引文格式:林其炜,李志伟,付海强.顾及地形特征的InSAR和光学DSM逐步回归融合[J].武汉大学学报(信息科学版),2026,51(7):1494-1506.DOI:10.13203/j.whugis20250131
Citation:LIN Qiwei,LI Zhiwei,FU Haiqiang.Stepwise Regression Fusion of InSAR and Optical DSM Considering Terrain Features[J].Geomatics and Information Science of Wuhan University,2026,51(7):1494-1506.DOI:10.13203/j.whugis20250131

顾及地形特征的InSAR和光学DSM逐步回归融合

林其炜[1] 李志伟[1] 付海强[1]
1 中南大学地球科学与信息物理学院,湖南 长沙,410083

摘 要:合成孔径雷达干涉测量(interferometric synthetic aperture radar, InSAR)和光学摄影测量是生成数字表面模型(digital surface model, DSM)的两种主要技术,但InSAR技术容易受到几何畸变影响,光学立体测量在云雨等恶劣天气下的观测能力受限。多源DSM融合能有效利用数据互补的优势,可提升单一数据源数据生成DSM的质量。然而,传统加权平均融合方法的权重分配依赖先验信息,且未充分考虑地形特征对高程误差的影响。因此,提出了一种顾及地形特征的InSAR和光学DSM逐步回归融合方法。以InSAR和光学DSM为研究对象,提取坡度、坡向、粗糙度等地形特征参数,采用逐步回归方法建立高程误差与地形特征间的回归模型,并基于模型自适应计算融合权重。采用美国山区的TanDEM-X DSM和AW3D30 DSM数据开展融合实验,并在两个实验区对结果进行了测试。实验结果表明,所提方法在两个测试区域中的均方根误差分别降低了14.3%和18%,优于传统加权平均融合方法;相比于随机森林方法,所提方法在测试区域达到更高融合精度的同时,具有一定的可解释性。

关键词:DSM;数据融合;地形特征;逐步回归
中图分类号:P237 **文献标识码**:A **收稿日期**:2025-07-21
DOI:10.13203/j.whugis20250131 **文章编号**:1671-8860(2026)07-1494-13

Stepwise Regression Fusion of InSAR and Optical DSM Considering Terrain Features

LIN Qiwei[1] *LI Zhiwei*[1] *FU Haiqiang*[1]
1 School of Geosciences and Info-Physics, Central South University, Changsha 410083, China

Abstract: **Objectives**: Interferometric synthetic aperture radar (InSAR) and optical photogrammetry are two primary techniques for generating digital surface model (DSM). However, InSAR is susceptible to geometric distortion, while optical photogrammetry is limited by cloudy and rainy weather. The fusion of multi-source DSM can effectively leverage the complementary advantages of the data to enhance the quality of DSM derived from a single data source. Traditional weighted average fusion methods rely on prior information for weight allocation and don't adequately consider the impact of terrain features on elevation errors. There is a need to develop a new fusion method. **Methods**: This paper proposes a stepwise regression fusion method for InSAR and optical DSM that takes topographic features into account. Taking InSAR and optical DSM as research subjects, and considering the commonalities and differences in error characteristics between InSAR and optical DSM, terrain geometric features, land cover features, and elevation difference features are extracted as training samples in the training area. Subsequently, the stepwise regression method is employed to construct a regression model between elevation errors and terrain features by incrementally adding or removing variables. Elevation errors are then predicted to adaptively determine the fusion weights. Finally, fusion experiments are conducted using TanDEM-X DSM and AW3D30 DSM data from mountainous

基金项目:国家自然科学基金重点项目(42330112);湖南省重点研发计划(2023SK2012)。
第一作者:林其炜,硕士,研究方向为多源DSM融合、InSAR DSM数据处理。qw_lin@163.com
通信作者:李志伟,博士,教授。zwli@csu.edu.cn

areas in the United States, and the results are tested in two different experimental regions. **Results:** The proposed method outperforms the traditional weighted averaging fusion method. The root mean square errors decrease by 14.3% and 18% in the two test regions, respectively. Compared with the random forest method, the proposed method achieves higher fusion precision while maintaining favorable interpretability. **Conclusions:** The proposed method effectively improves the quality of fused DSM by utilizing terrain features and achieves favorable performance in complex terrain areas, providing an effective solution for multi-source DSM fusion.

