

基于机器学习联合 TanDEM-X InSAR 和 ICESat-2 数据估计大范围林下地形

胡华参¹, 朱建军¹, 付海强¹, Lopez-Sanchez Juan Manuel²,
Gómez Cristina^{3,4}, 张涛¹, 刘奎¹

1. 中南大学 地球科学与信息物理学院, 长沙 410083;

2. 阿利坎特大学 计算机研究院, 阿利坎特 03690 西班牙;

3. 巴利亚多利德大学 可持续森林管理研究所, 杜克斯德索里亚 47002 西班牙;

4. 阿伯丁大学 地球科学学院, 苏格兰阿伯丁 AB24 3UE 英国

摘要: 双站 TanDEM-X InSAR 系统已成功应用于生产全球数字高程模型。然而, 受 X 波段 SAR 信号的穿透能力限制和森林体积散射的影响, 在森林地区提取的 DEM 包含严重森林信号。因此, 为降低 TanDEM-X InSAR 数据估计林下地形过程中森林体散射对 InSAR 测高的影响, 本研究提出了一种基于机器学习联合 TanDEM-X InSAR、ICESat-2 和 Landsat 8 数据估计林下地形的方案。为验证所提方法的有效性, 选用了两个具有不同地形条件和森林类型特征的试验区 (加蓬热带雨林试验区和西班牙北方试验区) 进行了测试, 并利用高精度机载 LiDAR DTM 进行精度评定。结果表明: 在加蓬热带雨林试验区, 所提方法估计林下地形在 2 个验证区域的 RMSE 为 5.45 m 和 5.91 m, 与 InSAR DEM 的估测结果 14.70 m 和 18.58 m 相比, 地形精度提高了 60% 以上; 在西班牙北方森林试验区, 林下地形估测的 RMSE 也从 6.05—9.10 m 降低到了 3.06—4.42 m。综上, 本研究为使用双站 X 波段 InSAR 系统准确估计大范围林下地形提供了一种有效且稳健的方案。

关键词: 林下地形, 合成孔径雷达干涉测量, 相位中心高度, 机器学习, TanDEM-X, ICESat-2

中图分类号: P2

引用格式: 胡华参, 朱建军, 付海强, Lopez-Sanchez Juan Manuel, Gómez Cristina, 张涛, 刘奎. 2025. 基于机器学习联合 TanDEM-X InSAR 和 ICESat-2 数据估计大范围林下地形. 遥感学报, 29(1): 191-202

Hu H C, Zhu J J, Fu H Q, Lopez-Sanchez J M, Gómez C, Zhang T and Liu K. 2025. Large-scale sub-canopy topography estimation from TanDEM-X InSAR and ICESat-2 data using machine learning method. National Remote Sensing Bulletin, 29(1):191-202[DOI:10.11834/jrs.20233152]

1 引言

数字高程模型 DEM (Digital Elevation Model) 是基础设施建设、自然资源开发和灾害预测不可或缺的输入数据 (Dias De Paiva 等, 2013; Fu 等, 2017)。激光雷达 LiDAR (Light Detection And Ranging)、光学摄影测量和合成孔径雷达干涉 InSAR (Interferometric Synthetic Aperture Radar) 技术已广泛用于 DEM 获取 (许家录, 2022; Rizzoli 等, 2017; Abrams 等, 2020)。其中 InSAR 技术凭借其全天时、全天候和高分辨率数据获取的优势, 已

成功应用于生产全球 DEM。

TerraSAR-X/TanDEM-X, 作为全球第一个星载双站 InSAR 系统, 成功获取了分辨率为 12 m 的全球 DEM (Gruber 等, 2012; Rizzoli 等, 2017)。然而, 在森林尤其是在高大茂密森林覆盖的区域, 受森林体散射的影响, 常规 InSAR 技术提取的 DEM 包含严重的森林高度信号, 即相位中心高度 PCH (Phase Center Height)。为了解决该问题, 已有研究提出了多种散射模型用于区分同一个分辨单元内不同散射体的贡献, 例如随机地体散射 RVoG (Random Volume over Ground) 模型已被广

收稿日期: 2023-05-09; 预印本: 2023-10-24

基金项目: 国家自然科学基金 (编号: 41820104005, 42030112); 湖南省自然科学基金 (编号: 2021JJ30808)

第一作者简介: 胡华参, 研究方向为基于 InSAR 技术的森林高度和林下地形估计。E-mail: csuhuacan@csu.edu.cn

通信作者简介: 朱建军, 研究方向为误差数据处理及其在 InSAR 中的应用。E-mail: zjj@csu.edu.cn

泛用于描述 SAR 信号在森林中的散射过程 (Cloude 和 Papathanassiou, 1998; 朱建军等, 2018)。基于 RVoG 模型, 文献 (Treuhart 等, 1996; Garestier 等, 2008; Fu 等, 2016, 2017) 提出了一些林下地形估计的算法并验证了其在不同森林条件下的可行性。然而, 以上算法需要极化 InSAR、多基线 InSAR 或时频 InSAR 数据进行解算, 且目前只能使用机载 InSAR 数据进行小范围验证, 不能应用于 TanDEM-X 标准模式下获取的单基线、单极化 InSAR 数据。

针对 TanDEM-X InSAR 系统的特点, Lei 等 (2021) 提出了一种基于少视直方图统计的方法估计林下地形, 但是统计窗口内的局部坡度的变化会造成经验函数的错误估计, 使得该方法难以应用于地形起伏明显的林区。此外, Wang 等 (2021) 通过简化 RVoG 模型将森林高度信号从 InSAR DEM 中移除。然而, 该方法是基于无限深假设模型提出的 (Dall, 2007; Schlund 等, 2019), 在低矮森林区域难以满足模型的假设条件。为了更好的估计林下地形, 已有研究中引入了冰、云和陆地高程卫星 ICESat (Ice, Cloud and Land Elevation Satellite) 获取的森林高度或地面高程信息作为控制, 结合光学遥感获取的森林信息通过回归分析或机器学习的方法估计相位中心高度, 并将其从 InSAR DEM 中移除获取林下地形 (Su 等, 2015; O'Loughlin 等, 2016; Kulp 和 Strauss, 2018; Magruder 等, 2021; Hawker 等, 2022)。但是, 由于 InSAR 信号具有一定的穿透能力, 其测量的森林特性难以被光学遥感数据完全替代, 且在高大茂密森林区光学遥感参数易出现饱和现象 (Rahman 等, 2005), 不能很好的移除 InSAR DEM 中的森林高度信号。

鉴于此, 本研究拟从 TanDEM-X InSAR 数据测量的相干性和后向散射系数信息入手, 深入剖析 SAR 信号在森林中的散射过程; 在此基础上, 结合星载 ICESat-2 提供的离散地表高程信息和 Landsat 8 光学数据提供的连续森林参数, 通过机器学习建立不同参数与相位中心高度之间的耦合关系; 最后, 在此基础上, 构建基于 TanDEM-X、ICESat-2 和 Landsat 8 数据提取大范围林下地形的技术框架; 最后, 利用位于非洲加蓬和西班牙的典型试验区对所提的算法进行了测试和验证, 以期待为双站 X 波段 InSAR 系统实现大范围林下地形估计提供理论参考。

2 研究方法

2.1 林下地形估计原理

在双站模式下, TerraSAR-X 和 TanDEM-X 卫星几乎同时获取了两幅 SAR 影像 S_1 和 S_2 , 复干涉相干性可表示为

$$\tilde{\gamma} = \frac{\langle S_1 S_2^* \rangle}{\sqrt{\langle S_1 S_1^* \rangle \langle S_2 S_2^* \rangle}} \quad (1)$$

式中, $\langle \cdot \rangle$ 代表期望处理, $*$ 表示共轭转置。在森林区域, 尽管双站模式下获取的 InSAR 数据不受时间去相干的影响, 但仍然存在多种去相干源会显著降低复相干性 (Lopez-Sanchez 等, 2017)。通过已知的策略对复相干性进行校正后 (Martone 等, 2015), 复相干性 $\tilde{\gamma}$ 可以认为直接代表了纯体去相干, 其表示了 InSAR 信号记录的森林形状、结构和介电属性, 是估计相位中心高度的关键 (张涛等, 2022)。

