

耦合空域重建的植被指数时空融合方法

唐以洁, 王群明

同济大学 测绘与地理信息学院, 上海 200092

摘要: 高时空分辨率归一化植被指数 NDVI (Normalized Difference Vegetation Index) 数据对于作物生长监测及参数反演等极其重要。卫星遥感数据的红波段和近红外波段可作为 NDVI 的重要数据源。然而, 受卫星传感器功率所限, 所获取遥感数据的空间分辨率和时间分辨率之间通常存在相互制约, 难以获得兼具高时空分辨率的 NDVI 数据。时空融合技术旨在融合具有高时间分辨率、但空间分辨率较低的 NDVI 数据和具有高空间分辨率、但时间分辨率较低的 NDVI 数据, 以生成兼具高时空分辨率 NDVI 数据。然而, NDVI 时空融合面临一较大挑战, 即已知辅助数据所在时刻与预测时刻之间, 地表覆盖通常呈现较大变化。因此, 为应对上述挑战, 本文提出了一种耦合空域重建的时空融合方法 STFSR (Spatio-Temporal Fusion then Spatial Reconstruction) 方法, 通过充分利用在时间上接近预测时刻、但含数据缺失 (遥感影像云覆盖导致) 的高空间分辨率数据来辅助 NDVI 时空融合预测。实验选取了位于法国巴黎、中国河北省和澳大利亚新南威尔士州的 3 个实验区的 Landsat 和 MODIS NDVI 数据, 并与已有的时空自适应反射率融合模型 STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model) 和基于空间加权和虚拟图像对的时空融合 VIPSTF-SW (Spatial Weighting-based Virtual Image Pair-based Spatio-Temporal Fusion) 方法对比, 进行验证。结果表明: 时域近邻辅助云影像的利用, 有效减小了 NDVI 变化对时空融合的影响; 在本研究所选取的 3 个实验区中, 所提出的 STFSR 方法被证明比常用的 STARFM 模型和 VIPSTF-SW 方法, 具有更高的精度。3 个实验区中 STFSR 方法的平均均方根误差 RMSE (Root Mean Square Error) 比 STARFM 和 VIPSTF-SW 分别降低了 0.0217 和 0.0188, 平均相关系数 CC (Correlation Coefficient) 分别提升了 0.0820 和 0.0742, 相对全局综合误差 ERGAS (Relative Global-dimensional Synthesis Error) 分别降低了 4.3170 和 3.8535。此外, 当辅助数据云区域面积增大时, STFSR 方法虽在精度上呈现下降态势, 但整体上仍优于 STARFM 和 VIPSTF-SW 方法。STFSR 方法为高时空分辨率 NDVI 数据的生成提供了一种新的思路, 且鉴于其原理与优势, 该模型于其他高时空分辨率植被指数数据, 例如增强型植被指数的生成亦具有可观的应用潜力。

关键词: 遥感, Landsat, MODIS, NDVI, 时空融合, 空域重建

中图分类号: TP75/P2

引用格式: 唐以洁, 王群明. 2025. 耦合空域重建的植被指数时空融合方法. 遥感学报, 29(8): 2602-2615

Tang Y J and Wang Q M. 2025. Spatial reconstruction-aided spatio-temporal fusion of NDVI images. National Remote Sensing Bulletin, 29(8): 2602-2615 [DOI: 10.11834/jrs.20254449]

1 引言

归一化植被指数 NDVI (Normalized Difference Vegetation Index) 是一种评估植被覆盖率和生长状况的常用指数, 且在估算植被生产力和碳储量等指标发挥着重要作用 (Chen 等, 2021; Chu 等, 2021)。大尺度的 NDVI 数据通常由遥感卫星数据的红波段和近红外波段计算获得。然而, 由于卫星传感器功率的限制, 遥感数据的时间分辨率和

空间分辨率之间通常存在相互制约。具体来讲, 具有较高空间分辨率的遥感影像通常具有较低的时间分辨率。例如, Landsat 8 OLI (Operational Land Imager) 和 Landsat 9 OLI2 影像的空间分辨率为 30 m, 然而其重返周期却长达 16 d (Cao 等, 2020; 王群明 等, 2023; Xu 等, 2023)。相反, Terra 和 Aqua 卫星上搭载的 MODIS (Moderate-resolution Imaging Spectroradiometer) 传感器时间分辨率小于 1 d, 但其用于 NDVI 生成的波段空间分

收稿日期: 2024-10-09; 预印本: 2025-03-14

基金项目: 国家自然科学基金 (编号: 42222108, 42171345)

第一作者简介: 唐以洁, 研究方向为遥感数据时空融合。E-mail: tangyijie@sina.com

通信作者简介: 王群明, 研究方向为遥感信息智能处理。E-mail: wqml1111@126.com

分辨率却仅有 250 m (张智昊 等, 2024)。为应对区域尺度下精细、实时监测(如作物生长、土地变化监测)的需求,通常需要兼具高时空分辨率(如 30 m、每天)的时序 NDVI 数据。

遥感数据时空融合技术通过融合具有高空间分辨率(如 30 m、16 d 的 Landsat)的 NDVI 数据和具有高时间分辨率(如 250 m、每天的 MODIS)的 NDVI 数据,获得兼具高时空分辨率(如 30 m、每天)的 NDVI 数据。近二十年来,时空融合技术取得了一定程度的发展(Zhu 等, 2018; 董文全和蒙继华, 2018; 刘建波 等, 2016)。现有的时空融合方法可分为 4 大类:(1) 基于权重函数的方法。时空自适应反射融合模型 STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model) (Gao 等, 2006) 是时空融合领域中较为经典且广泛使用的一种基于权重函数的方法。STARFM 通过一个综合考虑高、低分辨率影像之间空间、时间、光谱差异信息的加权函数对待预测时刻的高分辨率影像进行估算。针对该模型中存在的局限性不断改进与发展,自适应遥感图像时空融合变化检测方法(Hilker 等, 2009)、增强型时空自适应反射率融合模型 ESTARFM (Enhanced STARFM) (Zhu 等, 2010)、Fit-FC (Wang 和 Atkinson, 2018) 也相继被提出。(2) 基于空间解混的方法。这类方法主要是利用光谱解混模型,建立端元丰度、端元值和低分辨率影像灰度值之间的关系,求解类别反射率,并按照类别赋值给待求解高分辨率影像的各个像元(Amorós-López 等, 2013; Busetto 等, 2008; Gevaert 和 García-Haro, 2015; Maselli, 2001; Wu 等, 2012; Xu 等, 2015; Zurita-Milla 等, 2011)。最早的空间解混方法为多传感器多分辨率技术(MMT) (Zhukov 等, 1999), 后期的基于空间解混的方法主要是通过对原始解混模型加入不同的限制条件(Gevaert 和 García-Haro, 2015)以提升预测效果。(3) 基于学习的方法。这类方法主要是通过不同的学习模型来构建高低分辨率影像之间的映射关系。目前,稀疏表示模型(Huang 和 Song, 2012; Song 和 Huang, 2013)、随机森林模型(Ke 等, 2016)、深度学习模型(Das 和 Ghosh, 2016; Song 等, 2018) 等已被用于遥感数据的时空融合。(4) 综合方法。近年来,同时结合基于空间解混和权重函数两类方法的综合方法也得到了发展。其原理主要是先利用线性解混

模型求解类别反射率,并通过基于权重函数的方法进行残差补偿。该类方法中较为经典且常用的是自适应数据融合算法 FSDAF (Flexible Spatiotemporal Data Fusion) (Zhu 等, 2016)。基于 FSDAF 方法进行改进和发展,不同的综合方法也相继被提出(Guo 等, 2020; Li 等, 2020; Liu 等, 2019)。此外,另有研究通过挖掘时序植被指数数据构建物候学模型,用于生成高时空分辨率植被指数时序数据。例如:基于实际观测的气候学一致性调适模型 CACAO (Consistent Adjustment of the Climatology to Actual Observations) (Verger 等, 2013; Yin 等, 2017) 和减化比值植被指数模型 RSR (Reduced Simple Ratio) (Zhang 等, 2014)。现有的时空融合方法大多基于地表反射率数据提出,但其方法基本都适用于其他数据的融合(如地表温度、NDVI 等)。

尽管时空融合方法获得了迅速发展,但至今仍存在一因数据可获取性差而带来的核心挑战。通常,时空融合需要两个时刻的数据参与,即高空间分辨率数据缺失的待预测时刻和另一个高、低分辨率都存在的已知时刻。由于需要预测完整的待预测时刻高空间分辨率数据,在选取数据时,通常要求已知时刻的高、低分辨率数据具有空间完整性。然而,由于频繁的云覆盖,很难获取空间完整的已知时刻辅助数据。因此,通常获得的空间完整的辅助数据与待预测时刻之间存在较长的时间间隔。由于地物随时间不断发生变化,当两时刻数据之间的时间间隔较大时,两者之间的 NDVI 变化也会增大,给时空融合带来较大的挑战。事实上,距离待预测时刻较近,但存在部分空域缺失的 NDVI 数据(由于遥感影像的云覆盖)可为待预测时刻高空间分辨率影像的预测提供重要的辅助信息。然而,现有的时空融合方法往往忽略了这一点。