Key words: DSM; data fusion; terrain features; stepwise regression

数字表面模型(digital surface model, DSM)是一种重要的基础地形数据,广泛应用于地形测绘、水文分析、灾害监测等领域[1]。DSM数据常见的获取方法有野外实测、地形图数字化、合成孔径雷达干涉测量(interferometric synthetic aperture radar, InSAR)、光学摄影测量和激光雷达测高(light detection and ranging, LiDAR)等技术[2-3]。其中,InSAR和光学摄影测量技术因具备大范围、连续覆盖的能力,已成为全球DSM生产的主流技术。然而,DSM生产的数据质量常受限于观测手段、生成方法、天气条件和地形特征等因素[4]。光学数据易受云雾、水体干扰,导致异常值和数据缺失;雷达DSM数据在陡坡区域易出现数据空洞[5],且相位解缠困难会导致高程反演失败[6]。相比之下,多源DSM数据融合能综合异源数据的互补优势,生成更高质量的DSM数据[7]。因此,结合InSAR的云雾穿透能力和光学测量的高精度特性,通过InSAR DSM和光学DSM融合来满足高精度地形数据的应用需求,成为了近年来的研究热点之一。

多源DSM数据融合往往通过分配不同的权重来实现[8],国内外学者在这方面开展了大量工作。目前常用的方法可以分为两类,分别是机器学习融合方法和加权平均融合方法。机器学习融合方法即利用神经网络[9-12]构建高程误差模型,以此学习并优化DSM融合的权重分配。文献[13]通过对比多种机器学习方法,发现随机森林在校正航天飞机雷达地形测绘任务(shuttle radar topography mission, SRTM)DSM高程误差时的精度提升效果较好。然而,机器学习方法存在模型解释性不足的问题,难以满足实际应用需求。加权平均融合方法因融合效果好且计算成本低而被广泛使用[14],其中,最为简便的是等权的简单平均法[15]。如果数据都来自于同一个传感器,则平均后可以得到良好的结果[16-17],但该方法未考虑输入DSM的质量差异。另外有一种特例加权方法,即遍历权重寻优法[7,18-19],根据平均后相对于参考DSM的均方根误差(root mean square error, RMSE)表现,来遍历探求最佳的融合系数。相比于这两种平均方法,更为可靠的是考虑输入数据质量的加权平均融合方法,该方法通过权重来定量化描述输入数据的相对精度[20],以获取融合的最优解。文献[21]利用TanDEM-X数据采集过程中的高程误差图(height error map, HEM)确定权重系数,融合了升轨和降轨TanDEM-X DSM。文献[22]在不同区域的多幅TanDEM-X DSM上进行融合实验,基于DSM数量和质量参数设置优先级,再用HEM作为权重进行加权平均,显著提高了DSM的垂直精度。然而,加权平均融合方法在定权时存在局限性。一方面,高精度参考DSM和理论误差图难以获取,且HEM难以涵盖实际生产中的所有误差;另一方面,地形特征对DSM高程误差有显著影响[23-26],而现有融合研究未充分考虑建立特征-误差定量映射的回归模型融合方法,并且在特征选择方面表征维度单一,通常基于坡度划分地形类型,从而设定经验权重。例如,文献[27]依据坡度值划分地貌类型并设置固定权重,对先进星载热发射和反射辐射计全球数字高程模型和SRTM进行融合;文献[28]对相同地形类别的高程点采用相同的权重,未对各高程点逐一评估精度以确定权重。这种方法未能充分考虑地形特征与高程误差之间的复杂关系,也未能充分利用多源数据之间的差异性,导致在复杂地形区域的融合效果欠佳。

