并将最大值合成<sup>[9]</sup>(Maximum Value Composite, MVC)应用到了数据集的生产中, 但是最终生成的数据集仍会受到云、云影等的影响, 研究人员仍需要在众多数据中筛选出合适无云影像, 这限制了数

据的进一步应用。因此, 寻找一种可靠的重建算法来处理影像非常重要。

目前, 国内外运用比较广泛的方法有 HANTS<sup>[10]</sup>(Harmonic Analysis of NDVI Time Series)、S-G 滤波<sup>[11]</sup>(Savitzky-Golay Filter, S-G)、非对称高斯拟合法<sup>[12]</sup>(Asymmetric Gaussian model function, AG)、双逻辑拟合法<sup>[13]</sup>(Double Logistic function, DL)。Per Jönsson 和 Lars Eklundh 开发的 TIME-SAT 软件集成了 S-G、AG、DL 3 种算法, 且该软件运用广泛, 尤以 S-G 算法运用最多且效果好<sup>[11, 14-16]</sup>。从 2004 年 Chen 等<sup>[11]</sup>用迭代的 S-G 算法对 NDVI 序列进行滤波开始, 到 2012 年张慧芳等<sup>[17]</sup>的基于背景库的时序数据 S-G 滤波重建, 再到 2013

收稿日期: 2020-07-18; 修订日期: 2021-04-19

基金项目: 国家自然科学基金项目(41871226)。

作者简介: 吴川虎(1996—), 男, 四川巴中人, 硕士研究生, 主要从事植被指数时间序列研究。E-mail: shawfi@163.com

通讯作者: 陶于祥(1966—), 男, 重庆忠县人, 博士, 教授, 主要从事资源与环境经济学研究。E-mail: taoyx@cqupt.edu.cn

和 2019 年周增光等<sup>[18-20]</sup>结合影像质量的 S-G 滤波等,基于 S-G 滤波的重建算法取得了很大进展。然而,这些方法都在追求时间序列曲线的上包络线,容易出现过拟合的现象。实验结合克里金插值方法和质量权重,最大限度地保留了可靠像元,减少过拟合现象的出现。利用皮尔逊相关系数和新的平滑度指数进行对比,找出重庆市范围内二者表现都较好的 SG 滤波算法参数组合。

此外,大范围的长时间序列数据带来了数据量的激增,使用传统的本地环境和数据进行恢复重建会造成大量的数据冗余和时间浪费。而 GEE 是谷歌建立的一款云计算平台,可以直接在线处理 Landsat、MODIS、Sentinel 等卫星影像数据及其衍生数据集。目前,很多人已经使用这个平台来处理和分析数据,并取得了很多的成果<sup>[21-25]</sup>。因此,在带有大型数据库的云计算平台上处理和分析遥感影像已经成为一种新的趋势。但目前重建算法还没有在 GEE 平台上得到有效的应用。

## 2 研究区域概况及数据来源

### 2.1 研究区概况

重庆地处西南,跨 28°10′~32°13′N,105°11′~110°11′E,面积达 8.24 万 km<sup>2</sup>,年日照数 1 000~1 500 h,日照率仅为 25%~35%;其中冬春季日照更少,多为云雾天气。中西部为低山丘陵地带、东南部海拔较高为多山峰、东北部为高峡平湖,导致东南部与中西部全年云量较多,而东北部云量较少。因此对该地区植被指数长时间序列进行恢复重建,具有良好的现实意义。

### 2.2 数据源

实验数据源为 MODIS 全球数据集:MOD13Q1.006 (<https://lpdaac.usgs.gov/products/mod13q1v006/>),其时间分辨率为 16 d、空间分辨率为 250 m。选取 2014 年立春至 2018 年冬末(2014-03-22~2019-03-21),每年 23 景,共计 115 景。使用了 MOD13Q1.006 中的 NDVI 以及 DetailedQA 两个波段。其中 NDVI 值以 16 位有符号整型方式存储,范围从 -2 000 到 10 000,运用式(1)可求得所需 NDVI 值。

$$\text{NDVI} = \frac{\text{DN}}{10000} \quad (1)$$

DetailedQA 以 16 位无符号整型方式存储,在比特位 0~1 处为指数质量指标,具体如表 1 所示。

表 1 MOD13Q1.006 指数质量波段

Table 1 MOD13Q1.006 DetailedQA Band

比特位	数据说明	质量分级
00	指数质量良好,可以直接使用	高质量
01	指数质量一般,需要结合其它信息	中质量
10	有云、雾等干扰	低质量
11	数据丢失	噪声

## 3 研究方法

实验结合数据集的质量波段信息与空间插值和 S-G 滤波提出一种新方法 (Weight-SG, WSG) 从时空上均对被污染的像元值进行恢复重建。具体算法流程如图 1 所示。

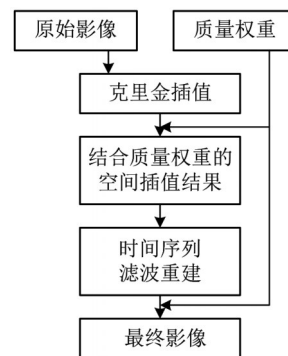


图 1 算法流程图

Fig.1 Flow chart of reconstruction

首先对时间段内的全球影像进行裁剪,仅保留重庆区域;同时统计影像质量波段信息,得出各质量等级占比后再转化为相应的质量权重。然后利用克里金插值法对原始影像进行空间插值,并按其权重对插值结果与原始影像进行合成同时更新权重;接着继续进行 S-G 滤波重建,再次结合质量权重得出最终的影像。

### 3.1 质量等级权重

S-G 滤波算法基于假设<sup>[26]</sup>(云、气溶胶等对 NDVI 值的影响往往倾向削弱 NDVI 值的方向)进行迭代来逼近原始 NDVI 曲线的上包络线从而获得最终结果。但在实际过程中并不完全满足该假设:在低 NDVI 值的背景区域,NDVI 值反而会有所升高,所以盲目的取较大值并不可取。实验结合像元本身的质量等级为其分配不同的权重,继而获得最终的结果如式(2)所示。

$$\text{NDVI}_i^* = \text{NDVI}_i \times w_i + \text{NDVI}_i' \times (1 - w_i) \quad (2)$$