因此,为了解决 NDVI 数据时空融合中时域近邻、空间完整辅助数据可获取性差的局限性,本文基于时域近邻、空间缺失辅助数据的充分利用,提出了先时空融合后空域重建 STFSR (Spatio-Temporal Fusion then Spatial Reconstruction) 方法,旨在耦合时间近邻(记为 t_m 时刻)、空域缺失辅助影像的剩余有效信息,在已知的完整 NDVI 数据(记为 t_k 时刻)和待预测时刻(记为 t_p 时刻) NDVI 数据相隔较远的情况下,进行更为可靠的时空融

合预测。STFSR 方法首先利用时空融合方法预测 t_p 时刻高空间分辨率 NDVI 数据部分区域（对应 t_m 时刻高空间分辨率 NDVI 数据的无云区域），再利用 STRF（Spatial-Temporal Random Forest）方法对剩余区域的高空间分辨率 NDVI 数据进行预测。STFSR 方法没有 t_m 时刻低空间分辨率数据空间完整性的要求，对于地表反射率、NDVI 数据等场景的时空融合也同样适用。其次，在利用较远时刻（ t_k ）去重建 t_p 时刻高空间分辨率 NDVI 数据时，同时有 t_p 时刻高空间分辨率 NDVI 数据部分辅助数据及两时刻对应低空间分辨率 NDVI 数据的加入，增强了当两时刻相隔较远时进行预测的可靠性。总结来说，所提出的 STFSR 方法为云覆盖影像中高空间分辨率辅助信息的充分利用提供了一有效的方案，克服了已知完整数据和待预测时刻之间 NDVI 变化较大带来的影响。

2 研究方法

以 Landsat 和 MODIS NDVI 数据融合为例，时空融合算法通常需要已知时刻（ t_k ）的 Landsat 和 MODIS NDVI 数据，以及待预测时刻（ t_p ）的 MODIS NDVI 数据作为输入，去预测 t_p 时刻的 Landsat NDVI 数据。通常， t_k 时刻的 Landsat NDVI 数据需要具有空间完整性。然而，由于频繁的云覆盖，所得到的空间完整的 Landsat NDVI 数据可能与待预测时刻相隔较远，NDVI 变化也会随之增大，预测挑战性也会增大，时空融合的精度难以

得到保证。在该情况下，与 t_p 时刻近邻的 t_m 时刻云覆盖 NDVI 数据可为时空融合提供重要的辅助数据。如图 1 所示，所提出的 STFSR 方法即充分利用 t_k 、 t_m 和 t_p 时刻的所有有效数据，对 t_p 时刻的 Landsat NDVI 数据进行更为可靠的预测。值得注意的是，由于 t_m 时刻的 MODIS NDVI 数据同样受云干扰，为保证有效数据的空间覆盖范围， t_m 时刻的 MODIS NDVI 数据由对应的 Landsat 影像按比例退化生成。为减小真实 MODIS NDVI 数据和模拟 MODIS NDVI 数据的差异，本研究对 t_k 和 t_p 时刻 MODIS NDVI 数据进行了线性变换。其中，线性变换系数由 t_k 时刻 Landsat 和 MODIS NDVI 数据线性回归获得。

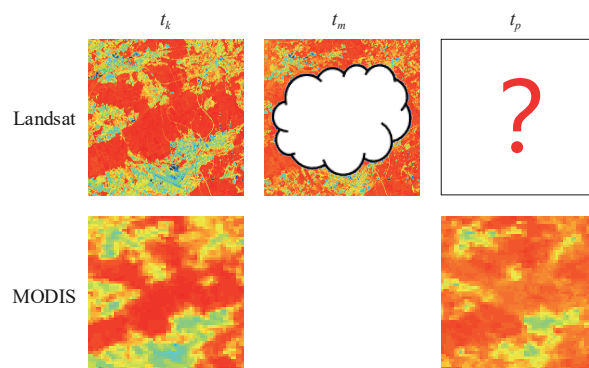


图 1 本文提出 STFSR 方法的输入数据
Fig. 1 Input data of the proposed STFSR method

STFSR 方法的实施主要分为两个步骤，如图 2 所示。

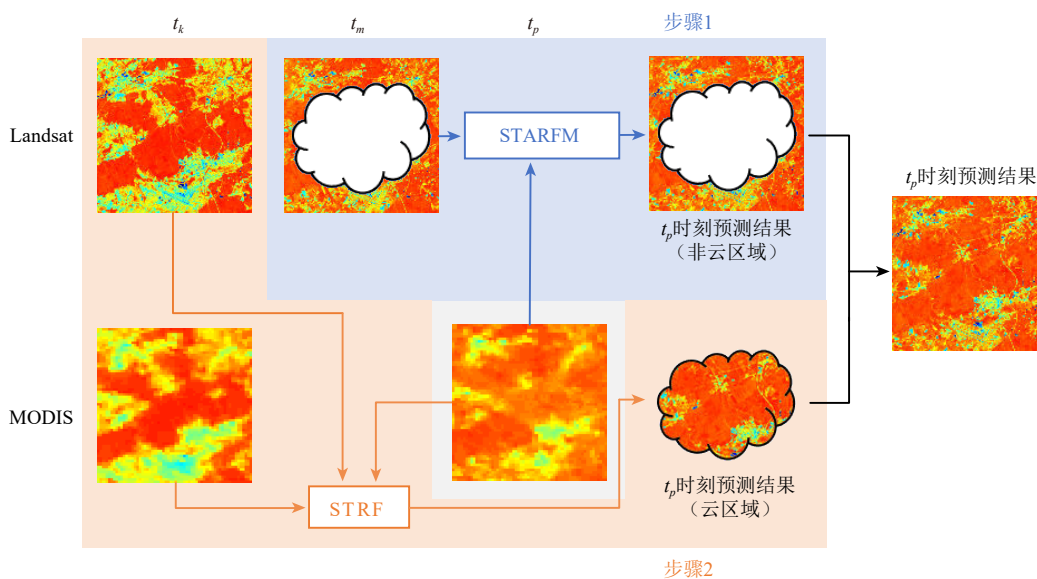


图 2 本文提出 STFSR 方法流程图
Fig. 2 Flowchart of STFSR method proposed in this study

步骤1：通过时空融合，重建 t_m 时刻影像非云区域对应位置下的 t_p 时刻数据。具体地，利用 t_m 时刻的Landsat和MODIS（由Landsat退化所得），以及 t_p 时刻的MODIS对应的非云区域数据输入时空融合，得到 t_p 时刻Landsat对应部分区域的预测结果。在本文中，利用广泛使用的STARFM时空融合方法进行该步骤。由于STARFM主要依赖于邻域信息进行预测，即使存在部分空域缺失，仍可保证待估算的每个像素具有一定的邻域信息，保证STARFM算法的正常实施。该步骤中预测出的 t_p 时刻Landsat NDVI数据可作为步骤2的输入。

步骤2：通过联合时空信息，重建 t_p 时刻剩余的数据。该步骤涉及 t_k 和 t_p 时刻的相关数据。在前期研究中，本研究提出了STRF方法（Wang等，2024），通过充分耦合已知时刻的高、低空间分辨率完整数据对，以及待预测时刻的部分有效数据，实现对剩余缺失区域的重建。由于步骤1已经重建了 t_p 时刻的部分数据，可视为 t_p 时刻的有效数据。具体地，STRF的实施主要分为两个部分：

（1）STRF模型训练。该步骤基于Landsat和MODIS NDVI数据对应的非云区域进行，利用随机森林模型进行训练。模型自变量为 t_k 时刻Landsat数据和MODIS差值数据（ t_p 时刻的MODIS数据减去 t_k 时刻的MODIS数据，并插值到Landsat尺度）下的 3×3 像元图像块（共18个自变量： t_k 时刻Landsat数据的9个像元的NDVI值和MODIS插值数据的9个像元的NDVI值）。因变量为 t_p 时刻单个Landsat NDVI数据对应的单个像元（ 3×3 像元图像块的中心像素）。利用随机森林模型进行训练得到

非线性关系 f 。

（2）STRF模型预测。该步骤基于Landsat和MODIS数据对应的云区域进行，通过输入各缺失像元对应的 t_k 时刻Landsat数据和MODIS差值数据，利用训练后的随机森林模型进行预测，获得 t_p 时刻云区域Landsat NDVI数据的重建结果。