因此,本文将回归分析引入到DSM融合过程中,并选取具有代表性的地形特征来构建高程误差回归模型,以实现动态权重分配。逐步回归方法能够自动筛选特征,可有效处理特征多重共线性并防止过拟合,因此,本文设计了一种顾及地形特征的InSAR与光学DSM数据逐步回归融合方法。首先,选取地形几何特征、地表覆盖和高程差异信息作为训练特征,以高精度LiDAR数据作为参考,构建高程误差回归模型。然后,将预测的高程误差作为定权的依据,从而实现多源DSM数据的自适应加权平均融合,得到更高精度的DSM数据。

1 顾及地形特征的DSM融合方法

1.1 地形相关特征提取

为了构建误差模型,本文系统构建了多维地形特征空间。考虑到InSAR和光学DSM在误差特性上的共性与差异,特征提取主要包含3类信息,分别为地形几何特征、地表覆盖特征和高程差异特征。

在众多地形因子中,本文选取了具有代表性的坡度、坡向、粗糙度和地形位置指数(topographic position index, TPI)作为训练的地形几何特征。坡度表示DSM每个像元的地面倾斜程度,是地表高程变化率的直接反映,通常基于像元与邻域像元的高程差来计算坡度,计算公式为:

$$S=\arctan\sqrt{f_x^2+f_y^2} \tag{1}$$

式中,S为坡度;f_x为X方向上的坡度;f_y为Y方向上的坡度。坡向是指地形坡面的朝向,即各像元高程变化率最大的下坡方向。坡向大小为[0°, 360°],按顺时针方向进行测量,计算公式为:

$$A=270^\circ+\arctan\frac{f_y}{f_x}-90^\circ\frac{f_x}{|f_x|} \tag{2}$$

式中,A为坡向。粗糙度是指地表某个特定曲面单元的表面积与其在水平面上的投影面积的比例,可以描述地表起伏变化程度,可用于揭示地形表面的侵蚀情况,计算公式为:

$$R=\frac{S_{\text{surf}}}{S_{\text{proj}}} \tag{3}$$

式中,R为地表粗糙度;S_{surf}为地表单元曲面的面积;S_{proj}为该单元在水平面上的投影面积。TPI是指某像元与预定范围内平均高程的差值,用于描述其与周围像元的高程相对位置关系。TPI值的正负表示该像元相对位置的高低,为了考虑数据的全局特性,采用DSM数据的整体均值进行计算,计算公式为:

$$I_{\text{TPI}}=Z_0-\frac{1}{n_R}\sum_{i\in R}Z_i \tag{4}$$

式中,Z_0为某一像元的高程值;R为设定的邻域范围;Z_i表示在该范围内各像元点的高程值;n_R为邻域内高程点的数量。地表覆盖特征用归一化差异植被指数(normalized difference vegetation index, NDVI)进行描述,NDVI是通过遥感图像计算得到的植被覆盖指标,反映了地表植被的密度和健康状况。NDVI值介于−1~1之间,值越高,表明植被覆盖越茂密;值越低,则表明植被覆盖越稀疏或为裸地。NDVI的计算公式为:

$$I_{\text{NDVI}}=\frac{B_{\text{NIR}}-B_{\text{red}}}{B_{\text{NIR}}+B_{\text{red}}} \tag{5}$$

式中,B_{NIR}为某一像元近红外(near infrared, NIR)波段的像元值;B_{red}为红光波段的像元值。此外,本文还考虑了高程差异特征,如DSM高程值、InSAR与光学DSM的高程差异以及DSM与外部参考DSM的残差,这些特征有助于模型识别不同高程区域的误差分布规律。在外部参考DSM的选择上,鉴于哥白尼DSM具有较高的精度和良好的可获取性,本文在特征提取时将其作为初步参考,为误差模型的准确构建提供保障。