其中:NDVI<sub>i</sub><sup>\*</sup>表示最终值;NDVI<sub>i</sub>表示该像元处理前的值;w<sub>i</sub>表示该像元质量的权重;NDVI<sub>i</sub>'表示该像元通过空间插值或滤波之后得到的拟合值。

### 3.2 空间插值

Tobler地理学第一定律<sup>[26]</sup>指出:任何事物与别的事物之间都是相关的,但近处的事物比远处的事物的相关性更强。S-G滤波不能对不同时间间隔的数据进行处理,所以在进行滤波重建之前对数据进行适当的插值不仅能方便进一步的处理,也能极大的提高中、低质量和噪声像元的可信度。

克里金插值法<sup>[27]</sup>(Kriging/Krige)最早是由法国地理学家 Matheron 和南非矿山工程师 Krige 提出的,主要用于矿山勘探。克里金插值法认为在空间中属性的变化是非常不规则,且不连续的,用简单的平滑函数将出现较多的误差,应该使用能反应空间变异性的变异函数。具体公式如式(3)所示:

$$Z(x_0, y_0) = \sum_{i=1}^n \lambda_i Z(x_i, y_i) \quad (3)$$

其中: $Z(x_0, y_0)$ 是n个采样点值的加权和; $Z(x_i, y_i)$ 是采样点 $(x_i, y_i)$ 处的值; $\lambda_i (i=1, 2, \dots, n)$ 为代求权重系数由变异函数确定。不同的变异函数拟合模型适用于不同的空间环境,常见的变异函数包括指数模型(Exponential)、高斯模型(Gaussian)以及球形模型(Spherical)等。重庆市多山地,地表覆盖类型变化较快,空间自相关性在一定距离范围内消失的也较快,且该距离范围也相对较小,所以使用球型模型更加合适。变异函数模型有3个参数:range是模型的变程,即插值结果的有效范围;sill是基台值,即超出范围后的恒定方差;nugget是块金值,即距离为0是的半方差值。重庆市土地覆盖类型变化较快所以在实验中三者分别设置为:500、1.0和0.1。

重庆市影像质量较差,使用部分像元进行小范围插值时会导致其它区域没有插值结果,因此实验中对全图使用了克里金插值,同时为了减少中、低质量像元对插值结果的影像,在插值后引入质量权重占比对结果进行再次处理。原始影像中高质量像元的值直接代替插值结果,而中、低质量像元按照权重占比来重新合成。同时对结果的质量进行新一轮判定,并在之后的SG滤波中沿用该等级。

### 3.3 改进后的SG滤波算法

S-G滤波是由 Savitzky 和 Golay 率先提出继而广泛应用于平滑时间序列数据的算法<sup>[11]</sup>。S-G滤波是一种基于最小二乘法多项式拟合法,通过一定长度的窗口大小,对待处理数据进行多项式加权拟合,继而求出最小均方根误差。其基本公式如式(4)所示:

$$NDVI_i^* = \frac{\sum_{j=-m}^{j=m} C_j NDVI_{j+i}}{N} \quad (4)$$

其中: $N$ 为窗口大小,满足 $N=2m+1$ ;  $C_j$ 表示S-G多项式拟合的系数。S-G滤波有两个参数: $m$ 是半窗大小, $d$ 是拟合的多项式项数。在陈晋等<sup>[11]</sup>提出的SG滤波中,通过实验指出 $m$ 的推荐取值范围为2~7, $d$ 的取值范围为2~4。实验提出的WSG算法滤波过程中使用的并非是和S-G滤波算法一样的原始遥感影像,而是空间插值结合质量权重后的影像;同时对滤波重建后的结果再次结合质量权重确保尽最大可能保留高质量像元值,而仅拟合中、低质量及噪声的像元NDVI值。

### 3.4 评价指标

#### 3.4.1 样本点

Pearson相关系数是一种常见的统计学系数,用来表示两组变量的相关程度。具体如式(5)所示:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\rho_X \rho_Y} \quad (5)$$

其中: $\text{cov}(X,Y)$ 是变量之间的协方差; $\rho_X \rho_Y$ 是标准差。实验中,用来计算样本点重建前后的相关程度。平滑度是本实验新提出的一种用于评价曲线平滑程度的指标,在时间序列曲线上,锯齿越大越多,曲线就越粗糙,NDVI值产生波动的范围也越大,且产生这种大范围波动的频率也就越高。锯齿在折线图中和极值类似,通过统计该段曲线内极值点的个数可以描述曲线的平滑程度S。在研究区域内,大体为单生长季植物,即在理想情况下:NDVI值从一年的春季开始单调递增,夏季达到最大之后再单调递减,一年内有3个极值点;若为双生长季植物,则一年内有2个波峰3个波谷,即5个极值点。综上所述,定义平滑度S如式(6)所示。

$$S = 1 - \frac{M - (2ny - 1)}{N * y} \quad (6)$$

其中: $M$ 为极值点个数; $y$ 为曲线所跨越的时间长度(按年计); $n$ 为当地所生长植被的生长季个数; $N$ 为数据集每年的影像总数。

#### 3.4.2 单幅图像

为了进一步比较S-G滤波算法以及WSG算法的重建结果,选取2014年春到2018年冬末影像质量较好的5幅,对原始影像引入10%的随机噪声,再对新的数据集分别进行S-G和WSG滤波。由于原始影像的质量较好,所以可以将原始影像作为参考影像,对比两种方法重建后的相关性以及噪声区



域的NDVI值变化,可以定量分析重建结果。

## 4 结果与分析

### 4.1 MODIS MOD13Q1数据集质量分析

利用MOD13Q1自带的指数质量波段,对5 a内的图像进行分析,如图2所示。结果表明:5 a内中质量像元占比数量最多达到了40.67%,而高质量和低质量像元分别占比30.36%和28.95%,噪声像元

仅占0.02%。2018年冬季高质量像元占比最低,仅为4.78%;2017年春季高质量像元占比最高,为61.82%。整体呈现出春夏季节高质量数据多而秋冬季节的低质量数据多的特征,从春季开始高质量数据占比逐减降低直到冬季达到最低。其中2014年春季质量高的像元占比较少,和上文提到的情况有所出入,详细浏览数据发现是由于该时间段内卫星数据缺失严重所引起的。