结合步骤1获得的 t_p 时刻非云区域和步骤2获得的 t_p 时刻云区域预测结果，获得 t_p 时刻Landsat NDVI数据的最终预测结果。

3 实验验证

3.1 实验数据及实验设计

本文利用Landsat和MODIS NDVI数据，选取3个不同的区域进行模拟的时空融合实验。3个研究区分别位于法国巴黎、中国河北省和澳大利亚新南威尔士州，如图3所示，3个区域基本是由植被和城市区组成。实验所用到的Landsat和MODIS NDVI数据如图4所示。可见3个区域中，第1时间点和待预测时间点（第3时间点）NDVI数据之间均发生了较大变化；与第3时间点相隔较近的第2个时间点和待预测时间点之间变化较小。Landsat NDVI的空间分辨率为30 m，而250 m MODIS NDVI被重采样到240 m，与Landsat NDVI保持8倍的缩放尺度；数据空间覆盖范围为 $14.4\text{ km}\times 14.4\text{ km}$ 。每个区域包含3个时刻的数据，即 t_k 时刻（与 t_p 时刻相隔较远）的完整Landsat和MODIS数据， t_m 时刻（与 t_p 时刻近邻）的云覆盖Landsat和MODIS NDVI数据，以及 t_p 时刻的Landsat（用于精度评价，不参与算法运行）和MODIS NDVI数据。

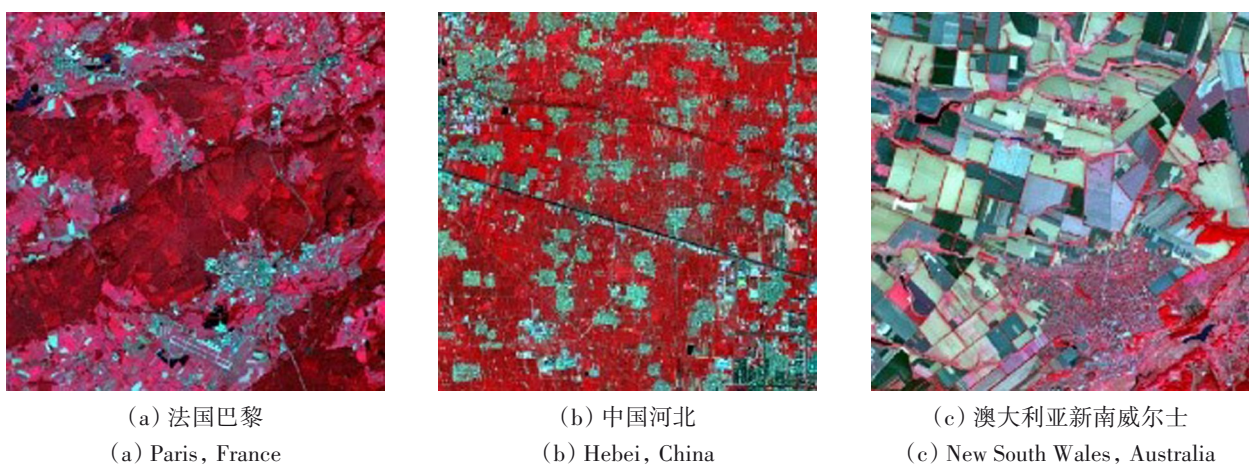


图3 3个研究区的Landsat影像图(以NIR、Red、Green为RGB组合)

Fig. 3 Landsat images of three study areas (NIR, Red and Green bands as RGB)

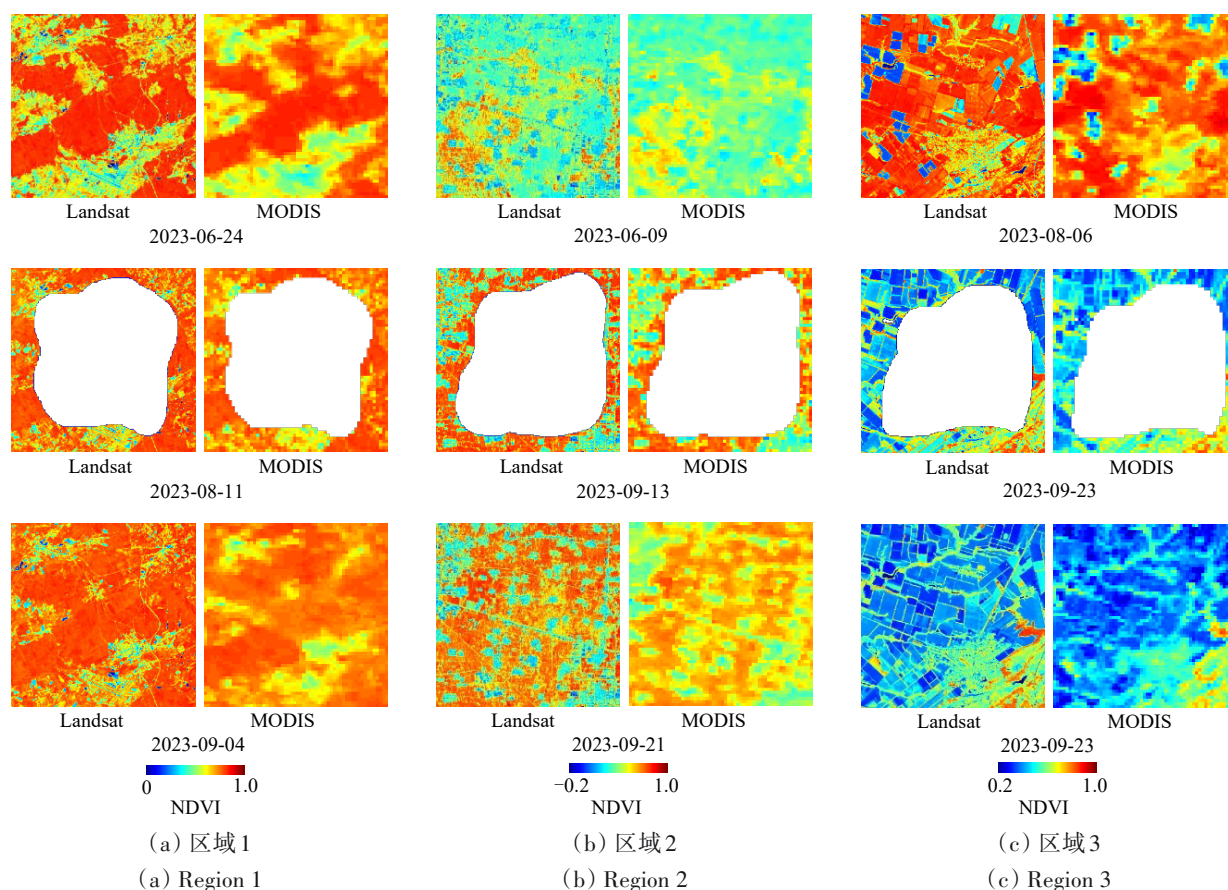


图4 不同日期下3个不同实验区域Landsat和MODIS NDVI数据

Fig. 4 Landsat and MODIS NDVI data for three different study areas on different acquisition dates

由于所提出的STFSR方法旨在解决地表覆盖较大情况下的NDVI时空融合挑战,本研究通过 t_k 与 t_p 时刻Landsat NDVI之间的相关系数CC (Correlation Coefficient) 来量化3个区域NDVI变化程度。经计算,3个区域的CC值分别为0.7057、0.6270和0.2541,表明均经历了较大程度的地表覆盖变化。对于STFSR方法,以上数据都用于算法运行。实验选取两种典型的时空融合方法(STARFM和基于空间加权的虚拟图像对时空融合VIPSTF-SW (Spatial Weighting-Based Virtual Image Pair-Based Spatio-Temporal Fusion) (Wang 等, 2020)) 进行对比。VIPSTF-SW方法是一种在已知和待预测时刻相距较远时,更具优势的时空融合方法。VIPSTF-SW通过构建虚拟数据对减小已知时刻与预测时刻NDVI图像之间的差异,有效降低了高空间分辨率增量估计的不确定性。对于STARFM和VIPSTF-SW方法,仅使用 t_k 和 t_p 时刻NDVI数据进行算法运行。在定量精度评价时,本

文使用了3种评价指标,即均方根误差RMSE (Root Mean Square Error)、CC和相对全局综合误差ERGAS (Relative Global-dimensional Synthesis Error) (Ranchin和Wald, 2000)。所有实验均依托于搭载Intel (R) Xeon (R) Silver 4110 CPU 2.10 GHz的工作站,并使用MATLAB (R2018b)与RStudio (1.1.456版本)编程软件开展。

3.2 验证结果及分析

3.2.1 第1步(非云区域重建)结果

与其他时空融合方法重建待预测时刻高分辨率数据不同,所提出的STFSR方法是分为两个步骤进行的。在步骤一中,需要对待预测时刻的非云区域进行预测,采用的是STARFM时空融合方法。步骤一的结果如图5所示。从视觉上来看,3个区域的非云区域都取得了较好的重建效果,与参考数据在纹理和细节上都十分相近。