此外, 在森林区域, 干涉相位可表示为

$$\phi_{\text{int}} = \phi_{\text{flat}} + \phi_{\text{top}} + \phi_{\text{forest}} + \phi_{\text{orb}} + \phi_{\text{noise}} \quad (2)$$

式中, ϕ_{flat} 表示平地相位, ϕ_{top} 是地形相位, ϕ_{orb} 是轨道误差相位, ϕ_{noise} 是噪声相位, ϕ_{forest} 是森林体散射引起的相位偏差, 即相位中心高度。InSAR 技术估计林下地形的关键在于如何将 ϕ_{top} 和 ϕ_{forest} 分离。

如图 1 所示, 采用常规 InSAR 技术, 将相位转换为高程后, 与森林冠层散射相关的相位中心高度 ϕ_{forest} 也可以表示为

$$H_{\text{PCH}} = H_{\text{DEM}} - H_{\text{Ground}} \quad (3)$$

式中, H_{PCH} 表示相位中心高度, H_{DEM} 表示 InSAR 获取的高程, H_{Ground} 表示真实的地表高程, 即林下地形。尽管 ICESat-2 等星载 LiDAR 系统的空间分辨率较低, 但它可以估算 InSAR DEM 中的一些相位中心高度:

$$H_{\text{PCH}} = H_{\text{DEM}} - H_{\text{ICESat-2}} \quad (4)$$

式中, $H_{\text{ICESat-2}}$ 表示 ICESat-2 地面点对应的高程, 这为提取连续的相位中心高度提供了可能。然而, 由于式 (2) 不能用于直接分离相位中心高度, 其它 InSAR 观测信息, 例如相干性幅度 $|\tilde{\gamma}|$ 以及与森林物理属性相关的后向散射系数应予以考虑 (Wang 等, 2021)。

基于 RVoG 模型, 可以发现相位中心高度是消光系数、地体幅度比和有效垂直波数的函数 (Cloude 和 Papathanassiou, 1998; Papathanassiou 和

Cloude, 2001), 这意味着相干性 $|\tilde{\gamma}|$ 或后向散射系数可以直接与相位中心高度相关联。然而, 尽管相干性和后向散射系数记录了对相位中心高度敏感的信息, 但只依赖其建模很难估计准确的相位中心高度, 原因包括以下两点: (1) 单极化、单基线条件下获取的 TanDEM-X InSAR 数据面临观测信息不足问题, 仅依赖相干性或后向散射系数难以支撑 RVoG 模型的解算; (2) 相干性作为 InSAR 记录森林信号的直接表达, 除森林高度信息外, 受森林类型、密度、结构和介电属性等多种因素的影响, 即相同的相干性可能对应不同的相位中心高度。

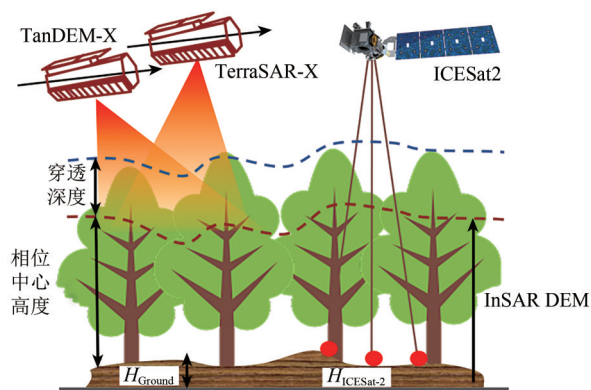


图1 相位中心高度估计示意图

Fig. 1 Schematic diagram of the estimation of the PCH height

为解决以上问题, 本文引入光学遥感数据获取的森林信号辅助 InSAR 参数更好的估计相位中心高度。光学遥感获取的归一化植被指数 NDVI (Normalized Difference Vegetation index)、增强植被指数 EVI (Enhanced Vegetation index)、植被连续场 VCF (Vegetation Continuous Fields) 和比值植被指数 RVI (Ratio Vegetation Index) 等提供了丰富的森林类型、密度等信息 (O’Loughlin 等, 2016; Son 等, 2014; Právělie 等, 2022; Hansen 等, 2008), 可以有效增强相干性和后向散射系数对相位中心高度的区分能力。此外, 雷达距离向的坡度已被证明对相干性和后向散射系数存在影响, 会扭曲他们与相位中心高度之间的关系, 因此必须考虑距离向坡度的影响 (Wang 等, 2021; Su 等, 2015)。

2.2 估计相位中心高度

为根据连续 TanDEM-X InSAR 观测信息、Landsat 8 植被参数估算连续的相位中心高度, 本文采用了机器学习方法构建了上述特征参数与离

散相位中心高度 (代表 InSAR DEM 与 ICESat-2 地离散面高程之间的高程差异) 间的复杂关联。为此, 本文选择了 CatBoost 机器学习模型 (Ben Jabeur 等, 2021), 该模型能够将众多不同强度的回归器整合在一起, 通过添加和训练梯度树得到最终的集成模型, 相比于传统的模型极大的提高了预测能力。为构建训练样本, 首先提取了 ICESat-2 ATL08 产品控制点在 InSAR DEM 中的位置分布。此外, 为了尽可能的消除定位误差带来的影响, 本文以 ICESat-2 高程的地理位置所在的 InSAR DEM 像素为中心构建了窗口大小为 3×3 的覆盖区域, 其中 DEM 分辨率为 10 m。之后, 提取窗口覆盖区 InSAR DEM 的高斯加权平均值作为对应 DEM 新的高程值。最后, 通过计算新高程点与 ICESat-2 高程之间的高度差作为相位中心高度样本。

基于上述分析, 本文采用的 8 个特征参数包括: 后向散射系数 σ_1 和 σ_2 、相干性 $|\tilde{\gamma}|$ 、距离向坡度以及 4 个植被特征参数 (NDVI、EVI、VCF 和 RVI)。以上参数样本被随机分成训练集 (70%) 和验证集 (30%) 用以模型的训练, 最后将训练后的模型用于预测 ICESat-2 未覆盖区域的相位中心高度。需要注意的是, 由于 SAR 特殊的侧视成像方式, 入射角、基线长度等不同会导致穿透深度不同, 即相位中心高度存在巨大差异。此外, 在大范围相位中心高度的估计中有必要考虑森林的空间异质性的影响。鉴于此, 本文针对每个 InSAR 干涉对分别进行机器学习模型的训练和预测相位中心高度。最终, 本文算法的整体处理流程如图 2 所示。

3 验证方法与结果分析

3.1 试验区及数据

3.1.1 试验区

本文选择了两个不同森林类型和地形条件的典型试验区对所提算法进行验证, 对应的地理位置如图 3 所示。第一个试验区位于非洲加蓬, 陆地面积约 6000 km^2 。该区域被茂密的热带森林覆盖, 其主要森林类型为海洋红树林和沼泽森林, 森林高度集中在 10—50 m; 试验区内地形较为平坦, 东北和西南方向海拔相对较高, 高程变化范围为 0—200 m。第二个试验区位于西班牙, 面积超过 50 万 km^2 。该区域内森林总体结构复杂, 成分多变, 主要以针叶