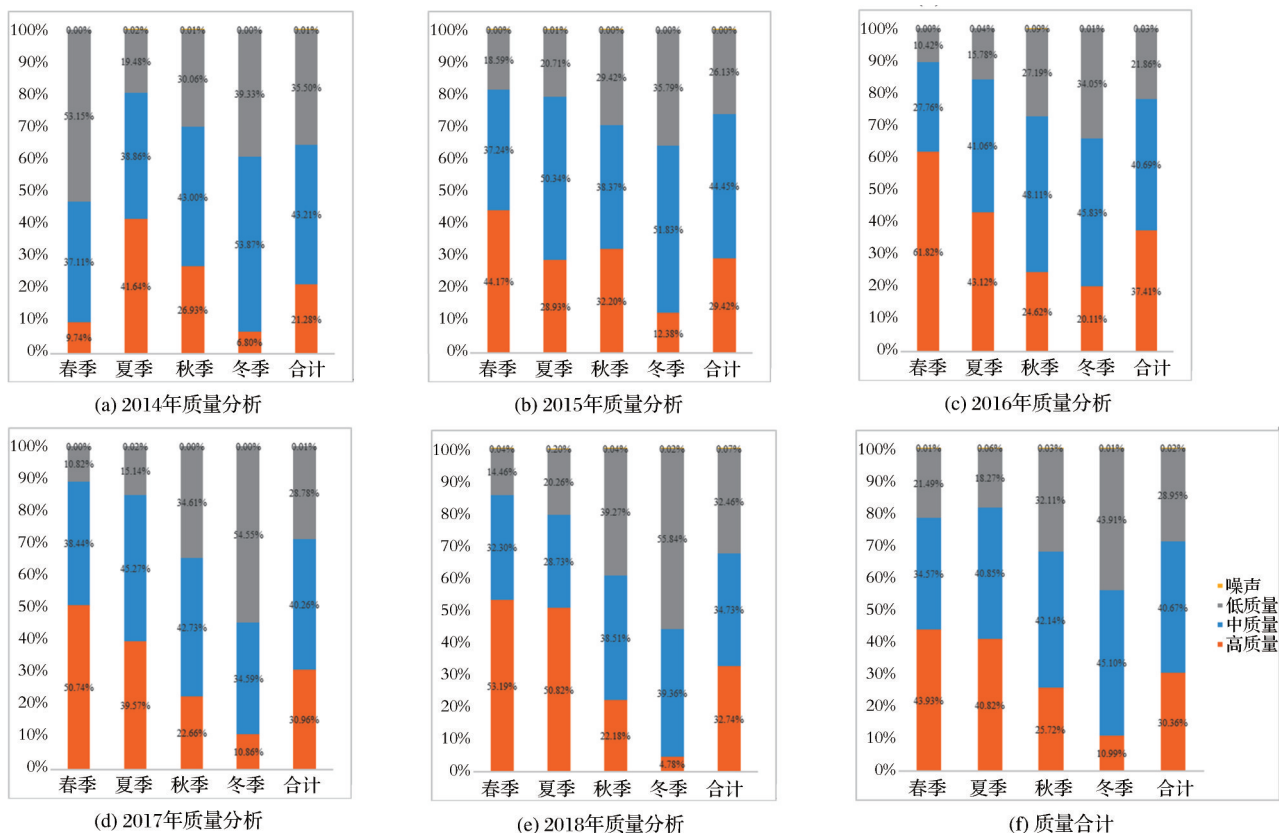


图2 质量分析图

Fig.2 Quality analysis chart

从整体上看中质量像元占比在40.67%,低质量像元占比在28.95%;所以将质量权重分别设置为0.4和0.3,具体的质量等级与权重转换表如表2所示。

### 4.2 空间插值结果分析

为了降低中、低质量与噪声像元对滤波重建结

果的影响,实验使用空间插值法对图像进行处理来填充缺失值,提高像元的可信度,并选取时间段内第一幅影像(时间:2014年3月22日)来说明插值结果。

#### 4.2.1 整体结果

插值后的结果如图3所示。图3(a)是原始图像,重庆东南部、中西部的NDVI值较低,整体受云雾影响较为严重。图3(b)是克里金插值后的影像,云污染区域面积有所减少,但大面积污染区域中心NDVI值仍旧较低;另外,插值后的结果有明显的块效应和边缘断层现象。而结合了质量权重的空间插值如图3(c)所示,相较于图3(b)块效应有所减少。

#### 4.2.2 局部细节

从图3可以看到,对中低质量像元进行插值后

表2 质量等级与权重

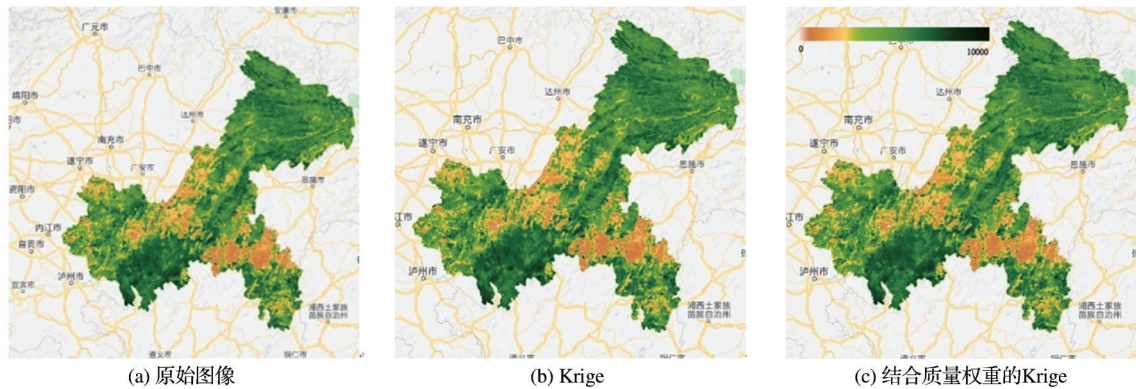
Table 2 Quality level and its weight

质量等级	权重大小
高	1
中	0.4
低	0.3
噪声	0

的影像部分区域仍旧污染严重,以重庆主城区为例,如图 4 所示。

对图 4(a)进行目视解译可知:蜿蜒的白色区域为长江、嘉陵江交汇处,周围橙色区域为主城区。图 4(b)、图 4(c)都能保持主城区的大概轮廓,但图 4(c)