为充分挖掘多源DSM融合的优势,本文还考虑了源数据的特有误差。例如,InSAR DSM误差与相干性密切相关,相干性直接决定干涉相位的质量,进而影响高程精度。在TanDEM-X全球DSM科学计划中,德国宇航局根据TanDEM-X DSM的相干性高程误差制作了HEM,并基于该误差图确定权重,实现大范围DSM融合。这种理论高程误差主要与干涉相干性有关,因此,本文方法也将相干性高程误差作为InSAR DSM的高程差异表征,TanDEM-X DSM相干性高程误差σ_j的计算公式为[29-30]:

$$\sigma_j=H_{\text{amb}}\frac{\sigma_{\phi,j}}{2\pi} \tag{6}$$

式中,H_{amb}是模糊高度;$\sigma_{\phi,j}$是基于干涉相干性和InSAR几何估计的干涉相位误差[31]。

在提取与地形相关的数据特征后,数据清洗是一个重要的处理步骤,旨在剔除数据中的离群点,减少噪声对模型构建及最终测试效果的干扰。采用稳健的归一化中值绝对偏差(normalized median absolute deviation, NMAD)[32]作为精度检测指标。具体而言,若InSAR和光学DSM数据与参考数据之间的残差超过两倍的NMAD,则将这些数据点视为不可靠的离群值,并予以剔除。

1.2 逐步回归模型构建

逐步回归是一种基于统计学习的特征选择方法,广泛应用于回归模型的构建[33]。它结合了向前选择和向后剔除两种策略,通过逐步添加或删除变量,优化模型的拟合效果和复杂度。在提取误差特征后,以LiDAR DSM作为高程真值,将InSAR或光学DSM与LiDAR DSM之间的误差设为因变量,提取的地形特征设为自变量,利用逐步回归方法分别构建InSAR和光学DSM的误差模型。

1)初始化回归模型

从包含所有特征变量的线性项和二次项的

二次多项式模型开始,建立误差-地形响应函数:

$$\hat{\varepsilon}=\beta_0+\sum_{i=1}^{p}\beta_i T_i+\sum_{j=1}^{q}\gamma_j T_j^2+\sum_{k<l}\theta_{kl}T_k T_l \quad (7)$$

式中,$\hat{\varepsilon}$为预测的误差值;β_0为模型的截距项;p和q分别表示线性项和二次项的数量;T_i为线性关系特征变量;T_j为二次项特征变量;T_k和T_l分别为存在交互关系的特征变量;β_i、γ_j和θ_{kl}分别为T_i、T_j和T_kT_l的系数。模型通过贝叶斯信息准则(Bayesian information criterion, BIC)动态选择显著项。

2)逐步选择变量

首先进行向前选择,从初始模型开始,逐步添加变量。每次添加一个使模型改进最大的变量,直到没有更多的变量可以显著改进模型。变量的选择基于统计检验或信息准则,这里采用计算所有候选特征的F统计量,计算公式为:

$$F=\frac{(E_r-E_f)/(d_r-d_f)}{E_f/d_f} \quad (8)$$

式中,E为误差平方和(sum of squared errors, SSE);d为自由度;下标r和f分别表示简化模型和完整模型。当$F>F_{\text{enter}}$时,添加最显著项,其中,F_{enter}为BIC动态调整的阈值。

随后进行向后剔除,从一个包含所有潜在变量的全模型开始,逐步删除变量,每次删除一个对模型贡献最小的变量,直到所有剩余变量都对模型有显著贡献。变量的剔除同样基于统计检验或信息准则,当$F<F_{\text{remove}}$时,移除最不显著项,其中,F_{remove}为BIC动态调整的阈值。BIC的计算公式为:

$$I_{\text{BIC}}=n\ln\frac{E}{n}+k\ln n \quad (9)$$

式中,n为训练样本数;k为模型参数个数。算法持续迭代,直至BIC值达到局部最小。

3)模型优化

通过设置上下界和选择标准,逐步回归方法在模型复杂度和拟合效果之间取得了平衡。上下界定义了模型的最小和最大形式,而选择标准则用于评估变量的显著性。常用的准则包括SSE、决定系数、调整后的决定系数等。本文将BIC作为选择标准来动态调整模型的复杂度。