林、阔叶林和混交林为主, 高度集中在5—35 m; 高程范围为0—3500 m。
试验区内地形起伏明显, 主要山脉呈东西走向,

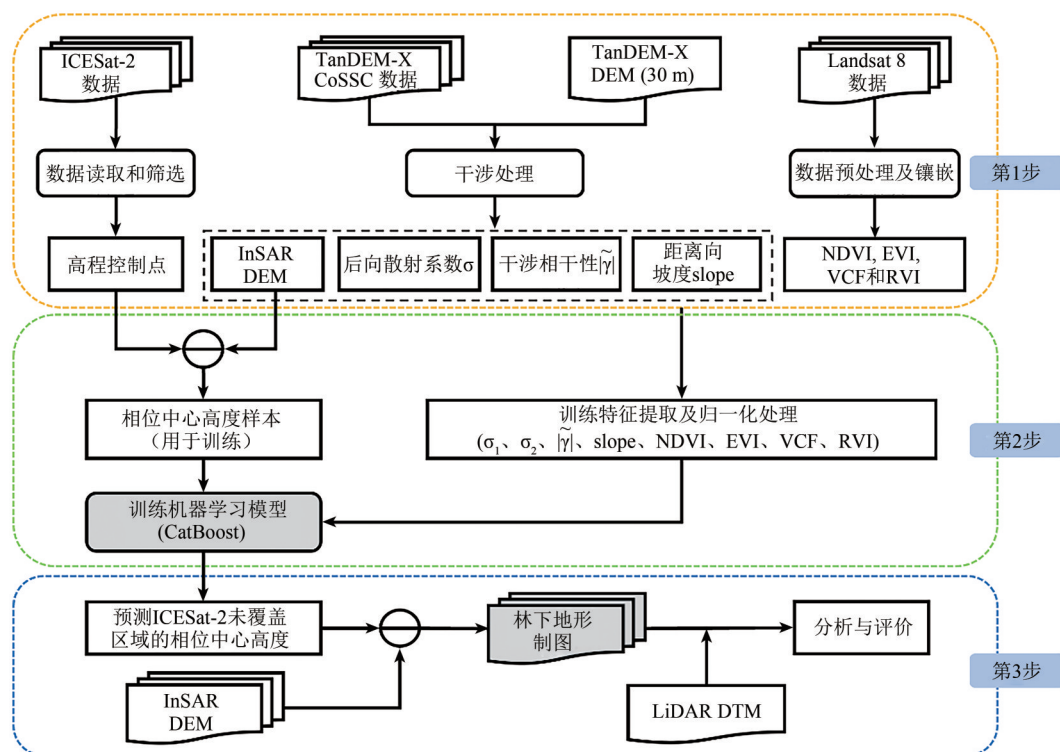


图2 所提方法估计林下地形流程图

Fig. 2 Flowchart of the proposed algorithm for sub-canopy topography estimation

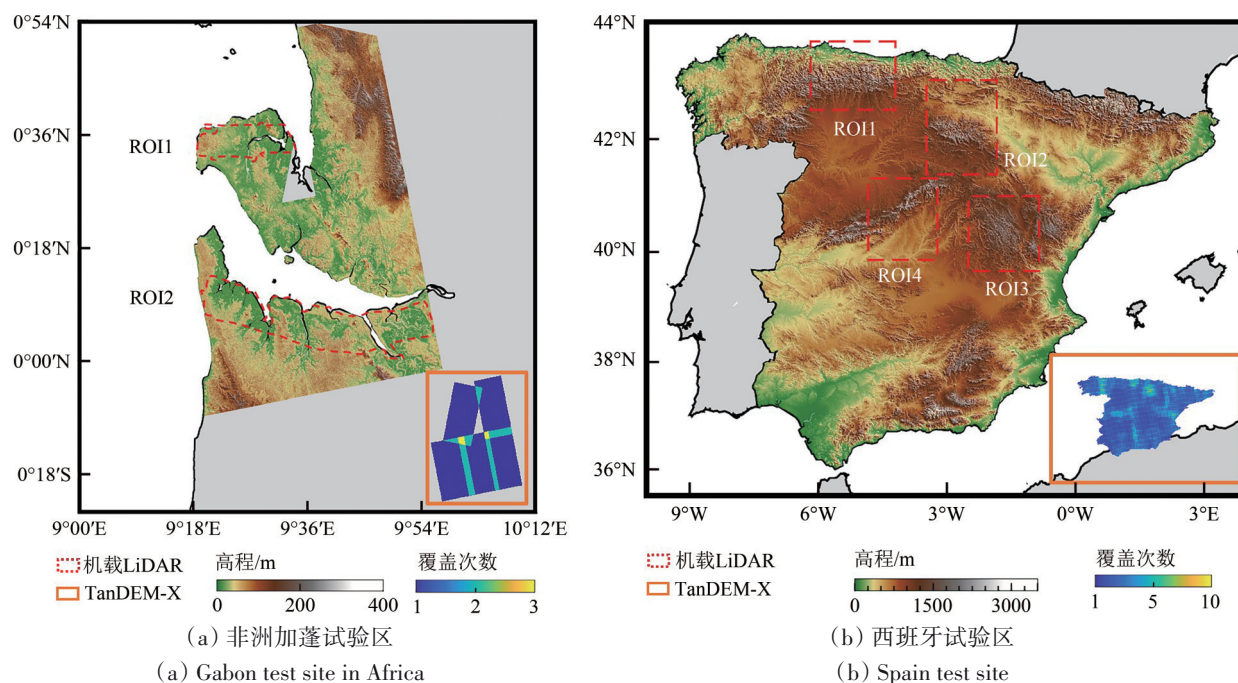


图3 试验区地理位置及所用数据覆盖示意图

Fig. 3 Geographical location of the test area and coverage of the data used

3.1.2 试验数据

图3右下角显示了每个试验区 TanDEM-X 干

涉对的覆盖情况, 其中非洲加蓬试验区包含5个干涉对、西班牙试验区包含1530个干涉对, 详细

信息如表 1 所示。本文采用的 ICESat-2 ATL08 数据是 2018 年 10 月—2022 年 3 月获取的，与 InSAR 数据的平均间隔时间超过 5 年，但是对于森林区域，地形的变化可以忽略不计。使用已有方法对 ATL08 数据进行筛选 (Neuenschwander 等, 2020)，两个试验区分别获取了 14705 个和 79800405 个高程控制点，控制点数量均小于试验区域总像素的 0.1%。

表 1 TanDEM-X InSAR 数据参数

Table 1 TanDEM-X InSAR data parameters

试验区	获取时间	数量	模糊高/m	入射角/(°)
非洲加蓬	2015-11—2015-12	5	73.64—101.21	43.38—46.84
西班牙	2011-09—2014-06	1530	30.67—60.34	30.75—45.86

两个试验区的 Landsat 8 数据分别于 2015 年和 2013 年获取，与 InSAR 数据平均时间相差<1 年，其森林属性的变化忽略不计。经过计算得到与 InSAR DEM 相同分辨率的 NDVI、EVI、VCF 和 RVI 产品 (Hansen 等, 2008; Son 等, 2014)，用以辅助 TanDEM-X 估计林下地形。

此外，NASA 在 2016 年的 AfriSAR 任务期间以及西班牙在 2010 年国家领土观测计划期间分别获取了两个试验区的高精度机载 LiDAR 数据 (Arozarena, 2006; Fatoyinbo 等, 2021)，其覆盖范围见图 3。经过处理后得到与 InSAR DEM 同坐标系和分辨率的 LiDAR DTM，以验证本文所提方法估计林下地形的性能与精度。

3.2 验证结果与评价

3.2.1 非洲加蓬试验区

图 4 是非洲加蓬试验区结果，其中图 4 (a) 是利用 TanDEM-X 数据通过干涉处理获取的 InSAR DEM，分辨率为 10 m。在 InSAR DEM 中可以看到明显的由森林引起的高程偏差，主要表现在裸地与植被区明显的高度梯度，尤其是试验区中部。图 4 (b) 是通过所提出的方法估计的林下地形，森林覆盖区域的高程显著降低，高程梯度明显趋于平滑。估计的相位中心高度如图 4 (c) 所示，范围为 0—30 m，最大值集中在验证区域的中部，这与实际森林高度情况非常相符。