比图 4(b)保留了更多细节,图 4(b)具有较为严重的块效应,但图 4(c)没有。小范围区域内,图 4(b)更倾向于同质化,而图 4(c)的 NDVI 值变化更加丰富,但仍是较为均匀的。如图 5 所示,是使用克里金插值法处理原始图像后各等级质量像元的分布情况。



审图号:GS(2019)3333

图 3 插值结果对比

Fig.3 Interpolation result graph

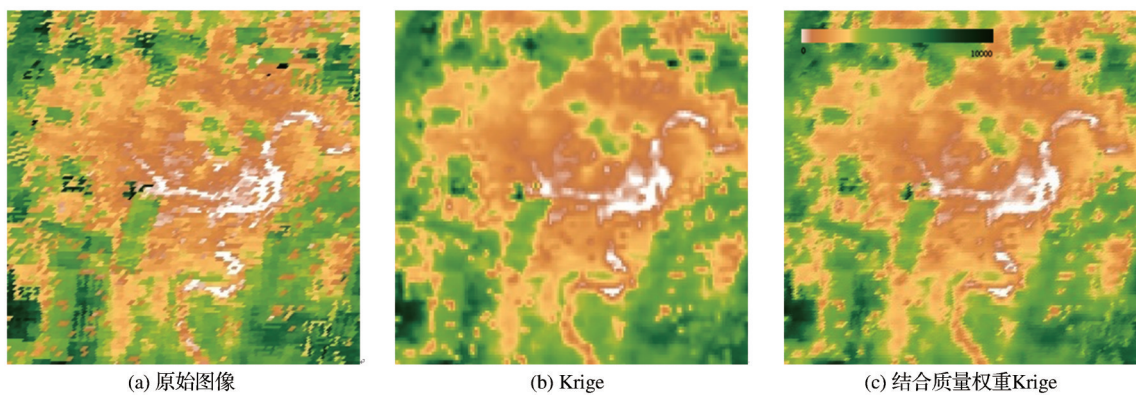
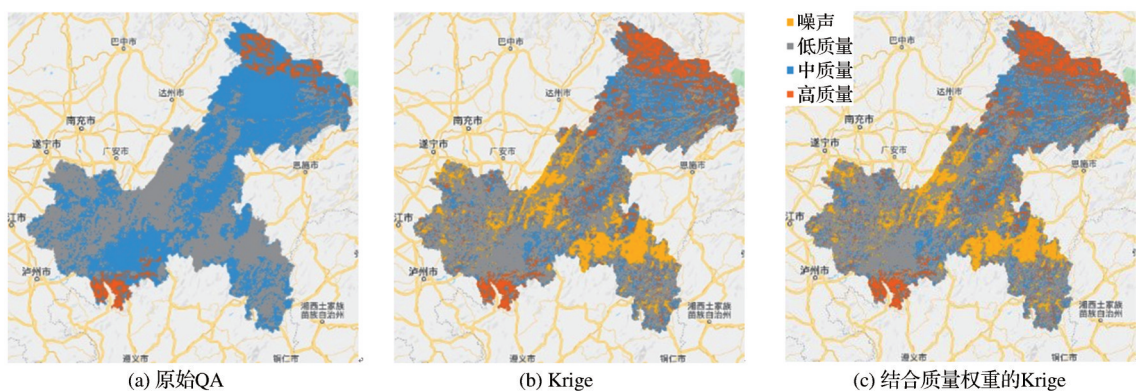


图 4 主城区结果对比

Fig.4 Comparison of results in the main city



审图号:GS(2019)3333

图 5 QA 质量等级图

Fig.5 QA quality detailed

原始影像中,高质量和噪声像元较少,中、低质量像元较多。而经过插值处理后的质量图 5(b):高

质量像元有少量增加,中质量像元大幅减少,低质量像元略有增加,噪声像元也大幅增加;图 5(c)与



图 5(b)变化情况相同,仅在变化幅度上有所差异。通过插值可以重建部分云污染区域,但仍有像元需要进一步处理。所以对插值后的影像进行滤波,从时间维度上提取更多的信息进行重建是有必要的。

### 4.3 滤波结果分析

#### 4.3.1 参数分析

S-G 算法半窗大小  $m$  一般取值范围在 2~7、多项式项数  $d$  在 2~4。为了探寻不同参数组合的 S-G 滤波在重庆市范围的不同结果,实验对图 6 所示的 50 个随机样本点,以及 18 种参数组合,做了共计 900 组实验。选取不同影像质量不同程度的 3 个样本点进行说明,如图 6 中红点所示:其中样本点 17 位于重庆东南部,影像质量一般,样本点 26 位于西部影像质量较差,样本点 39 位于东北部影像质量较好。



图 6 样本点分布

Fig.6 Sample points distribution

S-G 滤波结果的 NDVI 曲线相关性均值为 0.77、平滑度均值为 0.75,其中结果高于相关性均值的实验有 496 组,高于平滑度均值的实验有 459 组,高于均值的实验参数分布情况如表 3 所示。

表 3 参数分布情况

Table 3 Parameter distribution

相关性				平滑度			
$m$	个数	$d$	个数	$m$	个数	$d$	个数
2	132	2	199	2	0	2	177
3	105	3	199	3	36	3	177
4	83	4	225	4	74	4	105
5	70			5	97		
6	60			6	120		
7	46			7	132		

由结果分析可知:随着  $m$  的增加,相关性越低、平滑度越高;随着  $d$  的增加,相关性越高、平滑度越低。所以本实验将后续的 S-G 滤波以及 WSG 滤波参数  $m$ 、 $d$  均设置为 5 和 3。