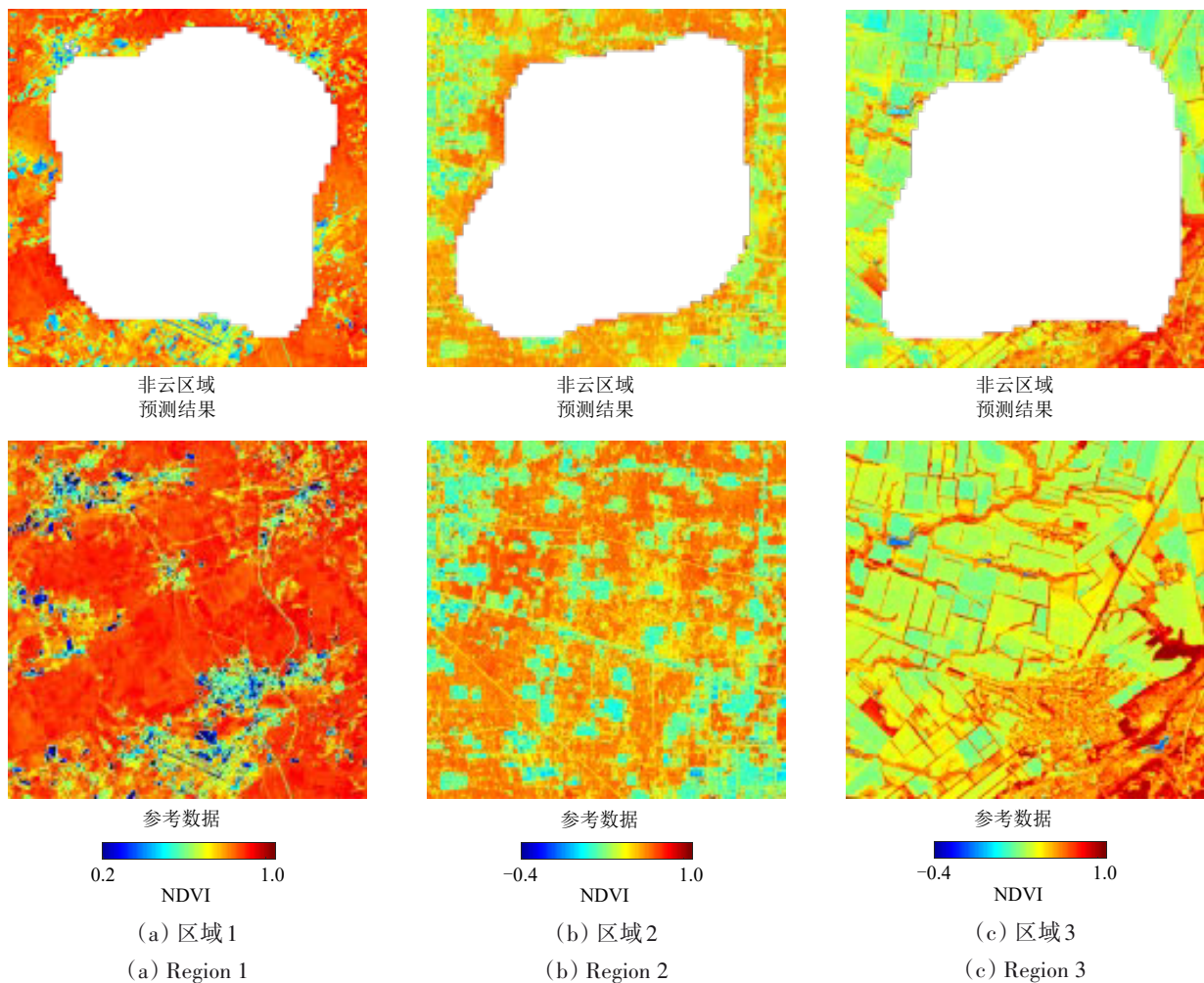


图5 非云区域预测结果

Fig. 5 Prediction for the non-cloud area

3.2.2 第2步(最终重建)结果

STFSR方法的步骤二，即 t_m 时刻数据云区域对应位置下的 t_p 时刻数据的重建，由STRF模型实现，结果如图6所示，对应的误差图如图7所示。从视觉上来看，在3个区域中，STFSR方法与参考数据在纹理和色调上更为相近，误差最小。由于已知时刻和待预测时刻数据在时间上相隔较远，地物变化较大，STARFM和VIPSTF-SW方法的大部分预测结果与参考数据相差较大。以区域1为例，STARFM和VIPSTF-SW预测结果将NDVI高值区域预测为低值，与参考数据之间存在较大色差，而STFSR在色调上与参考数据最为接近。且在误差图中，STARFM和VIPSTF-SW也出现了大面积的深红和深蓝，表明与参考数据之间存在较大误差。从误差分布来看，对于纹理较为简单的植被区，

3种方法的预测结果预测误差均较小。然而，STARFM和VIPSTF-SW在城市区域的预测结果相对参考图像出现了较大的低估，而在部分纹理较为复杂的植被区域出现了较大的高估。在区域2中，STARFM和VIPSTF-SW方法所预测的结果中的纹理与参考数据差别较大，且出现了不同程度的模糊效应。从误差图来看，STARFM和VIPSTF-SW在城市区域都产生了不同程度的低估，而在植被区域产生了不同程度的低估。STFSR方法误差分布与其他两种方法相似，但误差整体较小，且在非云区域的预测误差小于云区域。而在区域3中，由于地物变化极大，给3个方法都带来了一些挑战，都出现了不同程度的模糊效应，但STFSR方法仍然取得了与参考数据更为相近的结果。反观STARFM方法预测结果在色调上与参考数据相

差较大, 而VIPSTF-SW方法虽与参考数据在色调上较为相似, 但丢失了较多纹理信息。从误差图来看, STARFM和VIPSTF-SW在细小地物的NDVI预测中出现了较大的低估, 而STFSR方法误差整体较小且分布较为均匀。为了定量评价各方法的精度, 对3个区域各个方法的预测结果与参考数据计算了RMSE、CC和ERGAS, 如表1所示。在所选取的3个区域中, STFSR方法都取得了最大CC, 最小RMSE和ERGAS, 获得了最高精度。以区域1为例, STFSR方法的RMSE为0.0806, 比STARFM

和VIPSTF-SW方法分别低0.0166和0.0091。对于CC而言, STFSR方法比STARFM和VIPSTF-SW方法分别高出0.0522和0.0419, ERGAS同样也比STARFM和VIPSTF-SW方法分别降低2.1258和1.1697。同时, 本研究对比了区域1—3在3种方法下的运行效率, 如表2所示。可见所提出的STFSR方法虽相较于STARFM和VIPSTF-SW方法略显耗时, 然而鉴于其在精度方面的优势, STFSR方法于NDVI数据时空融合为一种相对更优的选择。