在进行DSM融合时,采用逐步回归方法可以突破传统固定权重融合的局限,通过非线性回归捕捉地形-误差的复杂关系,并基于统计显著性检验来剔除冗余特征。逐步回归模型的构建流程如图1所示。

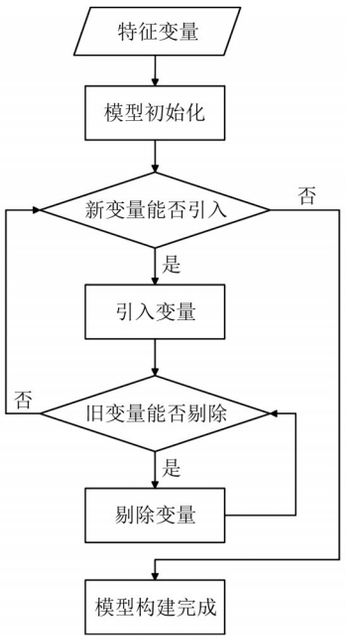


图1　逐步回归模型的构建流程

Fig. 1　Construction Process of Stepwise Regression Model

1.3　InSAR和光学DSM加权平均融合

利用逐步回归模型预测出InSAR和光学DSM的误差估计值后,依据这些误差估计值校正并实现权重的自适应分配。具体而言,误差越小,表明该数据源在相应位置的可靠性越高,因此,应赋予其更大的权重。基于误差估计进行权重分配,进而完成多源DSM的加权平均融合。为了确保权重的合理性,需对权重进行归一化处理,归一化后的权重计算公式为:

$$w_1=\frac{s_1^2}{s_1^2+s_2^2} \quad (10)$$

$$w_2=\frac{s_2^2}{s_1^2+s_2^2} \quad (11)$$

式中,w_1为InSAR DSM的权重;w_2为光学DSM的权重;s_1、s_2分别为InSAR DSM、光学DSM预测误差的倒数。

将归一化后的权重应用于每个数据源的高程值,得到加权后的高程值。利用计算得到的权重,对InSAR和光学DSM进行加权平均融合,融合高程h_{fusion}的计算公式为:

$$h_{\text{fusion}}=w_1\times h_1+w_2\times h_2 \quad (12)$$

式中,h_1和h_2分别是InSAR DSM和光学DSM的高程值。将所有像元位置的高程值进行加权平均,生成最终的融合DSM。本文所提方法的主要流程见图2。

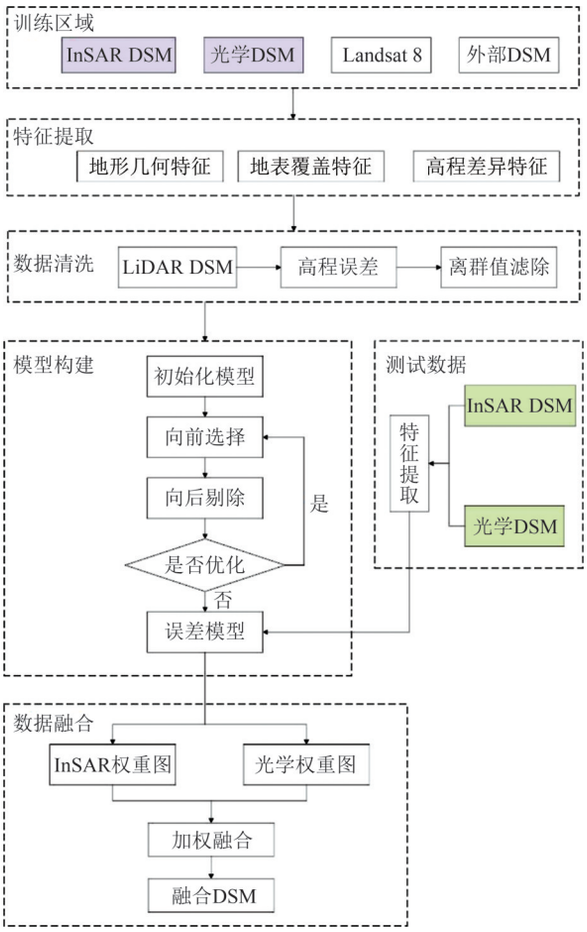


图2 顾及地形特征的InSAR和光学DSM逐步回归融合方法流程图

Fig. 2 Flowchart of Stepwise Regression Fusion Method of InSAR and Optical DSM Considering Terrain Features