#### 4.3.2 滤波结果

对 5 a 内的数据进行滤波后,结果如图 7 所示。样本点的时间序列曲线能直观反映各样本点的影像质量:样本点 17,波动较大、但产生大波动的频率不如样本点 26;样本点 26 波动小于 17,但产生大波动的频率远远高于 17;而样本点 39 大波动较少,频率也较低。完全符合选取的样本点实际情况。3 个样本点中,S-G 方法均存在一定程度上的低估或高估,为了平滑丢失了部分细节,虽然平滑度较高,但相关性很低。尤其以样本点 26 最为严重,因为影像波动过于频繁,导致拟合值与原始值差距较大。而 WSG 方法虽然平滑度不如 S-G 但高于原始曲线,且 WSG 相关性高于 S-G。即 WSG 方法的相关性很高,对原始 NDVI 序列的改变较小;S-G 滤波结果具有一定的过拟合现象。具体 3 个样本点的对比如表 4 所示。

表 4 样本点的相关性与平滑度

Table 5 Correlation and smoothness of three sample points

样本点	SGPearson	WSGPearson	原始平滑度	SG平滑度	WSG平滑度
17	0.53	0.72	0.52	0.77	0.56
26	0.71	0.76	0.52	0.82	0.51
39	0.69	0.82	0.44	0.76	0.55

依旧选取时间段内第一幅影像进行说明,图 8(a)、图 8(b)分别是 SG 滤波和 WSG 滤波结果,与图 3(a)的原始影像相比二者都重建了较大范围,但 WSG 滤波比 S-G 保留了更多的细节,且没有出现明显的过拟合现象。虽然从直观上 S-G 滤波结果看起来更好,植被区域更广阔,但是其与原始影像的相关系数仅为 0.62,数据的可信度大大降低,并且低 NDVI 值的建筑区域存在较为严重的过拟合现象;而 WSG 结果的相关系数为 0.78,比 S-G 滤波更加贴近实际情况。图 8(c)、图 8(d)分别是其对应的主城区,S-G 滤波结果更复杂,在同一地物的建筑区域,NDVI 变化频繁,且范围过大不符合实际情况,但河流的轮廓比 WSG 清晰一些;WSG 滤波结果在建筑区域,NDVI 变化更加均匀,恢复结果略好于 S-G。

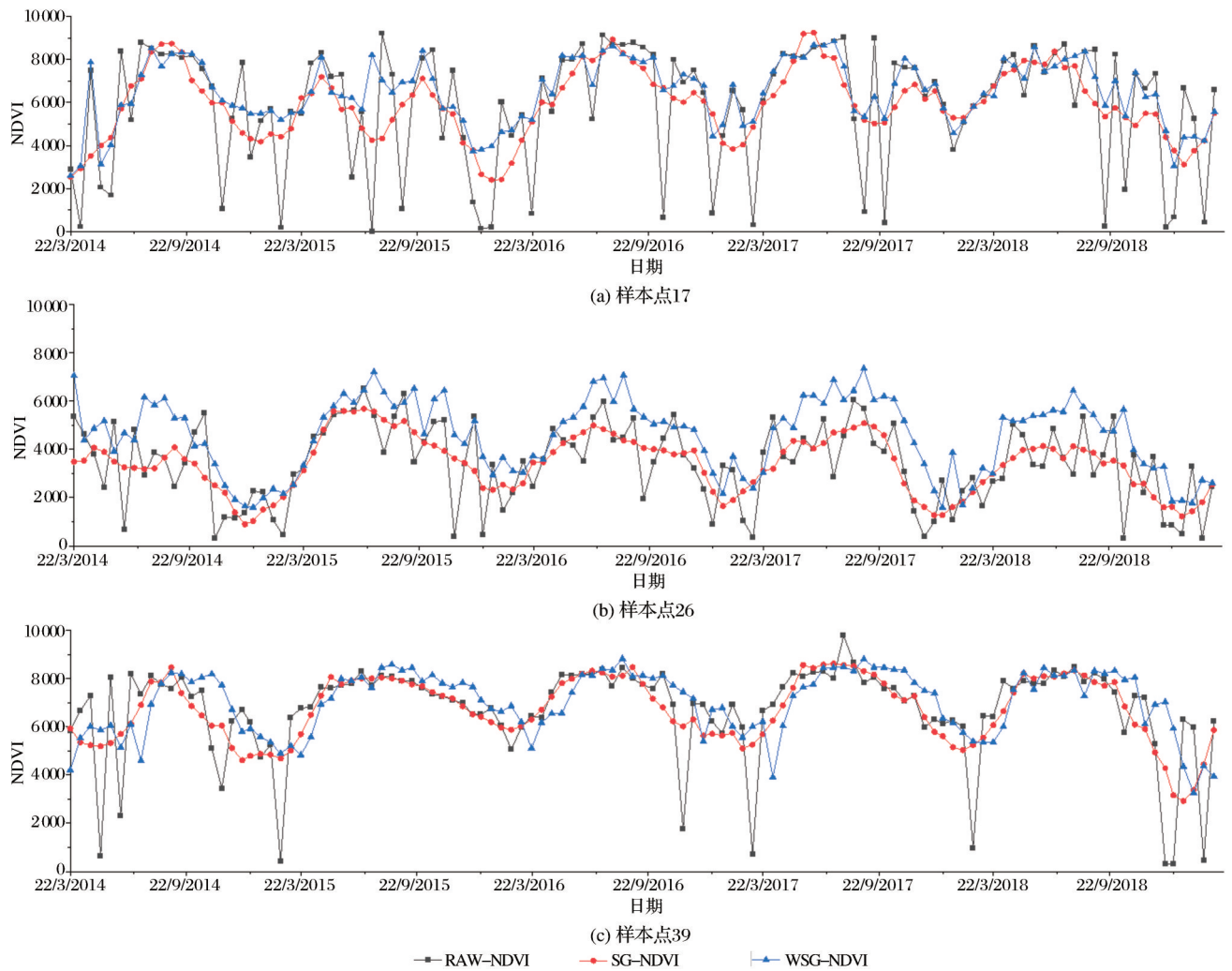
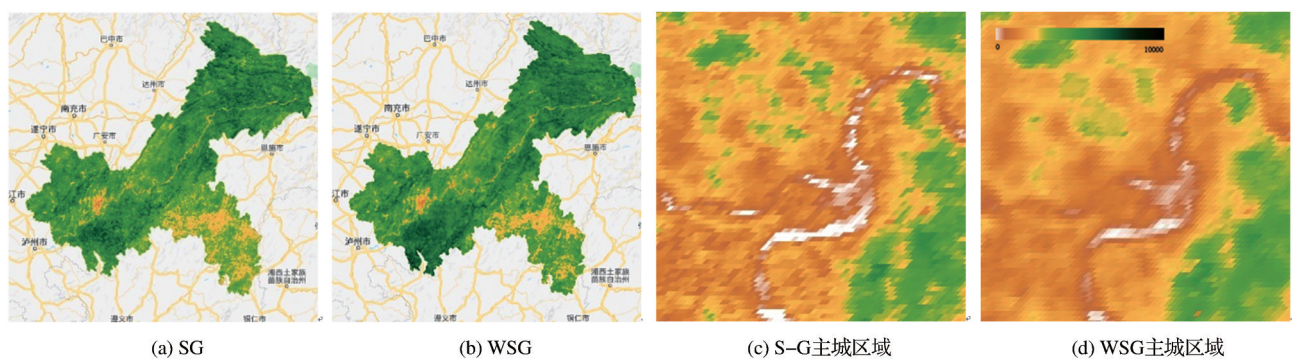