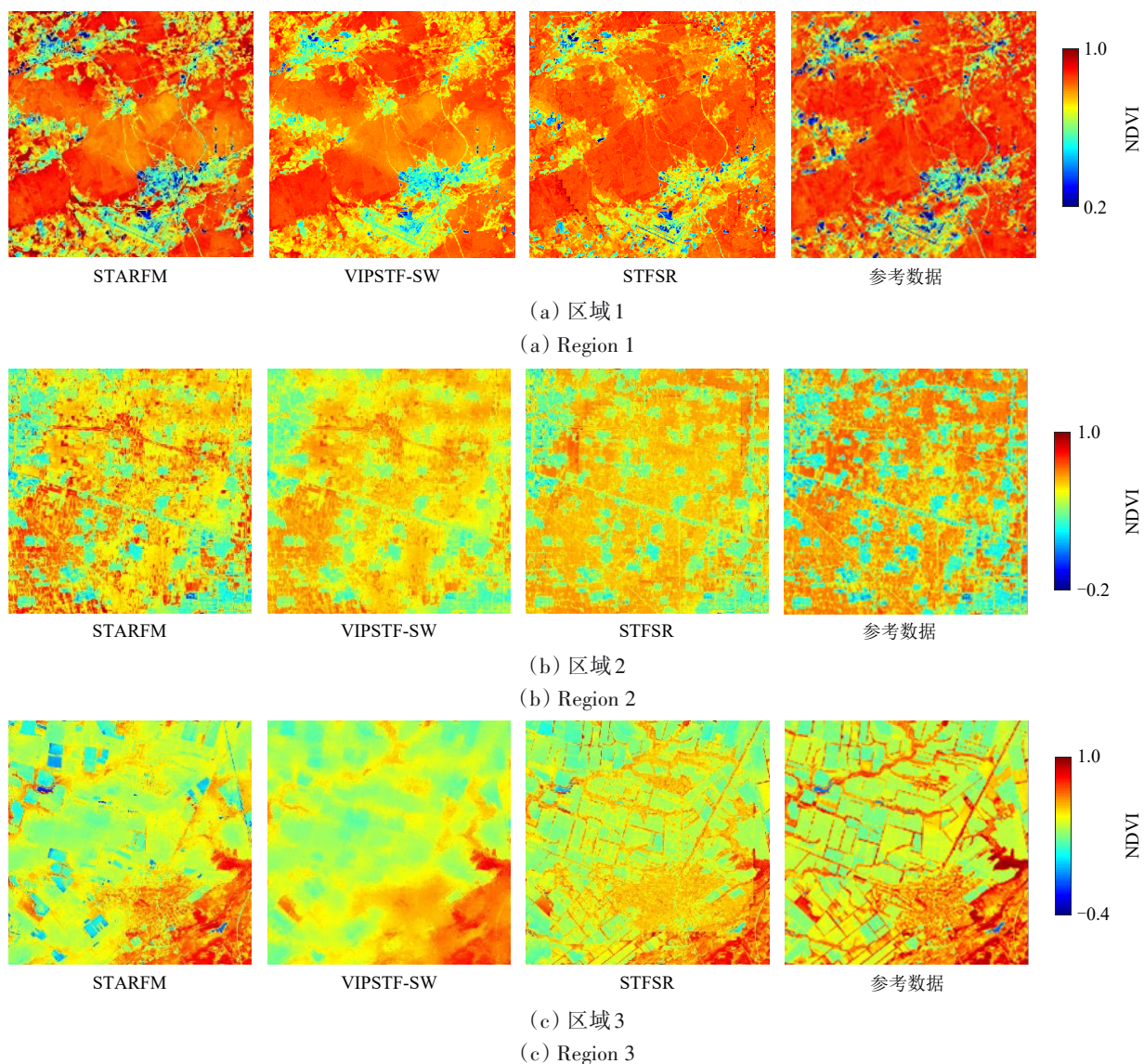


图6 3种方法对3个实验区域整体预测结果的视觉对比

Fig. 6 Visual comparison between predictions of three methods for the whole area of the three regions

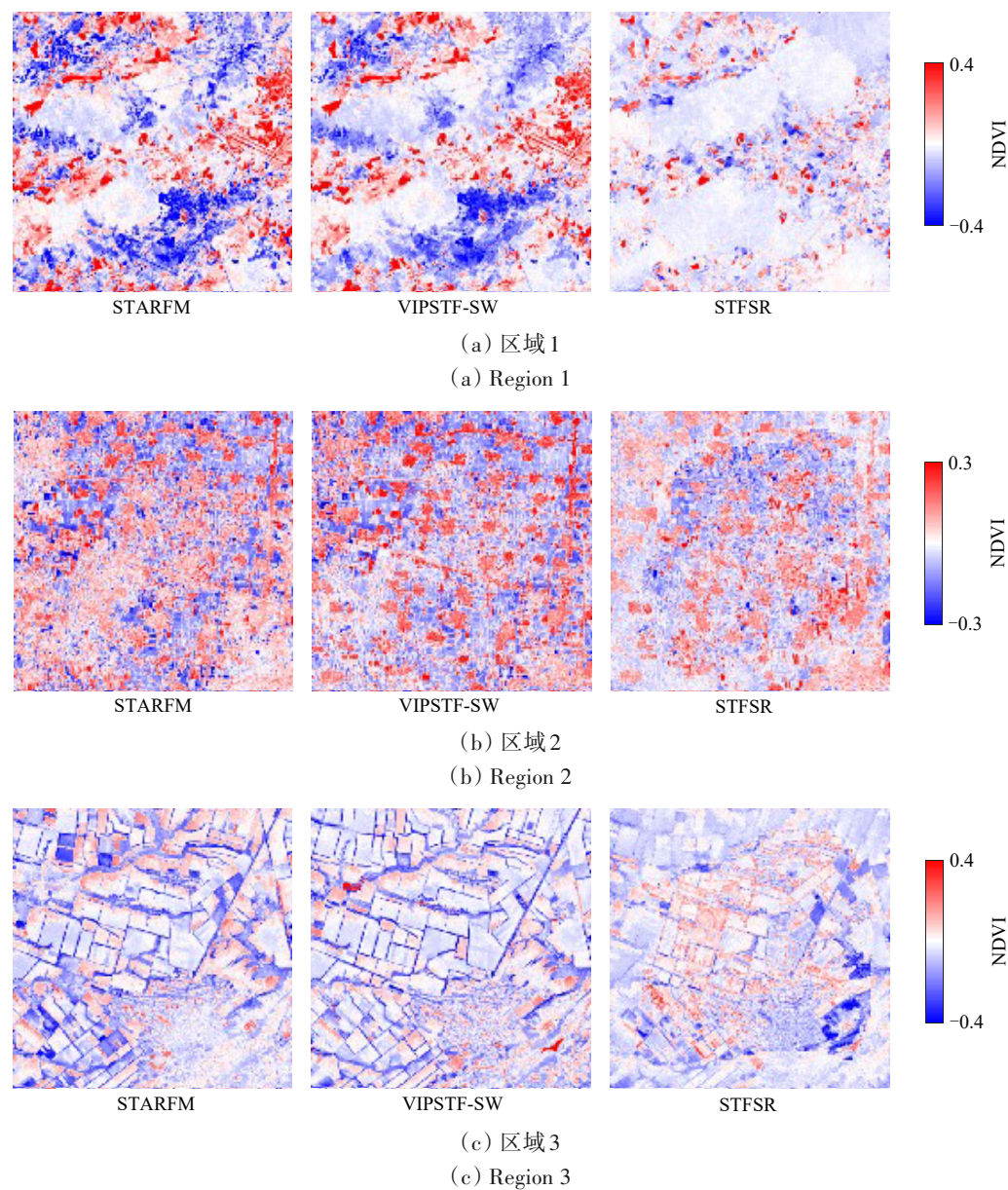


图7 3种计算方法预测结果的误差图对比

Fig. 7 Comparison between error images of predictions of the three methods

表1 3种方法预测结果的精度评价

Table 1 Accuracy assessment of the prediction by three methods

区域	方法	RMSE	CC	ERGAS
		0	1	0
区域1	STARFM	0.0972	0.7833	12.4451
	VIPSTF-SW	0.0898	0.7935	11.4890
	STFSR	0.0806	0.8355	10.3193
区域2	STARFM	0.1282	0.7906	21.6053
	VIPSTF-SW	0.1301	0.8130	21.9274
	STFSR	0.1049	0.8835	16.9333
区域3	STARFM	0.1240	0.6962	30.2091
	VIPSTF-SW	0.1209	0.6869	29.4526
	STFSR	0.0987	0.7972	24.0560

注：加黑数据表示精度最高的方法。

表2 3种方法的运行效率对比
Table 2 Comparison of operational efficiency among three methods

方法	效率/s		
	区域1	区域2	区域3
STARFM	62.8	60.7	60.2
VIPSTF-SW	60.2	61.4	60.8
STFSR	78.4	81.3	78.8

3.2.3 非云区域和云区域重建精度对比

由于STFSR方法对于云区域和非云区域采用不同的方法，分为两步进行重建。对于最终重建结果中的云区域和非云区域，也分别对其精度进行了评价，结果如图8所示。可见在所有3个区域

的结果中，无论是云区域还是非云区域，所提出的STFSR方法都取得了最高的精度。其中，在区域2非云区域STFSR的CC比STARFM和VIPSTF-SW分别高出0.128和0.116；而在云区域，STFSR的CC比STARFM和VIPSTF-SW分别高出0.079和0.042。参与非云区域数据重建的是 t_m 和 t_p 两个时刻的高低分辨率数据，两时刻数据相隔较近，易于获得更高的重建精度。而对于非云区域的重建，所参与的是 t_k 和 t_m 两个时刻的高低分辨率数据。由于云区域重建结果作为非云区域重建的基础，其中包含误差传递的过程，且 t_k 和 t_m 两个时刻数据相隔较远，其重建的难度本就高于非云区域重建。