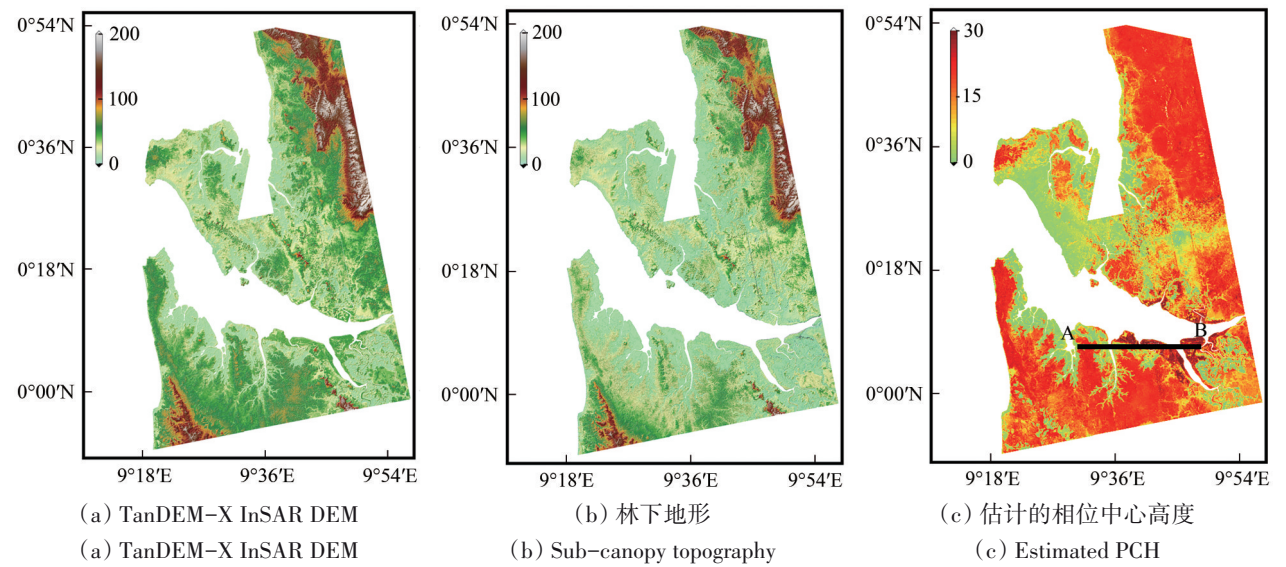


图 4 非洲加蓬试验区试验结果
Fig. 4 The results for the Gabon test site in Africa

此外，通过分析剖面 AB，进一步将获得的林下地形与 LiDAR DTM 进行了比较。如图 5 所示，林下地形比 InSAR DEM 更接近 LiDAR DTM，并且地形趋势几乎一致。然而，可以注意到剖面的右侧部分，林下地形仍然高于真实的地表高程，森林高度信号未完全消除。这是因为在茂密高大的

森林地区，ICESat-2 获取的高程并未在真实的地表，而是高于地表的某一个位置，该现象在已有的研究中已经得到了证实 (董佳臣 等, 2021)。

为了定量评估试验的结果，使用图 3 (a) 所示两个验证区域的 LiDAR DTM 作为参考，统计了森林覆盖区域的林下地形精度。图 6 为 InSAR DEM 和林

下地形与LiDAR DTM的差异分布直方图,并计算了均方根误差RMSE (Root Mean Square Error)、平均绝对误差MAE (Mean Absolute Error)以及标准差STD (STandard Deviation)。可见,相较于LiDAR DTM,两个感兴趣区域ROI (Region Of Interest)获取InSAR DEM包含显著的植被信号,对应的MAE分别为10.13 m和14.46 m,去除森林信号后MAE接近1 m。此外,ROI1和ROI2验证区估计的林下地形的RMSE分别为5.45 m和5.91 m,相比于InSAR DEM (14.70 m和18.58 m)降低了60%以上。

3.2.1 西班牙试验区

图7 (a) 和 (b) 分别显示了西班牙试验区估计的相位中心高度和光学影像。由图7 (a) 可见估计的相位中心高度范围为0—20 m, 较大的值主

要出现在北部沿海地区。由图7 (b) 光学影像可见,估计的相位中心高度与植被的空间分布特征展示出了较好的一致性。

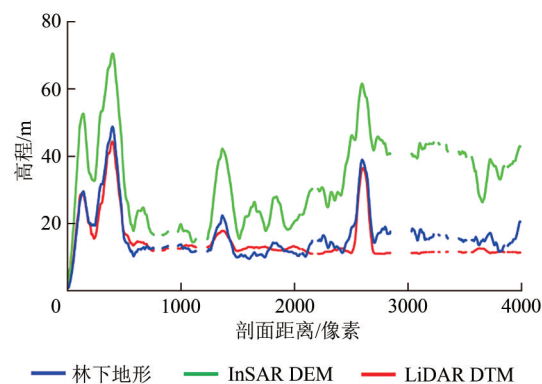


图5 在图4(c)中标注的林下地形剖面AB

Fig. 5 Sub-canopy topography profile AB marked in Fig. 4(c)

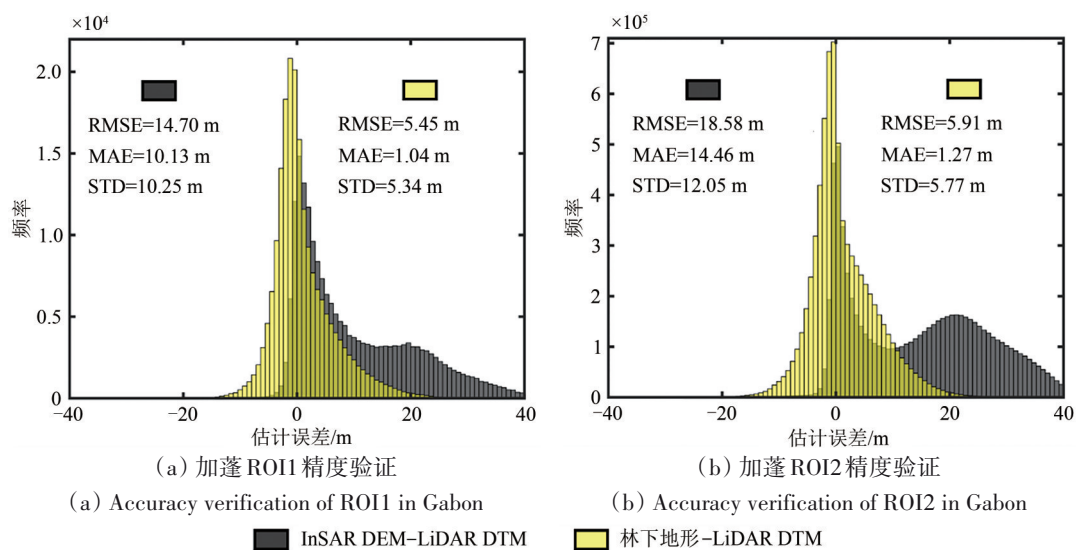


图6 非洲加蓬试验区InSAR DEM和林下地形与LiDAR DTM的高差统计直方图

Fig. 6 Statistical histogram of elevation difference between InSAR DEM and forest terrain and LiDAR DTM in the Gabon test site in Africa

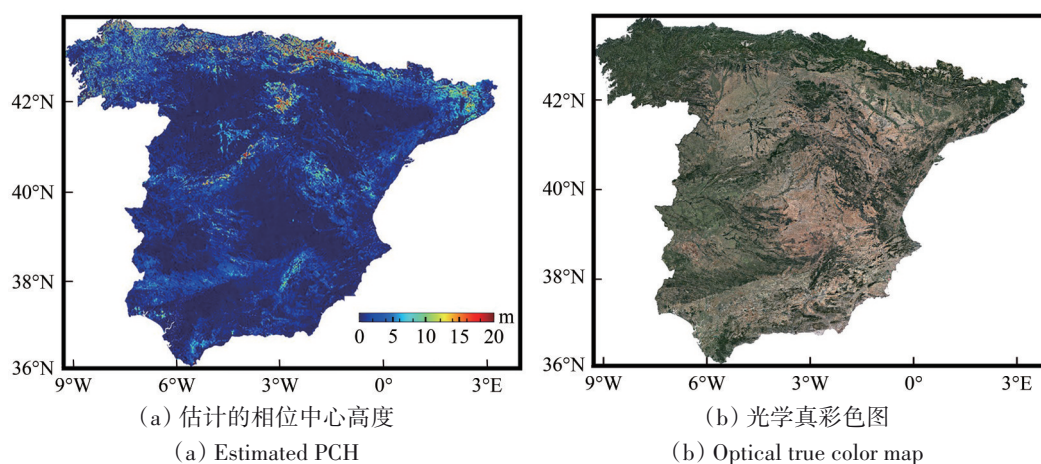


图7 西班牙试验区相位中心高度估计结果

Fig. 7 The estimated PCH of the Spanish test area

图 8 为去除相位中心高度前后与 LiDAR DTM 的对比, 其中图 8 (a) 是西班牙试验区 4 个验证区域的 LiDAR DTM; 图 8 (b) 是 InSAR DEM 和 LiDAR DTM 的差异图。可见, 由于森林体散射的影响, 采用常规 InSAR 技术提取的 DEM 明显高于 LiDAR DTM。图 8 (c) 为所提方法估计的林下地形与 LiDAR DTM 的差异图, 与图 8 (b) 相比, 由森林冠层散射引起的相位中心高度被很好的移除。此外, 对比图 8 (b) 和图 8 (c) 可以观察到, 移除相位中心高度前后非森林区域的高程没有显著变化, 其主要原因在于本文中光学植被参数的引入, 不仅为区分不同的森林属性提供了参考, 也为区分不同的地表覆盖类型 (尤其森林与非森林的区分) 提供了必要的信息。