图 7 3 个样本点的重建结果

Fig.7 Three sample points details



审图号:GS(2019)3333

图 8 滤波重建结果

Fig.8 Filtered reconstruction results

#### 4.3.3 人工插入噪声

实验选取 2016 年 5 月 24 日与 2017 年 4 月 7 日两幅影像来进行说明,两幅原始影像添加了 10% 的随机噪声后如图 9 所示。

通过 S-G 滤波以及 WSG 滤波后的结果如图 10

所示:从直观上来讲,WSG 重建后的结果比 S-G 更好,有效的重建了噪声区域的 NDVI 值。而 S-G 的重建结果整体 NDVI 值偏高较为严重,且能明显的观察到噪声像元存在过的痕迹。对二者进行相关性分析,S-G 的 Pearson 相关系数分别为 0.65 和



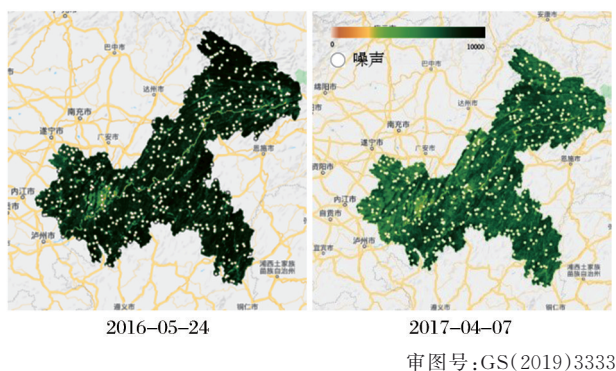


图9 模拟噪声位置分布图

Fig.9 Image with 10% noise

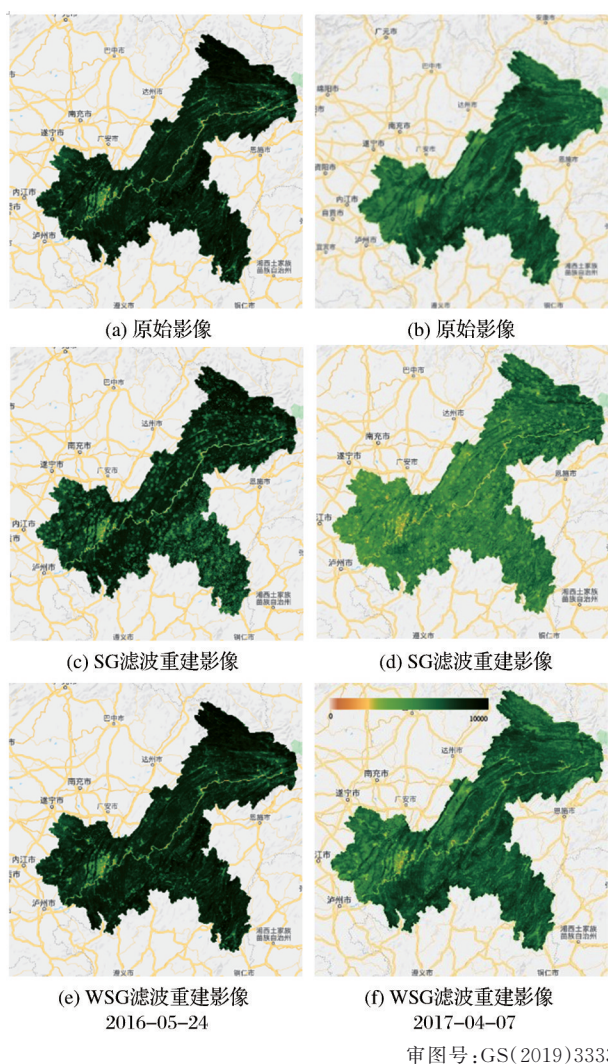
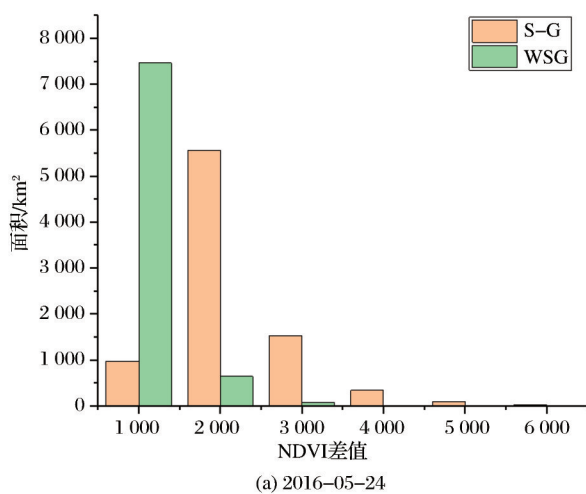