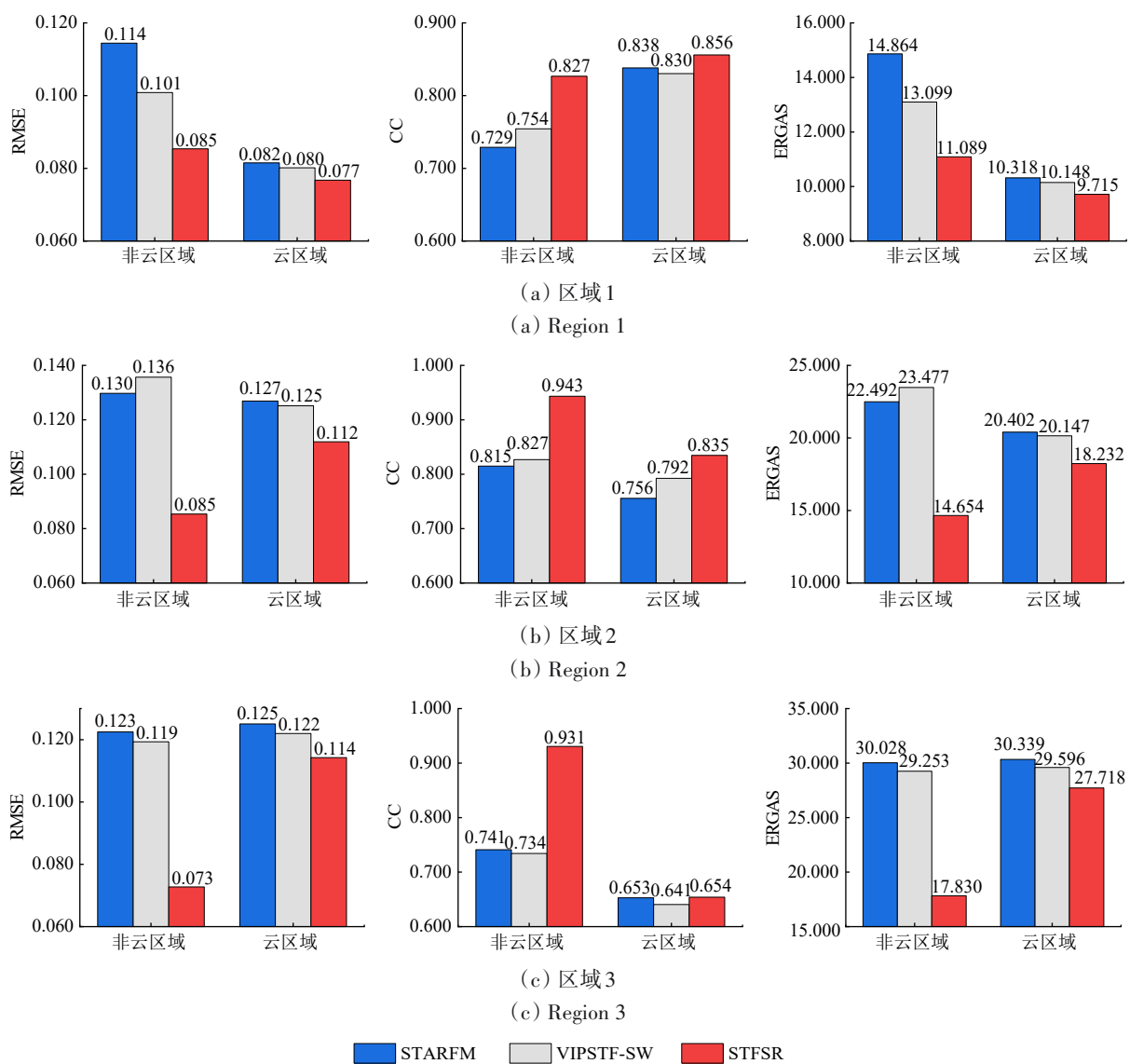


图8 3种方法云和非云区域精度评价

Fig. 8 Accuracy assessment of the cloud and non-cloud areas of the three methods

3.2.4 云尺寸的影响

此外，由于云区域的重建结果会对非云区域的重建产生影响，进而影响总体重建精度。考虑到云尺寸的大小可能对结果的精度产生影响，本研究也研究了使用不同大小的云掩膜时STFSR方法重建精度的变化。本文以30 m像元为单位，依次选取了从图像中心像元往外延伸的正方形云掩膜，以此模拟不同云量。本次实验共考虑5种云尺寸情况，分别是192×192（云量约15%）、256×256（云量约25%）、320×320（云量约40%）、384×384（云量约60%）以及448×448（云量约80%）。图9

展示了不同云尺寸下的STFSR方法精度。同时，也将STARFM和VIPSTF-SW方法的精度作为比较。可见：随着云尺寸增大，STFSR方法的精度在3个区域出现下降；总体来说，STFSR方法的精度在几乎所有情况下都高于STARFM和VIPSTF-SW方法。该实验表明，即使在 t_m 时刻云数据较大时，所提出的STFSR方法也可保持较高的融合精度。在实际应用中，为了保证STFSR预测的精度，本研究中应尽量避免所选取缺失面积过大（如>80%）的辅助影像。

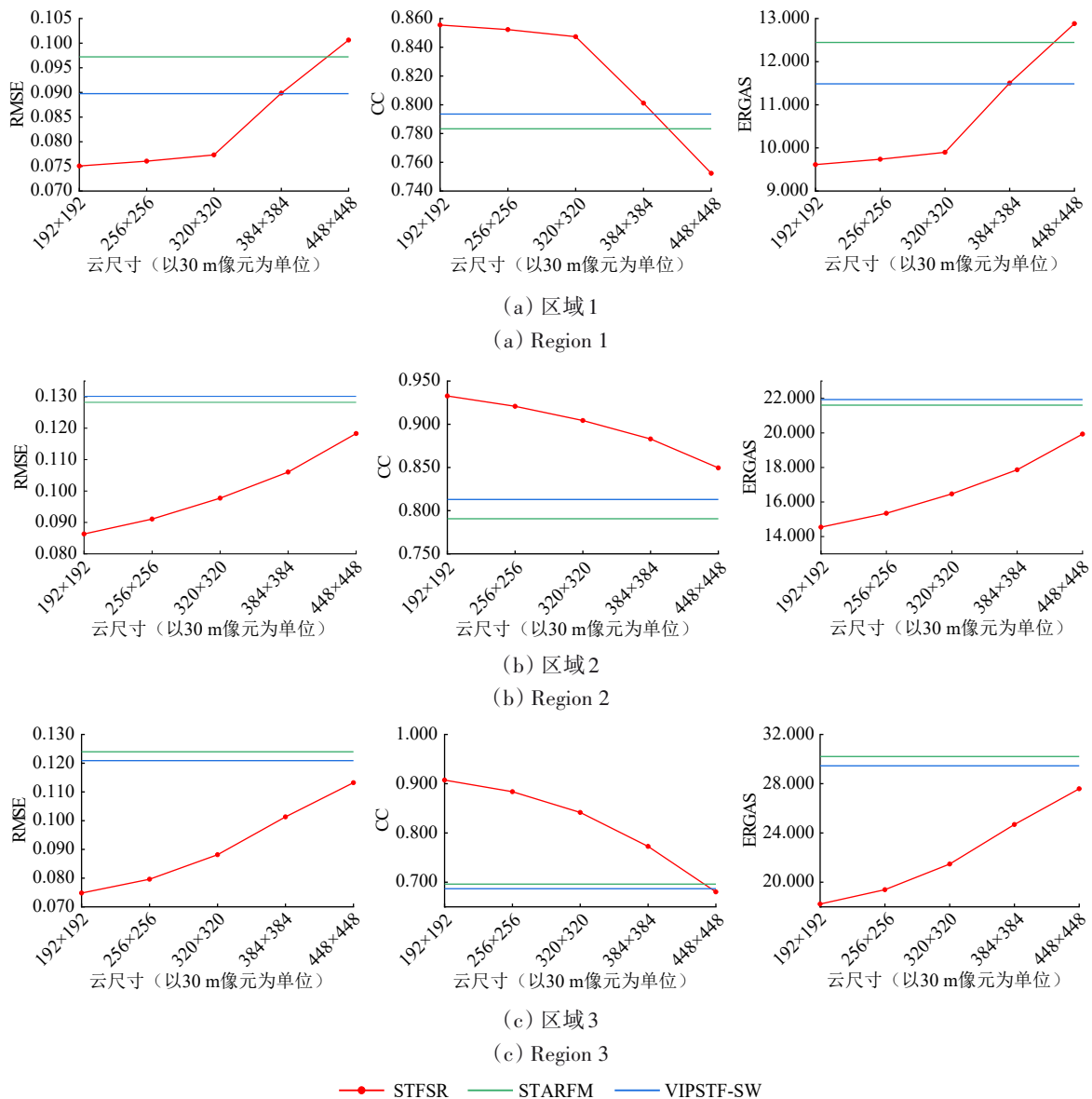


图9 3个区域不同云尺寸下3种方法的精度评价

Fig. 9 Accuracy assessment of three methods under different cloud sizes for regions 1-3

4 讨论

4.1 STFSR 方法的优势

本文所提出的 STFSR 方法结合距预测时刻较近的云数据和距预测时刻较远的完整数据，分别通过时空融合和空域重建方法对缺失的高空间分辨率 NDVI 数据进行分布预测。在辅助数据的选择方面，现有时空融合方法通常追求数据完整性，选取距离预测时刻 t_p 较远的 t_k 时刻的完整 NDVI 数据作为辅助数据。由于两时刻相隔较远，通常地表覆盖变化较大，给预测带来极大挑战。而本文所提出的 STFSR 方法不依赖于辅助数据的完整性，协同距预测时刻较近、但存在云覆盖的 t_m 时刻部分有效数据进行预测。这一做法避免了重要有效数据的浪费，降低了地表覆盖变化的影响，提高了非云区域预测的准确性。表 3 统计了区域 1—3 非云区域数据相关性，以 CC 作为评价指标。可以发现，区域 1—3 中 t_m 和 t_p 时刻非云区域数据相关性远大于 t_k 和 t_p 时刻的相关性。因此，对于非云区域的预测，STFSR 方法较其他时空融合方法预测精度更高。另一方面，在云区域的重建上，STFSR 方法充分利用第一步得到的精度较高的非云区域重建结果协同预测，一定程度上增强了预测的可靠性，在云区域的预测中相较其他时空融合方法也获得了更高的精度（如图 8 结果所示）。

表 3 3 个区域非云区域数据相关性
Table 3 Correlation coefficient for data in non-cloud parts of the three regions