为了定量评估林下地形的精度, 使用了图 3 (b) 所示 4 个覆盖区域的 LiDAR DTM 作为参考。图 9 显示了去除相位中心高度前后在森林区域的高程误差分布直方图, 并显示了 RMSE、MAE 和 STD 用于直观评价林下地形的精度。对于 4 个验证区域, 与常规的 InSAR DEM 相比, 本文算法提取林下地形的各项指标有了显著的改进。其中 MAE 指标提升最为明显, 林下地形的平均绝对误差接近于零。相比于 InSAR DEM, 林下地形与 LiDAR DTM 差异的标准差也有了显著的下降, 平均下降了 1.58 m。此外, 4 个验证区域的 RMSE 由 6.05—9.10 m 降低到 3.09—4.42 m, 精度提升超过 50%。其中 ROI1 和 ROI4 验证区域的精度指标略微逊色, 主要原因是这两个区域的森林主要分布在地形较为复杂的高山地区。

此外, 按不同的坡度将地表分为平地、丘陵、山地和高山地, 并以 LiDAR DTM 为参考, 对所有像素统计计算了 RMSE, 结果见表 2。相比于 InSAR DEM, 每个地形类别的精度都有了显著的提升, 且在认为 LiDAR DTM 足够准确的条件下, 该试验区估计的林下地形精度可达到 1:2.5 万数字高程模型一级精度标准。

3.3 结果分析及讨论

3.3.1 本文所提方法的优势

基于机器学习联合 ICESat-2 和光学遥感参数的方法已被用于去除 TanDEM-X 任务获取的 30 m 分辨率的哥白尼 DEM 中森林信号 (Hawker 等, 2022), 结果发布在 <https://data.bris.ac.uk/data/>

[dataset/25wfy0f9ukoge2gs7a5mqpq2j7](https://data.bris.ac.uk/data/dataset/25wfy0f9ukoge2gs7a5mqpq2j7) [2023-10-24]。将本文所估计的林下地形重采样到 30 m 分辨率, 控制统计流程和区域一致与已有产品进行了对比, 结果见表 3。可知, 本文所提方法获取的林下地形的精度均优于现有的产品, 特别是西班牙试验区的 RMSE 有显著的提升。精度更优的原因主要包括: (1) 本文所提方法除使用光学遥感参数外, 还使用了 InSAR 自身获取的相干性、后向散射系数以及距离向坡度等参数, 从而克服了单一数据源难以准确刻画相位中心高度空间变化的限制; (2) 与已有研究只训练一个机器学习模型不同 (Kulp 和 Strauss, 2018; Hawker 等, 2022), 本文针对每景 SAR 影像单独进行训练和预测森林信号, 原因在于在区域、国家甚至是全球尺度上, 森林具有异质性, 一个模型难以顾及不同的森林和地形条件。综上, 本文所提方法估计的林下地形更优且在大范围估计时更具稳定性。

3.3.2 ICESat-2 对林下地形估计的影响

本文所提方法估计林下地形的另一个关键在于高精度 ICESat-2 地面高程点的使用, 表 4 总结了每个 LiDAR DTM 覆盖区域的 ICESat-2 精度。可以看出, 非森林区域的 RMSE 接近 1 m, 北方森林区的 RMSE 为 2.15—3.34 m, 热带雨林区的 RMSE 约为 5 m。为了验证 ICESat-2 精度对林下地形估计的影响, 本文选取了西班牙试验区一个地形起伏明显的干涉对进行定量评估。首先提取干涉对范围内所有 ICESat-2 控制点的地理坐标 (经度和纬度); 然后利用二次样条插值提取该坐标的 LiDAR DTM 作为模拟 ICESat-2 控制点高程; 最后控制后续流程一致估计林下地形并进行比较。使用真实和模拟 ICESat-2 控制点估计得到的林下地形结果与 LiDAR DTM 的高程误差分布图如图 10 所示。可以观察到, 使用模拟控制点估计的林下地形更接近于 LiDAR DTM, 尤其是试验区的北部, 相位中心高度被很好的移除, 并且地形复杂地区低估的现象得到了较大改善。此外, 定量评估结果表明, 使用模拟 ICESat-2 地面控制点提取的林下地形, 其 RMSE 由 9.75 m 提高到 2.94 m; 相较于使用真实 ICESat-2 数据提取林下地形 (RMSE=4.51 m), 精度提高了 1.57 m。因此, 如何提升高精度地面参考数据对于提升本文算法的表现至关重要。