图10 重建结果

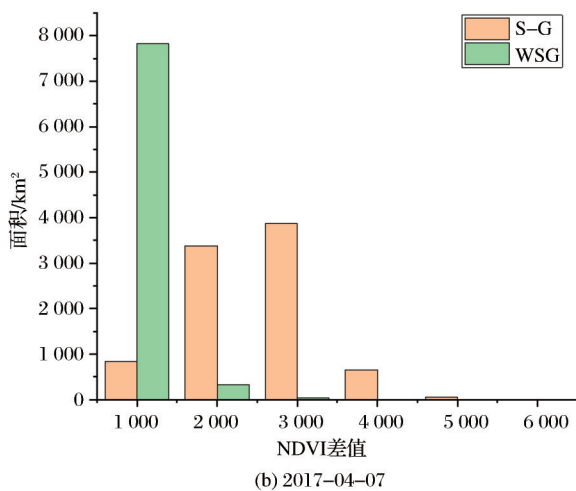
Fig.10 Reconstructed image

0.61, 而 WSG 为 0.87 和 0.94 明显高于 S-G 的相关性, 重建结果在视觉上更接近于真实影像。

为了更直观地展示重建结果, 对比原始影像和重建影像中噪声部分的 NDVI 值变化, 具体如图 11 所示(为方便展示, NDVI 均乘以一万)。



(a) 2016-05-24



(b) 2017-04-07

图11 NDVI 差值

Fig.11 NDVI Difference

由图 11 可以明显看出, 在人为添加的噪声区域中, WSG 重建结果与真实影像更加接近, 差值集中在 1 000 以内, 少部分在 1 000~2 000, 超过 3 000 的很少; 而 S-G 重建结果相对较差, 大部分在 1 000~3 000, 最大值甚至接近 6 000。所以, 整体来讲, WSG 方法的恢复重建效果优于 S-G。

## 5 结论

### 5.1 结论

实验基于 GEE 云平台实现了经典 S-G 滤波算法, 并在其基础上, 结合空间插值法和质量权重提出了 WSG 方法, 进一步在参数不变的情况下提升了重建结果的相关性。同时提出了一个平滑度指标来描述时间序列曲线的平滑程度。通过对研究结果进行分析主要得出以下结论:

(1) 使用空间插值法对原始影像中的低质量及噪声像元进行预处理后, 能够提升滤波重建结果的



相关性,但要尽量减小因插值带来的块效应。S-G滤波算法和WSG都能够较好地重建被云雾覆盖等原因污染的NDVI值,同参数下S-G滤波在相关性上低于WSG,在平滑度上略高于WSG。

(2)GEE云平台能够完成一部分的空间信息处理与分析工作,降低了遥感信息处理的门槛,但相较于传统工具(ArcGIS、QGIS、ENVI等)还有一定的差距。本研究使用时遇到的主要问题体现在:①分配给用户的云计算资源略有不足;②没有较为完善的制图、出图工具;③基于C/S架构,用户运行的结果往往是异步的,如果使用较多的循环等会导致运行超时网页无响应;④对矢量文件的运算处理缺少优化,计算复杂边界时会导致运算量加大。

## 5.2 展望

时间序列数据重建对于生成更加可靠的数据,进行下一步研究具有重大意义,结合云计算可以在减少数据收集、预处理时间的同时加快重建算法的运行效率。且随着越来越多的科研人员加入与技术的逐步发展,时间序列重建甚至可以在做到与原始数据一样的更新频率的情况下获得质量更好的影响,为后续的土地利用类型分类、植被覆盖度分析等奠定良好的基础。

## 参考文献(References):

- [1] Chen Jin, Chen Yunhao, He Chunyang, *et al.* Sub-pixel model for vegetation fraction estimation based on land cover classification[J]. *Journal of Remote Sensing*, 2001, 5(6): 416-422.[陈晋, 陈云浩, 何春阳, 等. 基于土地覆盖分类的植被覆盖率估算亚像元模型与应用[J]. *遥感学报*, 2001, 5(6): 416-422.]
- [2] Shao Y, Lunetta R S, Wheeler B, *et al.* An evaluation of time-series Smoothing Algorithms for land-cover classifications using MODIS-NDVI multi-temporal data[J]. *Remote Sensing of Environment*, 2016, 174: 258-265.
- [3] Atzberger C, Eilers P H. A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America[J]. *International Journal of Digital Earth*, 2011, 4(5): 365-386.
- [4] Stöckli R, Vidale P L. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset[J]. *International Journal of Remote Sensing*, 2004, 25(17): 3303-3330.
- [5] Esau I, Miles V V, Davy R, *et al.* Trends in Normalized Difference Vegetation Index (NDVI) associated with urban development in northern west Siberia[J]. *Atmospheric Chemistry and Physics*, 2016, 16(15): 9563-9577.
- [6] Mao J, Ribes A, Yan B, *et al.* Human-induced greening of the northern extratropical land surface[J]. *Nature Climate Change*, 2016, 6(10): 959.
- [7] Lunetta R S, Joseph F K, Jayantha E, *et al.* Land-cover change detection using multi-temporal MODIS NDVI data[J]. *Remote sensing of environment*, 2006, 105(2): 142-154.
- [8] Li Ru, Zhang Xia, Liu Bo, *et al.* Review on methods of remote sensing time-series data reconstruction[J]. *Journal of Remote Sensing*, 2009, 13(2): 335-341.[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.]
- [9] Holben B N. Characteristics of maximum-value composite images from temporal AVHRR data[J]. *International Journal of Remote Sensing*, 1986, 7(11): 1417-1434.
- [10] Lovell J, Graetz R. Filtering pathfinder AVHRR land NDVI data for australia[J]. *International Journal of Remote Sensing*, 2001, 22(13): 2649-2654.
- [11] Chen J, 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(3-4): 332-344.
- [12] Jonsson 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): 1824-1832.
- [13] Ma M, Veroustraete F. Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China[J]. *Advances in Space Research*, 2006, 37(4): 835-840.
- [14] Song Chunqiao, Ke Linghong, You Songcai, *et al.* Comparison of three NDVI time-series fitting methods based on TIMESAT-Talking the grassland in Northern Tibet as case[J]. *Remote Sensing Technology and Application*, 2011, 26(2): 147-155.[宋春桥, 柯灵红, 游松财, 等. 基于TIMESAT的3种时序NDVI拟合方法比较研究——以藏北草地为例[J]. *遥感技术与应用*, 2011, 26(2): 147-155.]
- [15] Cheng Linlin, Li Yuhu, Sun Haiyuan, *et al.* Applicability of fitting and reconstruction method of MODIS long-time enhanced vegetation index in Beijing-Tianjin-Hebei[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(11): 148-158.[程琳琳, 李玉虎, 孙海元, 等. 京津冀MODIS长时序增强型植被指数拟合重建方法适用性研究[J]. *农业工程学报*, 2019, 35(11): 148-158.]
- [16] Chen Yang. Research on medium and high spatial resolution[D]. Chengdu: University of Electronic Science and Technology of China, 2019.[陈洋. 中高空间分辨率卫星NDVI时间序列数据重建技术研究[D]. 成都: 电子科技大学, 2019.]
- [17] Zhang Huifang, Gao Wei, Shi Runhe. Reconstruction of high-quality LAI time-series product based on long-term historical database[J]. *Journal of Remote Sensing*, 2012, 16(5): 986-999.[张慧芳, 高伟, 施润和. 基于背景库的高质量LAI时间序列数据重建[J]. *遥感学报*, 2012, 16(5): 986-999.]
- [18] Zhou Zengguang, Tang Ping. VI-Quality-Based Savitzky-Golay method for filtering time series data[J]. *Remote Sensing Technology and Application*, 2013, 28(2): 232-239.[周增光, 唐婷. 基于质量权重的Savitzky-Golay时间序列滤波方法[J]. *遥感技术与应用*, 2013, 28(2): 232-239.]