2 DSM融合实验

2.1 研究区与实验数据

本文选取美国俄勒冈州喀斯喀特山脉某区域作为研究区,该区域位于北美洲太平洋沿岸山脉,地势起伏显著,涵盖高山、峡谷、森林和河流等多种地形类型,海拔范围从低海拔河谷到超过3 000 m的高山。植被以针叶林为主,覆盖密集。为了提取多样的地形特征,本文获取了多种来源的遥感数据,研究区地形与数据覆盖情况如图3所示。采用1景双星InSAR获取的TanDEM-X raw DSM与1景光学立体测量获取的AW3D30 DSM数据开展融合实验。其中,蓝框为TanDEM-X覆盖区域,黄框为AW3D30覆盖区域,紫框为LiDAR DSM覆盖区域。在重叠区选取训练数据和测试数据进行融合实验,其中,$A1$为训练区域,$B1$为以高山为主的测试区域1,$B2$为涵盖了山地和平地的测试区域2。Landsat 8影像用于评估研究区域的植被覆盖情况,哥白尼DSM作为外部参考DSM,LiDAR DSM用于精度评定。研究区实验数据如表1所示。

2.2 数据处理

2.2.1 预处理

在进行融合之前,对异源DSM数据进行预处理是确保融合精度的关键步骤。TanDEM-X raw DSM由TanDEM-X干涉影像对处理得到,存在较多粗差和空洞,因此,对其采取了粗差剔除和空洞填补[34]处理,以确保融合数据的完整性和可靠性。另外,不同来源的数据在分辨率、水平基准和垂直基准等方面存在显著差异,需要消除这些不一致性。采用双线性插值法对数据进行重采样[35],将分辨率统一为10 m,将所有数据转换到WGS84坐标系,使得InSAR DSM和光学DSM基于同样的参考系统。将InSAR DSM和光学DSM进行配准处理,以确保DSM在像素层面的一致性。

2.2.2 融合处理

在训练区域内提取地形特征作为回归模型的训练样本,在训练区提取的坡度、坡向、粗糙度和TPI等地形因子如图4所示,提取的NDVI如图5所示。地形因子由InSAR DSM和光学DSM数据分别处理得到,用于描述地形的几何特征,NDVI则通过Landsat 8影像计算得出。由图4可知,DSM的差异在坡度和粗糙度的表征上尤为明显。

此外,本文还提取了InSAR DSM和光学DSM的高程值、两者之间的高程差异、各自数据与参考DSM的残差和InSAR DSM的相干性高程误差,具体结果如图6所示。这些特征在一定程度上能够反映异源DSM数据在几何精度和可靠性方面的差异性,为后续融合建模奠定了数据基础。

2.3 精度评定

为了验证融合结果的效果,使用LiDAR DSM作为参考DSM进行精度评估。本文采用了两种主要的精度指标,包括反映整体误差的RMSE和反映抗粗差影响的NMAD,二者的计算公式分别为:

$$Z_{\mathrm{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\Delta h_i^2} \tag{13}$$

$$Z_{\mathrm{NMAD}}=\mathrm{median}\left(\left|\Delta h_i-m_{\Delta h}\right|\right)\times 1.482\,6 \tag{14}$$