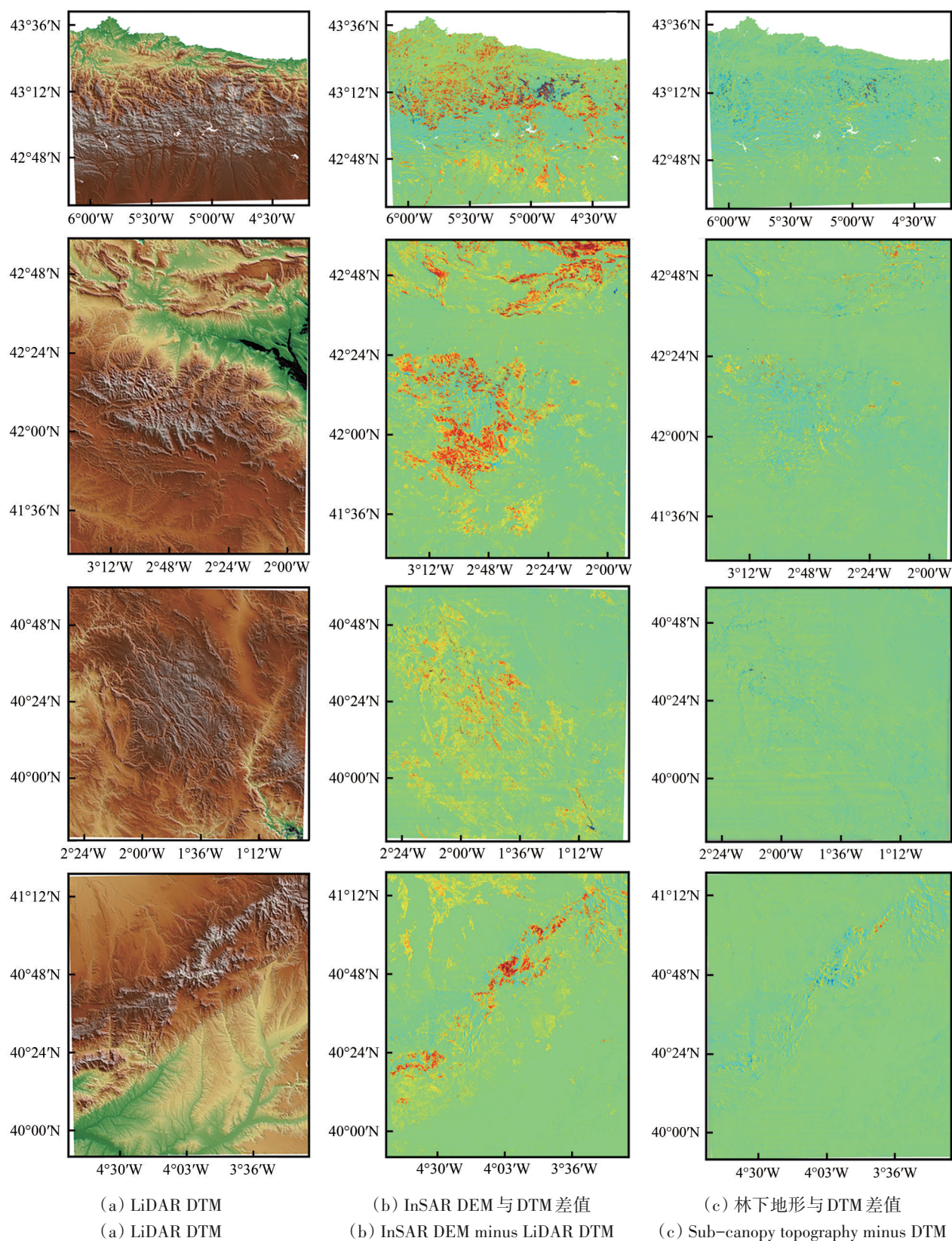


图8 西班牙试验区使用4个ROI的LiDAR DTM与InSAR DEM和林下地形进行对比

Fig. 8 Using LiDAR DTM compared to InSAR DEM and sub-canopy topography in four ROI areas

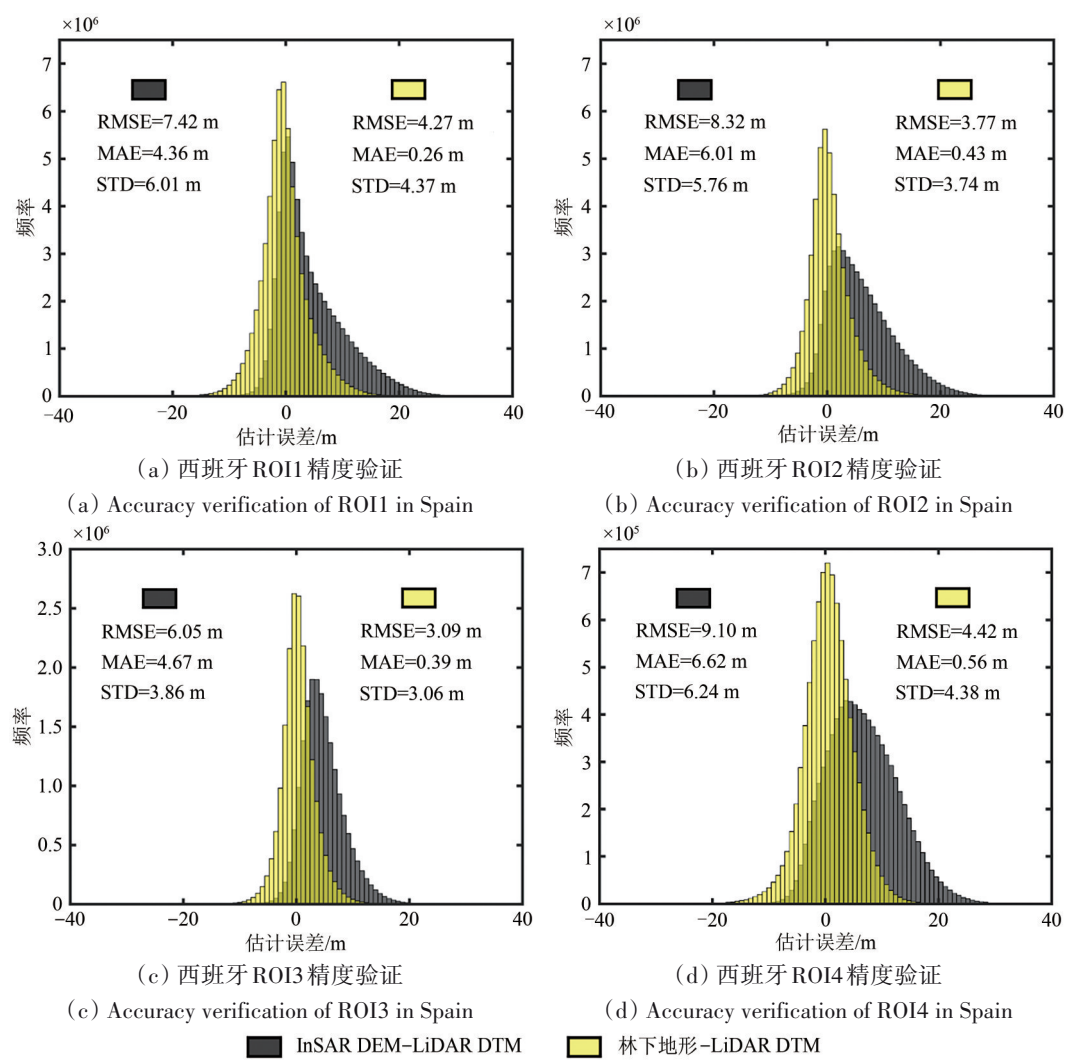


图9 西班牙试验区InSAR DEM和林下地形与LiDAR DTM的高差统计直方图

Fig. 9 Statistical histogram of elevation difference between InSAR DEM and sub-canopy topography and LiDAR DTM in the Spanish test area

表2 西班牙试验区不同地形类别下估计的InSAR DEM和林下地形的精度(RMSE)对比

Table 2 Comparison of estimated InSAR DEM and sub-canopy topography accuracy (RMSE) under different terrain categories in Spain test site

坡度	地形类型	ROI1		ROI2		ROI3		ROI4	
		DEM	林下地形	DEM	林下地形	DEM	林下地形	DEM	林下地形
<2°	平地	2.45	1.33	2.02	1.04	1.52	0.95	2.10	0.99
2°—6°	丘陵	3.21	1.80	2.70	1.32	2.01	1.11	1.75	1.00
6°—25°	山地	4.15	2.76	4.37	2.09	2.67	1.64	3.44	2.06
>25°	高山地	6.98	4.77	6.63	3.65	4.41	3.26	5.84	3.79

表3 所提方法估计林下地形与现有方法对比

Table 3 The sub-canopy topography estimated by the proposed method compared with existing method

试验区	验证区域	InSAR DEM		林下地形		FAB_DEM	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
非洲加蓬	ROI1	14.25	10.11	5.53	1.05	5.66	2.71
	ROI1	18.79	14.06	5.68	1.23	7.46	3.44
西班牙	ROI1	7.31	4.24	4.31	0.19	5.65	0.87
	ROI2	8.51	5.95	3.75	0.51	5.94	0.58
	ROI3	6.01	4.53	3.02	0.35	4.78	0.11
	ROI4	9.05	6.57	4.45	0.61	6.23	0.26

表4 使用LiDAR DTM验证筛选后的ICESat-2控制点精度
Table 4 Validation of screened ICESat-2 control point accuracy using LiDAR DTM

地表类型	非洲加蓬		西班牙			
	ROI1	ROI2	ROI1	ROI2	ROI3	ROI4
非森林	0.81	0.75	0.97	1.38	1.23	1.16
森林	5.23	4.39	3.25	2.81	2.15	3.34