- [19] Zhou Minyue, Shen Runping, Chen Jun, *et al.* Filtering and reconstruction of LAI time series data by filter based on pixel quality analysis and outlier detection [J]. *Remote Sensing Technology and Application*, 2019, 34(2): 323-330. [周旻悦, 沈润平, 陈俊, 等. 基于像元质量分析和异常值检测的 LAI 时序数据 S-G 滤波重建研究[J]. *遥感技术与应用*, 2019, 34(2): 323-330.]
- [20] Hu Shunshi, Huang Chunxiao, Yang Bin, *et al.* Reconstruction of MODIS vegetation index time series by adaptive weighted Savitzky-Golay filter[J]. *Science of Surveying and Mapping*, 2020, 45(4): 105-116. [胡顺石, 黄春晓, 杨斌, 谭子芳. 自适应加权 Savitzky-Golay 滤波重构 MODIS 植被指数时间序列[J]. *测绘科学*, 2020, 45(4): 105-116.]
- [21] Pei Jie, Niu Zheng, Wang Li, *et al.* Monitoring to variations of vegetation cover using long-term time series remote sensing data on the Google Earth Engine cloud platform[J]. *Carsologica Sinica*, 2018, 37(4): 608-616. [裴杰, 牛铮, 王力, 等. 基于 Google Earth Engine 云平台的植被覆盖度变化长时间序列遥感监测[J]. *中国岩溶*, 2018, 37(4): 608-616.]
- [22] Long Shuang, Guo Zhengfei, Xu Li, *et al.* Spatiotemporal variations of fractional vegetation coverage in China based on Google Earth Engine [J]. *Remote Sensing Technology and Application*, 2020, 35(2): 326-334. [龙爽, 郭正飞, 徐粒, 等. 基于 Google Earth Engine 的中国植被覆盖度时空变化特征分析[J]. *遥感技术与应用*, 2020, 35(2): 326-334.]
- [23] Li Peilin, Liu Xiaoping, Huang Yinghui, *et al.* Mapping impervious surface dynamics of Guangzhou downtown based on Google Earth Engine [J]. *Journal of Geo-information Science*, 2020, 22(3): 638-648. [李培林, 刘小平, 黄应淮, 等. 基于 GEE 平台的广州市主城区不透水面时间序列提取[J]. *地球信息科学学报*, 2020, 22(3): 638-648.]
- [24] Kong D, Zhang Y, Gu X, *et al.* A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 155: 13-24.
- [25] Park J, Tateishi R. Correction of time series NDVI by the method of temporal window operation[C]// 19th Asian Conference on Remote Sensing, Manila, 1998.
- [26] Tobler, Waldo R. A computer movie simulating urban growth in the detroit region [J]. *Economic Geography*, 1970, 46: 234-240.
- [27] Oliver M A, Webster R. A tutorial guide to geostatistics: Computing and modelling Variograms and Kriging [J]. *Catena*, 2014, 113: 56-69.

## Reconstruction of Chongqing's Long Time-series NDVI through an Improved S-G Filter based on Google Earth Engine

Wu Chuanhu<sup>1,2</sup>, Tao Yuxiang<sup>1,2</sup>, Luo Xiaobo<sup>1,2</sup>

(1. Department of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Spatial Big Data Research Center, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** Due to factors such as clouds and atmosphere, there are certain errors in the Normalized Difference Vegetation Index (NDVI) time series data set. The Savitzky-Golay (S-G) filtering method can reduce this error to a certain extent and suppress the sudden drop of low-quality pixel values, but it is lacking in the suppression of high-value low-quality pixels and the protection of high-quality pixel values. , And cannot be used well in time series image reconstruction at different time intervals. Based on the Google Earth Engine (GEE) cloud platform, a comprehensive use of spatial interpolation, temporal filtering, and pixel quality analysis to reconstruct the 250m resolution MOD13Q1 long-term data set in Chongqing from spring 2014 to winter 2018. At the same time, the Pearson correlation coefficient (Pearson), the newly proposed smoothness index and the difference of NDVI change are used to quantitatively compare the reconstruction results of a single sample point and a single image. Research shows that under the same parameters, the correlation between the time series reconstructed by the new method and the original image is higher than that of the S-G method; in the simulated noise experiment, the correlation between it and the two simulated noise images are 0.87 and 0.94, respectively, while the correlation of the S-G method is only 0.65 and 0.61.

**Key words:** S-G filter; Vegetation Index; Time Series Reconstruction; Google Earth Engine