非云区域 数据相关性	区域 1	区域 2	区域 3
$CC_{\text{非云}}(t_k, t_p)$	0.7169	0.6538	0.3456
$CC_{\text{非云}}(t_m, t_p)$	0.7869	0.9649	0.9454

4.2 STFSR 方法的不确定性

STFSR 方法聚焦于充分利用距预测时刻较近，但受云覆盖影响的高分辨率 NDVI 数据中的有效数据。具体而言，STFSR 首先利用时空融合方法 STARFM 重建部分区域，后采用 STRF 方法重建剩余区域。在实验中，也对第 1 步的重建结果进行了精度评价。尽管 3 个区域第 1 步的重建结果都取得了较高的精度，然而第 1 步客观存在的误差不可避免会传递到第 2 步的估计中。在后续研究中，对步

骤一中的不确定性进行量化，并引入步骤二的模型构建中，或为一种可行的手段。此外，由于 STFSR 采用分步式的实施策略，第 1 步中的 STARFM 和第 2 步中的 STRF 算法在模型架构和数据输入方面均存在差异，导致两步所获得的结果存在一定程度的色调差异（图 7）。相较而言，其他两种时空融合方法通过直接采用距预测时刻较远的完整影像进行预测，避免了色调差异的产生。但无论从非云区域和云区域的精度来看，所提出的 STFSR 方法都取得了更高的精度。因此，如何在兼顾算法精度的同时，有效消除两步预测的差异，获得从视觉和精度方面都更为可靠的 STFSR 预测结果，也是后续研究的重点。

此外，STFSR 模型采用 Landsat NDVI 数据升尺度获得的模拟 MODIS NDVI 含云数据作为输入数据，这一做法主要由于两种卫星在同一地区的过境时间存在差异，导致两种数据中的云覆盖区域呈现不一致性。因此，为了保证算法的有效运行，本研究选取了模拟的 MODIS NDVI 含云数据，且模型仍取得了较高的时空融合精度。模拟 MODIS NDVI 相较真实 MODIS NDVI 含云数据用于 STFSR 模型的差异，是后续研究值得探讨的问题。

4.3 研究区属性对精度的影响

本文选取了 3 个研究区对 STFSR 方法进行验证。3 个研究区均为植被覆盖区，其中区域 1 具有较强的空间匀质性，而区域 2 和区域 3 异质性更强。从表 1 中的 3 种评价指标来看，STFSR 方法在区域 1 总体取得了更高的精度。因此，由于异质性区域纹理重建的复杂性，相较于异质性区域，STFSR 方法在匀质性更强的区域精度更高。此外，在异质性更强的区域 2 和区域 3，非云区域的重建相对云区域都取得了更高的精度。这是由于在云区域重建时，已知和待预测时刻相隔更远，对于复杂地物纹理的重建挑战性更大。然而，在区域 1 的重建中，由于整体区域纹理较为简单，且从视觉判断可以发现，云区域的匀质性强于非云区域（图 3）。因此，对于云区域的预测反而取得了更高的精度。

4.4 展望

在 STFSR 模型中，步骤一和二分别使用经典的时空融合方法 STARFM 和耦合多尺度数据的空

域重建方法 STRF。事实上，STFSR 模型为高时空分辨率 NDVI 数据的重建提供了一种高度灵活的框架。在未来的研究中，STFSR 方法可考虑引入其他时空融合和空域重建方法，进一步提升算法的性能。其次，在数据层面，后续研究可考虑将更多异源数据引入 STFSR 模型，以进一步扩充有效数据量，提升算法精度。此外，本文研究了 STFSR 方法在 NDVI 数据时空融合中的应用，对于其他高时空分辨率植被指数（例如增强型植被指数（EVI））数据生成的应用场景仍值得探索。

5 结 论

为应对现有时空融合算法在生成兼具高时空分辨率的 NDVI 数据时，已知时刻和待预测时刻数据之间的 NDVI 变化大的挑战性问题，本文提出了一种耦合空域重建的时空融合方法（STFSR），旨在充分利用与待预测时刻 t_p 相隔较近（ t_m 时刻）但含有部分云覆盖的高分辨率数据，而非无云覆盖但相隔较远的 t_k 时刻数据，协同获得更为可靠的 NDVI 时空融合结果。STFSR 方法通过两个步骤实现：通过时空融合，重建 t_m 时刻 NDVI 数据非云区域对应位置下的 t_p 时刻数据；利用 STRF，重建 t_m 时刻 NDVI 数据云区域对应位置下的 t_p 时刻数据。主要结论如下：

（1）在 3 个实验区中，所提出的 STFSR 总体精度高于常用的 STARFM 和 VIPSTF-SW 时空融合方法，且无论是在云区域还是非云区域，STFSR 都有着明显优势。

（2）当云区域面积逐渐增大时，STFSR 方法精度出现下降，但其精度在几乎所有情况下仍高于 STARFM 和 VIPSTF-SW 方法。

总结来说，STFSR 方法为 NDVI 数据时空融合中存在数据缺失的辅助数据的利用提供了一种有效的解决方案，减小了 NDVI 变化对时空融合精度的影响，有效地提升了时空融合的可靠性。在未来研究中，可通过引入更先进的时空融合与空域重建方法、整合多源异构数据，以进一步提升算法性能。

参考文献 (References)

Amorós-López J, Gómez-Chova L, Alonso L, Guanter L, Zurita-Milla R, Moreno J and Camps-Valls G. 2013. Multitemporal fusion of

- Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*, 23: 132-141 [DOI: 10.1016/j.jag.2012.12.004]
- Busetto L, Meroni M and Colombo R. 2008. Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series. *Remote Sensing of Environment*, 112(1): 118-131 [DOI: 10.1016/j.rse.2007.04.004]
- Cao R Y, Chen Y, Chen J, Zhu X L and Shen M G. 2020. Thick cloud removal in Landsat images based on autoregression of Landsat time-series data. *Remote Sensing of Environment*, 249: 112001 [DOI: 10.1016/j.rse.2020.112001]
- Chen Y, Cao R Y, Chen J, Liu L C and Matsushita B. 2021. A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky – Golay filter. *ISPRS Journal of Photogrammetry and Remote Sensing*, 180: 174-190 [DOI: 10.1016/j.isprsjprs.2021.08.015]
- Chu D, Shen H F, Guan X B, Chen J M, Li X H, Li J and Zhang L P. 2021. Long time-series NDVI reconstruction in cloud-prone regions via spatio-temporal tensor completion. *Remote Sensing of Environment*, 264: 112632 [DOI: 10.1016/j.rse.2021.112632]
- Das M and Ghosh S K. 2016. Deep-STEP: a deep Learning approach for spatiotemporal prediction of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 13(12): 1984-1988 [DOI: 10.1109/lgrs.2016.2619984]
- Dong W Q and Meng J H. 2018. Review of spatiotemporal fusion model of remote sensing data. *Remote Sensing for Natural Resources*, 30(2): 1-11 (董文全, 蒙继华. 2018. 遥感数据时空融合研究进展及展望. *国土资源遥感*, 30(2): 1-11) [DOI: 10.6046/gtzyyg.2018.02.01]
- Gao F, Masek J, Schwaller M and Hall F. 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8): 2207-2218 [DOI: 10.1109/tgrs.2006.872081]
- Gevaert C and García-Haro F J. 2015. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment*, 156: 34-44 [DOI: 10.1016/j.rse.2014.09.012]
- Guo D Z, Shi W Z, Hao M and Zhu X L. 2020. FSDAF 2.0: improving the performance of retrieving land cover changes and preserving spatial details. *Remote Sensing of Environment*, 248: 111973 [DOI: 10.1016/j.rse.2020.111973]
- Hilker T, Wulder M A, Coops N C, Linke J, McDermid G, Masek J G, Gao F and White J C. 2009. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*, 113(8): 1613-1627 [DOI: 10.1016/j.rse.2009.03.007]
- Huang B and Song H H. 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 50(10): 3707-3716 [DOI: 10.1109/TGRS.2012.2186638]
- Ke Y H, Im J, Park S and Gong H L. 2016. Downscaling of MODIS one kilometer evapotranspiration using Landsat-8 data and machine learning approaches. *Remote Sensing*, 8(3): 215 [DOI: 10.