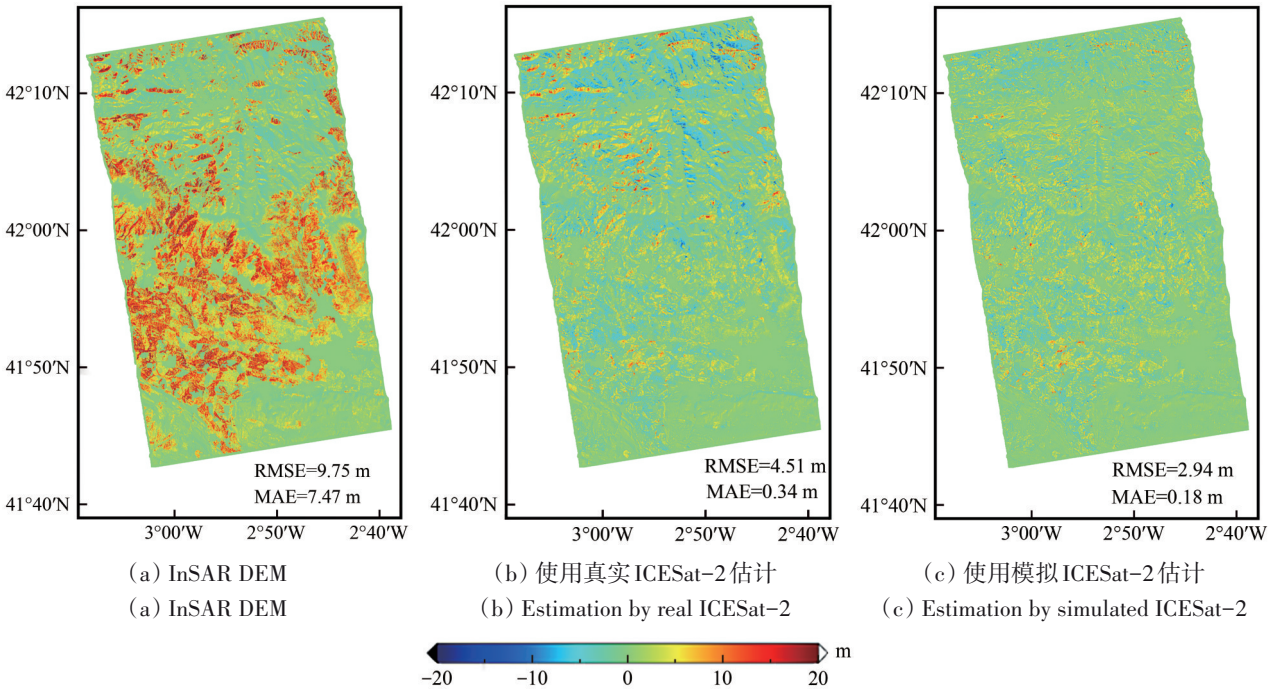


图10 使用真实和模拟的ICESat-2控制点估计的林下地形结果与LiDAR DTM的差值图
Fig. 10 Difference map between estimated sub-canopy topography and LiDAR DTM using real and simulated ICESat-2 control points

4 结 论

针对常规 InSAR 技术无法准确获取森林区林下地形这一问题，本研究提出了一种联合 TanDEM-X InSAR 数据、ICESat-2 和 Landsat 8 数据的大范围林下地形提取方法，并选取了位于非洲加蓬和西班牙两个试验区对本研究模型进行验证。主要结论如下：

- (1) 由于 InSAR 数据的观测信息有限，仅使用 InSAR 观测值难以准确的分离 DEM 中包含的森林信号，引入光学遥感获取的森林参数增加了 InSAR 观测值对相位中心高度的区分能力，更好的估计了林下地形。此外，与现有产品相比，本文方法估计的林下地形精度有显著的提升。
- (2) ICESat-2 控制点的精度是制约估计高精度林下地形的主要因素。我们的结果表明，针对北方森林区域，即使使用现有的控制点筛选策略

也能基本满足 1：2.5 万比例尺 DEM 的要求。然而，为了进一步提升林下地形的估计精度，后续工作不仅需要保证控制点的数量，如何提高地面控制点的精度也值得进一步研究。

(3) 本文方法的实际意义在于开发了单极化、单基线 InSAR 数据条件下估计大范围林下地形的稳定框架，能为国内的天绘系列、陆探一号双站 SAR 卫星以及未来的 TanDEM-L 任务等估计林下地形提供方法支撑。

志 谢 感谢德国宇航局 DLR (German Aerospace Center) 为本项研究工作提供的 TanDEM-X 数据；感谢西班牙阿利坎特大学和林业部门提供的西班牙试验区 LiDAR 实测数据；感谢美国航空航天局 NASA (National Aeronautics and Space Administration) 提供的非洲加蓬试验区 LiDAR 实测数据。

参考文献 (References)

- Abrams M, Crippen R and Fujisada H. 2020. ASTER global digital elevation model (GDEM) and ASTER global water body dataset (AST-WBD). *Remote Sensing*, 12(7): 1156 [DOI: 10.3390/rs12071156]
- Arozarena A. 2006. The Spanish National Territory Observation Program: current states and next steps. (in Spanish) *Mapping Inter*: 16-22
- Ben Jabeur S, Gharib C, Mefteh-Wali S and Ben Arfi W. 2021. Cat-Boost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, 166: 120658 [DOI: 10.1016/j.techfore.2021.120658]
- Cloude S R and Papathanassiou K P. 1998. Polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5): 1551-1565 [DOI: 10.1109/36.718859]
- Dall J. 2007. InSAR elevation bias caused by penetration into uniform volumes. *IEEE Transactions on Geoscience and Remote Sensing*, 45(7): 2319-2324 [DOI: 10.1109/TGRS.2007.896613]
- Dias De Paiva R C, Buarque D C, Collischonn W, Bonnet M P, Frappart F, Calmant S and Bulhões Mendes C A. 2013. Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 49(3): 1226-1243 [DOI: 10.1002/wrcr.20067]
- Dong J C, Ni W J, Zhang Z Y and Sun G Q. 2021. Performance of ICESat-2 ATL08 product on the estimation of forest height by referencing to small footprint LiDAR data. *National Remote Sensing Bulletin*, 25(6): 1294-1307 (董佳臣, 倪文俭, 张志玉, 孙国清. 2021. ICESat-2 植被冠层高度和地表高程数据产品用于森林高度提取的效果评价. *遥感学报*, 25(6): 1294-1307) [DOI: 10.11834/jrs.20219449]
- Fatoyinbo T, Armston J, Simard M, Saatchi S, Denbina M, Laval M, Hofton M, Tang H, Marselis S, Pinto N, Hancock S, Hawkins B, Duncanson L, Blair B, Hansen C, Lou Y L, Dubayah R, Hensley S, Silva C, Poulsen J R, Labrière N, Barbier N, Jeffery K, Kenfack D, Herve M, Bissengou P, Alonso A, Moussavou G, White L T J, Lewis S and Hibbard K. 2021. The NASA AfriSAR campaign: Airborne SAR and lidar measurements of tropical forest structure and biomass in support of current and future space missions. *Remote Sensing of Environment*, 264: 112533 [DOI: 10.1016/j.rse.2021.112533]
- Fu H Q, Wang C C, Zhu J J, Xie Q H and Zhang B. 2016. Estimation of pine forest height and underlying DEM using multi-baseline P-band PolInSAR data. *Remote Sensing*, 8(10): 820 [DOI: 10.3390/rs8100820]
- Fu H Q, Zhu J J, Wang C C, Wang H Q and Zhao R. 2017. Underlying topography estimation over forest areas using high-resolution P-band single-baseline PolInSAR data. *Remote Sensing*, 9(4): 363 [DOI: 10.3390/rs9040363]
- Garestier F, Dubois-Fernandez P C and Champion I. 2008. Forest height inversion using high-resolution P-band pol-InSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11): 3544-3559 [DOI: 10.1109/TGRS.2008.922032]
- Gruber A, Wessel B, Huber M and Roth A. 2012. Operational TanDEM-X DEM calibration and first validation results. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 39-49 [DOI: 10.1016/j.isprsjprs.2012.06.002]
- Hansen M C, Roy D P, Lindquist E, Aducci B, Justice C O and Altstatt A. 2008. A method for integrating MODIS and landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sensing of Environment*, 112(5): 2495-2513 [DOI: 10.1016/j.rse.2007.11.012]
- Hawker L, Uhe P, Paulo L, Sosa J, Savage J, Sampson C and Neal J. 2022. A 30 m global map of elevation with forests and buildings removed. *Environmental Research Letters*, 17(2): 024016 [DOI: 10.1088/1748-9326/ac4d4f]
- Kulp S A and Strauss B H. 2018. CoastalDEM: a global coastal digital elevation model improved from SRTM using a neural network. *Remote Sensing of Environment*, 206: 231-239 [DOI: 10.1016/j.rse.2017.12.026]
- Lei Y, Treuhaft R and Gonçalves F. 2021. Automated estimation of forest height and underlying topography over a Brazilian tropical forest with single-baseline single-polarization TanDEM-X SAR interferometry. *Remote Sensing of Environment*, 252: 112132 [DOI: 10.1016/j.rse.2020.112132]
- Lopez-Sanchez J M, Vicente-Guijalba F, Erten E, Campos-Taberner M and Garcia-Haro F J. 2017. Retrieval of vegetation height in rice fields using polarimetric SAR interferometry with TanDEM-X data. *Remote Sensing of Environment*, 192: 30-44 [DOI: 10.1016/j.rse.2017.02.004]
- Magruder L, Neuenschwander A and Klotz B. 2021. Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry. *Remote Sensing of Environment*, 264: 112621 [DOI: 10.1016/j.rse.2021.112621]
- Martone M, Brautigam B and Krieger G. 2015. Quantization effects in TanDEM-X data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2): 583-597 [DOI: 10.1109/TGRS.2014.2325976]
- Neuenschwander A, Guenther E, White J C, Duncanson L and Montesano P. 2020. Validation of ICESat-2 terrain and canopy heights in boreal forests. *Remote Sensing of Environment*, 251: 112110 [DOI: 10.1016/j.rse.2020.112110]
- O'Loughlin F E, Paiva R C D, Durand M, Alsdorf D E and Bates P D. 2016. A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sensing of Environment*, 182: 49-59 [DOI: 10.1016/j.rse.2016.04.018]
- Papathanassiou K P and Cloude S R. 2001. Single-baseline polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11): 2352-2363 [DOI: 10.1109/36.964971]
- Prăvălie R, Sîrodoev I, Nita I A, Patriche C, Dumitrașcu M, Roșca B, Tișcovschi A, Bandoc G, Săvulescu I, Mănoiu V and Birsan M V. 2022. NDVI-based ecological dynamics of forest vegetation and its relationship to climate change in Romania during 1987-2018. *Ecological Indicators*, 136: 108629 [DOI: 10.1016/j.ecolind.2022.108629]
- Rahman M M, Csaplovics E and Koch B. 2005. An efficient regression strategy for extracting forest biomass information from satellite sensor data. *International Journal of Remote Sensing*, 26(7): 1511-1519 [DOI: 10.1080/01431160500044705]
- Rizzoli P, Martone M, Gonzalez C, Wecklich C, Tridon D B, Bräutigam B, Bachmann M, Schulze D, Fritz T, Huber M, Wessel B,