- 3390/rs8030215]
- Li X D, Foody G M, Boyd D S, Ge Y, Zhang Y H, Du Y and Ling F. 2020. SFSDAF: an enhanced FSDAF that incorporates sub-pixel class fraction change information for spatio-temporal image fusion. *Remote Sensing of Environment*, 237: 111537 [DOI: 10.1016/j.rse.2019.111537]
- Liu J B, Ma Y, Wu Y T and Chen F. 2016. Review of methods and applications of high spatiotemporal fusion of remote sensing data. *Journal of Remote Sensing (in Chinese)*, 20(5): 1038-1049 (刘建波, 马勇, 武易天, 陈甫. 2016. 遥感高时空融合方法的研究进展及应用现状. *遥感学报*, 20(5): 1038-1049) [DOI: 10.11834/jrs.20166218]
- Liu M, Yang W, Zhu X L, Chen J, Chen X, Yang L Q and Helmer E H. 2019. An improved flexible spatiotemporal data fusion (IFSDF) method for producing high spatiotemporal resolution normalized difference vegetation index time series. *Remote Sensing of Environment*, 227: 74-89 [DOI: 10.1016/j.rse.2019.03.012]
- Maselli F. 2001. Definition of spatially variable spectral endmembers by locally calibrated multivariate regression analyses. *Remote Sensing of Environment*, 75(1): 29-38 [DOI: 10.1016/S0034-4257(00)00153-X]
- Ranchin T and Wald L. 2000. Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, 66(1): 49-61
- Song H H and Huang B. 2013. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4): 1883-1896 [DOI: 10.1109/tgrs.2012.2213095]
- Song H H, Liu Q S, Wang G J, Hang R L and Huang B. 2018. Spatiotemporal satellite image fusion using deep convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3): 821-829 [DOI: 10.1109/JSTARS.2018.2797894]
- Verger A, Baret F, Weiss M, Kandasamy S and Vermote E. 2013. The CACAO method for smoothing, gap filling, and characterizing seasonal anomalies in satellite time series. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4): 1963-1972 [DOI: 10.1109/tgrs.2012.2228653]
- Wang Q M and Atkinson P M. 2018. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sensing of Environment*, 204: 31-42 [DOI: 10.1016/j.rse.2017.10.046]
- Wang Q M, Tang Y J, Tong X H and Atkinson P M. 2020. Virtual image pair-based spatio-temporal fusion. *Remote Sensing of Environment*, 249: 112009 [DOI: 10.1016/j.rse.2020.112009]
- Wang Q M, Tang Y J, Tong X H and Atkinson P M. 2024. Filling gaps in cloudy Landsat LST product by spatial-temporal fusion of multi-scale data. *Remote Sensing of Environment*, 306: 114142 [DOI: 10.1016/j.rse.2024.114142]
- Wang Q M, Zhang Z H and Zhang C Y. 2023. Downscaling GF-5 hyperspectral images by fusing with Sentinel-2 images. *National Remote Sensing Bulletin*, 27(8): 1936-1950 (王群明, 张智昊, 张成媛. 2023. 融合 Sentinel-2 数据的高分五号高光谱数据降尺度. *遥感学报*, 27(8): 1936-1950) [DOI: 10.11834/jrs.20211420]
- Wu M Q, Niu Z, Wang C Y, Wu C Y and Wang L. 2012. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model. *Journal of Applied Remote Sensing*, 6(1): 063507 [DOI: 10.1117/1.Jrs.6.063507]
- Xu H Q, Ren M J and Yang L J. 2023. Evaluating the consistency of surface brightness, greenness, and wetness observations between Landsat-8 OLI and Landsat-9 OLI2 through underfly images. *International Journal of Applied Earth Observation and Geoinformation*, 124: 103546 [DOI: 10.1016/j.jag.2023.103546]
- Xu Y, Huang B, Xu Y Y, Cao K, Guo C L and Meng D Y. 2015. Spatial and temporal image fusion via regularized spatial unmixing. *IEEE Geoscience and Remote Sensing Letters*, 12(6): 1362-1366 [DOI: 10.1109/lgrs.2015.2402644]
- Yin G F, Li A N, Jin H A, Zhao W, Bian J H, Qu Y H, Zeng Y L and Xu B D. 2017. Derivation of temporally continuous LAI reference maps through combining the LAI Net observation system with CACAO. *Agricultural and Forest Meteorology*, 233: 209-221 [DOI: 10.1016/j.agrformet.2016.11.267]
- Zhang H K, Chen J M, Huang B, Song H H and Li Y R. 2014. Reconstructing seasonal variation of Landsat vegetation index related to leaf area index by fusing with MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(3): 950-960 [DOI: 10.1109/jstars.2013.2284528]
- Zhang Z H, Wang Q M and Ding X Y. 2024. FVC-Net: a fusion network for producing fine spatial resolution fractional vegetation cover. *National Remote Sensing Bulletin*, 28(12): 3184-3196 (张智昊, 王群明, 丁欣宇. 2024. FVC-Net: 高空间分辨率植被覆盖度的融合网络. *遥感学报*, 28(12): 3184-3196) [DOI: 10.11834/jrs.20243112]
- Zhu X L, Cai F Y, Tian J Q and Williams T K A. 2018. Spatiotemporal fusion of multisource remote sensing data: literature survey, taxonomy, principles, applications, and future directions. *Remote Sensing*, 10(4): 527 [DOI: 10.3390/rs10040527]
- Zhu X L, Chen J, Gao F, Chen X H and Masek J G. 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114(11): 2610-2623 [DOI: 10.1016/j.rse.2010.05.032]
- Zhu X L, Helmer E H, Gao F, Liu D S, Chen J and Lefsky M A. 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment*, 172: 165-177 [DOI: 10.1016/j.rse.2015.11.016]
- Zhukov B, Oertel D, Lanzl F and Reinhackel G. 1999. Unmixing-based multisensor multiresolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3): 1212-1226 [DOI: 10.1109/36.763276]
- Zurita-Milla R, Gomez-Chova L, Guanter L, Clevers J G P W and Camps-Valls G. 2011. Multitemporal unmixing of medium-spatial-resolution satellite images: a case study using MERIS images for land-cover mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11): 4308-4317 [DOI: 10.1109/tgrs.2011.2158320]

Spatial reconstruction-aided spatio-temporal fusion of NDVI images

TANG Yijie, WANG Qunming

College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China

Abstract: Normalized Difference Vegetation Index (NDVI) images with fine spatial and temporal resolutions are important data for real-time precise vegetation monitoring. Remote sensing image, as an important data source for producing NDVI data, however, always present a trade-off between the spatial and temporal resolutions due to the limitation in the power of satellites. Generally, sensors with fine spatial resolution always have a long revisit time (e.g., Landsat images), while sensors with a short revisit period always have coarse spatial resolution (e.g., Moderate-resolution Imaging Spectroradiometer (MODIS) images). Spatio-temporal fusion technique can be applied to generate NDVI images with both fine spatial and temporal resolutions, by fusing NDVI images acquired from these two categories of sensors. The existing spatio-temporal fusion methods, however, suffer from a long-standing challenge, that is, the NDVI change between the images at the known and prediction times, which restricts the accuracy of spatio-temporal fusion prediction greatly. In this paper, a Spatio-Temporal Fusion Then Spatial Reconstruction (STFSR) method was proposed to cope with the NDVI change issue in predicting the 30 m Landsat NDVI images.

Generally, when predicting the missing Landsat NDVI image by spatio-temporal fusion, a pair of spatially complete fine and coarse spatial resolution NDVI images is also required (probably temporally far from the prediction time). Except for the original auxiliary images, the proposed STFSR method also included the fine spatial resolution image temporally closer to the prediction time, but with different degrees of data loss caused by cloud cover (hereafter, simplified as auxiliary cloudy NDVI image) in prediction. The implementation of STFSR is divided into two steps: (1) Reconstructing the non-cloud area (the corresponding non-cloud area in the auxiliary cloudy image, but in the image to be predicted) using the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM). (2) Reconstructing the cloud area (the corresponding cloud area in the auxiliary cloudy image, but in the image to be predicted) by a Spatial-Temporal Random Forest (STRF) algorithm, a spatial reconstruction method integrating the information from both fine and coarse spatial resolution NDVI images.

In the experiments in three regions, the effectiveness of the proposed STFSR method was evaluated, by comparing with two commonly used spatio-temporal fusion methods, the STARFM and the Spatial Weighting-based Virtual Image Pair-based Spatio-Temporal Fusion (VIPSTF-SW) algorithms. The results demonstrate that the proposed STFSR can produce greater accuracy than the other two methods for all three regions. Furthermore, when the cloud coverage increases to a certain percentage (e.g., 80%) in the auxiliary cloudy image, the STFSR method can still provide a more satisfactory prediction than two benchmark methods. Specifically, the average Root Mean Square Error (RMSE) of STFSR is 0.0217 and 0.0188 smaller than that of STARFM and VIPSTF-SW, respectively. The corresponding average Correlation Coefficient (CC) is 0.0820 and 0.0742 larger, and the corresponding average Relative Global-dimensional Synthesis Error (ERGAS) is 4.3170 and 3.8535 smaller.

The proposed STFSR method takes full advantage of the important information in the cloudy, but temporally closer NDVI image, which fails to be utilized in existing spatio-temporal fusion methods. Generally, the proposed STFSR method provides a flexible solution to deal with the NDVI change in spatio-temporal fusion. Moreover, this model has great potential for the generation of other vegetation index data with fine spatial and temporal resolutions, such as the Enhanced Vegetation Index (EVI) and the Leaf Area Index (LAI).

Key words: remote sensing, Landsat, MODIS, NDVI, spatio-temporal fusion, spatial reconstruction

Supported by National Natural Science Foundation of China (No. 42222108, 42171345)