- Krieger G, Zink M and Moreira A. 2017. Generation and performance assessment of the global TanDEM-X digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132: 119-139 [DOI: 10.1016/j.isprsjprs.2017.08.008]
- Schlund M, Baron D, Magdon P, Erasmi S, 2019. Canopy penetration depth estimation with TanDEM-X and its compensation in temperate forests. *Isprs Journal of Photogrammetry and Remote Sensing*, 147: 232-241 [DOI:10.1016/j.isprsjprs.2018.11.021]
- Son N T, Chen C F, Chen C R, Minh V Q and Trung N H. 2014. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agricultural and Forest Meteorology*, 197: 52-64 [DOI: 10.1016/j.agrformet.2014.06.007]
- Su Y J, Guo Q H, Ma Q and Li W K. 2015. SRTM DEM correction in vegetated mountain areas through the integration of spaceborne LiDAR, airborne LiDAR, and optical imagery. *Remote Sensing*, 7(9): 11202-11225 [DOI: 10.3390/rs70911202]
- Treuhaft R N, Madsen S N, Moghaddam M and Van Zyl J J. 1996. Vegetation characteristics and underlying topography from interferometric radar. *Radio Science*, 31(6): 1449-1485 [DOI: 10.1029/96RS01763]
- Wang H Q, Fu H Q, Zhu J J, Liu Z W, Zhang B, Wang C C, Li Z W, Hu J and Yu Y N. 2021. Estimation of subcanopy topography based on single-baseline TanDEM-X InSAR data. *Journal of Geodesy*, 95(7): 84 [DOI: 10.1007/s00190-021-01519-3]
- Xu J L. 2022. Discussion on key technologies of obtaining high-precision DEM based on airborne LiDAR data. *Geospatial Information*, 20(11): 5-8 (许家录. 2022. 基于机载 LiDAR 数据获取高精度 DEM 的关键技术探讨. *地理空间信息*, 20(11): 5-8) [DOI: 10.3969/j.issn.1672-4623.2022.11.002]
- Zhang T, Zhu J J, Fu H Q and Wang C C. 2022. Forest height inversion with single-baseline TanDEM-X InSAR coherence. *Acta Geodaetica et Cartographica Sinica*, 51(9): 1931-1941 (张涛, 朱建军, 付海强, 汪长城. 2022. 单基线 TanDEM-X InSAR 相干性反演森林高度. *测绘学报*, 51(9): 1931-1941) [DOI: 10.11947/j.AGCS.2022.20210298]
- Zhu J J, Fu H Q and Wang C C. 2018. Methods and research progress of underlying topography estimation over forest areas by InSAR. *Geomatics and Information Science of Wuhan University*, 43(12): 2030-2038 (朱建军, 付海强, 汪长城. 2018. InSAR 林下地形测绘方法与研究进展. *武汉大学学报(信息科学版)*, 43(12): 2030-2038) [DOI: 10.13203/j.whugis20180266]

Large-scale sub-canopy topography estimation from TanDEM-X InSAR and ICESat-2 data using machine learning method

HU Huacan¹, ZHU Jianjun¹, FU Haiqiang¹, LOPEZ-SANCHEZ Juan Manuel², GÓMEZ Cristina^{3,4}, ZHANG Tao¹, LIU Kui¹

1. School of Geosciences and Info-Physics, Central South University, Changsha 410083, China;

2. Institute for Computer Research, University of Alicante, Alicante 03690, Spain;

3. iuFOR-EiFAB, University of Valladolid, Campus Duques de Soria 47002, Spain;

4. School of Geoscience, University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK

Abstract: Digital Elevation Models (DEMs) are indispensable data sources for natural resource investigation, climate change analysis, and disaster monitoring and assessment. TanDEM-X mission, as the first twin-satellite Interferometric Synthetic Aperture Radar (InSAR) system, has successfully obtained a high-precision global DEM with 12 m resolution. However, limited by the penetration capability of shortwave signals, DEMs acquired in dense forest areas are usually contaminated by forest canopy signals and are difficult to meet practical applications. The Phase Center Height (PCH) caused by forest volume scattering needs to be removed from InSAR-derived DEM to obtain sub-canopy topography. Unfortunately, TanDEM-X acquires single-baseline, single-polarization data in the global standard mode, which is difficult to meet the needs of existing model solutions and requires the introduction of external data. In this study, we propose a machine learning-based method to estimate sub-canopy topography by combining TanDEM-X InSAR, ICESat-2, and Landsat 8 OLI data. The effectiveness of the proposed method was tested and validated in the Gabon rainforest and the Spanish boreal forest. In the Gabon rainforest test site, compared with that of the airborne LiDAR Digital Terrain Model (DTM), the Root-Mean-Square errors (RMSEs) of the InSAR DEMs corresponding to two locations are 14.70 m and 18.58 m. After PCH removal, the accuracy is improved to 5.54 m and 5.86 m, which represents an improvement of over 60%. In the Spanish northern forest test site with complex terrain, the RMSE of sub-canopy topography decreased from 6.05—9.10 m to 3.06—4.42 m. In addition, we investigate the necessity of the proposed method to use InSAR observations and the effect of the accuracy of the ICESat-2 control points used on the sub-canopy topography estimation. These satisfactory results demonstrate the potential of the proposed method in estimating sub-canopy topography for future spaceborne InSAR missions (e.g., TanDEM-L and LT-1) when only single-baseline, single-polarization data are available. Furthermore, by combining the high resolution of TanDEM-X and the strong penetration of BIOMASS, the proposed method has the potential to estimate sub-canopy topography with higher accuracy and resolution in the future.

Key words: sub-canopy topography, InSAR, phase center height, machine learning, TanDEM-X, ICESat-2

Supported by National Natural Science Foundation of China (No. 41820104005, 42030112); National Natural Science Foundation of Hunan (No. 2021JJ30808)