式中,median()用于求取中位数;Δh_i代表DSM与参考DSM的高程误差;$m_{\Delta h}$代表误差中值。

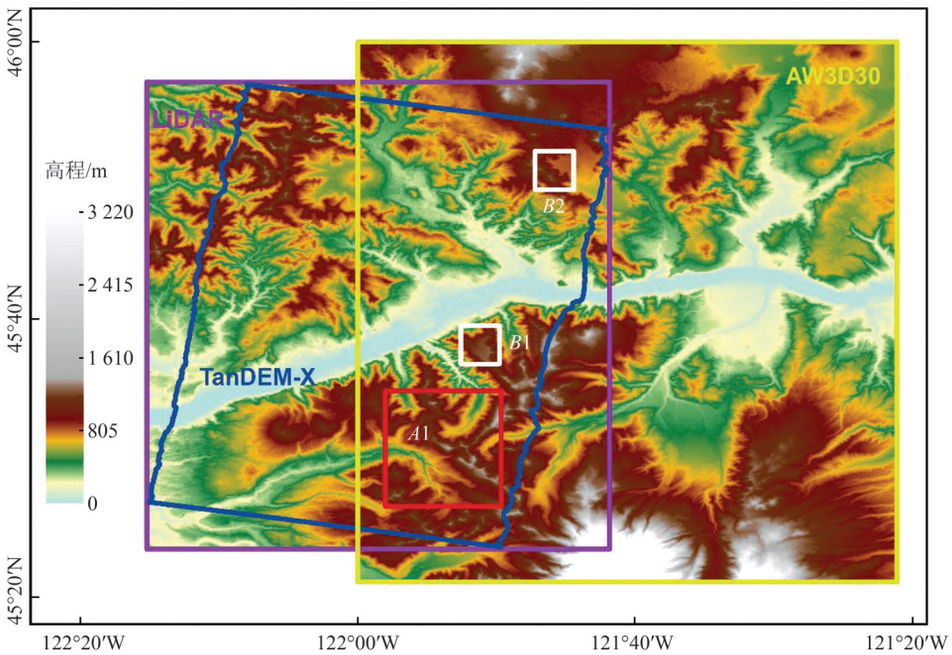


图 3　研究区地形概况与数据分布

Fig. 3　Topographic Overview and Data Distribution of the Study Area

表 1　研究区实验数据

Table 1　Experimental Data in the Study Area

数据	水平基准	垂直基准	分辨率/m	尺寸/像素
TanDEM-X	WGS84	EGM08	10	6 054×6 006
AW3D30		EGM96	30	7 000×7 000
LiDAR DSM		NAD83	10	6 054×6 006
哥白尼DSM		EGM08	30	10 800×20 600
训练区域*A*1的数据	WGS84	EGM08	10	1 500×1 500
测试区域*B*1的数据				500×500
测试区域*B*2的数据				
Landsat 8	UTM-WGS84		30	7 901×7 781

3　实验结果分析与讨论

3.1　融合结果

为了验证本文方法改善DSM质量的可行性和有效性,采用多种融合方法进行对比,包括简单平均法[17]、遍历权重寻优法[21]、随机森林[15]和回归分析方法,其中,回归方法分别采用了线性回归法、二次回归法和本文的逐步回归法来构建模型。在构建模型时,随机森林法和回归方法所使用的地形特征相同。本文所用对比方法的原理如表2所示。

测试区域*B*1、*B*2的融合结果分别如图7、图8所示。从视觉对比来看,本文逐步回归融合方法与其他融合方法以及原始的InSAR和光学DSM数据之间并未表现出明显的差异,这主要是因为实验区域的高程跨度较大,接近1 000 m,而融合前后DSM的垂直高程变化通常仅在数十米范围内。因此,融合前后的垂直高程变化在地形渲染图中难以直观体现。但从DSM的整体高程分布可以看出,本文方法生成的DSM能够有效维持地形的整体稳定性,未对宏观地形特征产生显著影响。

InSAR原始DSM在山体边缘区域存在明显的高程噪声,通过简单平均法和遍历权重寻优法融合后,这些异常现象仍然部分保留;而回归方法和随机森林等其他融合方法则能较好地消除这些噪声,实现更平滑的高程视觉效果。

3.2　误差分布定性分析

测试区域*B*1和*B*2的DSM误差分布情况分别如图9、图10所示。误差分布图由目标DSM与LiDAR DSM的高程进行差分计算得到,并对该差值进行分级分类处理,以颜色区分误差正负与量级。

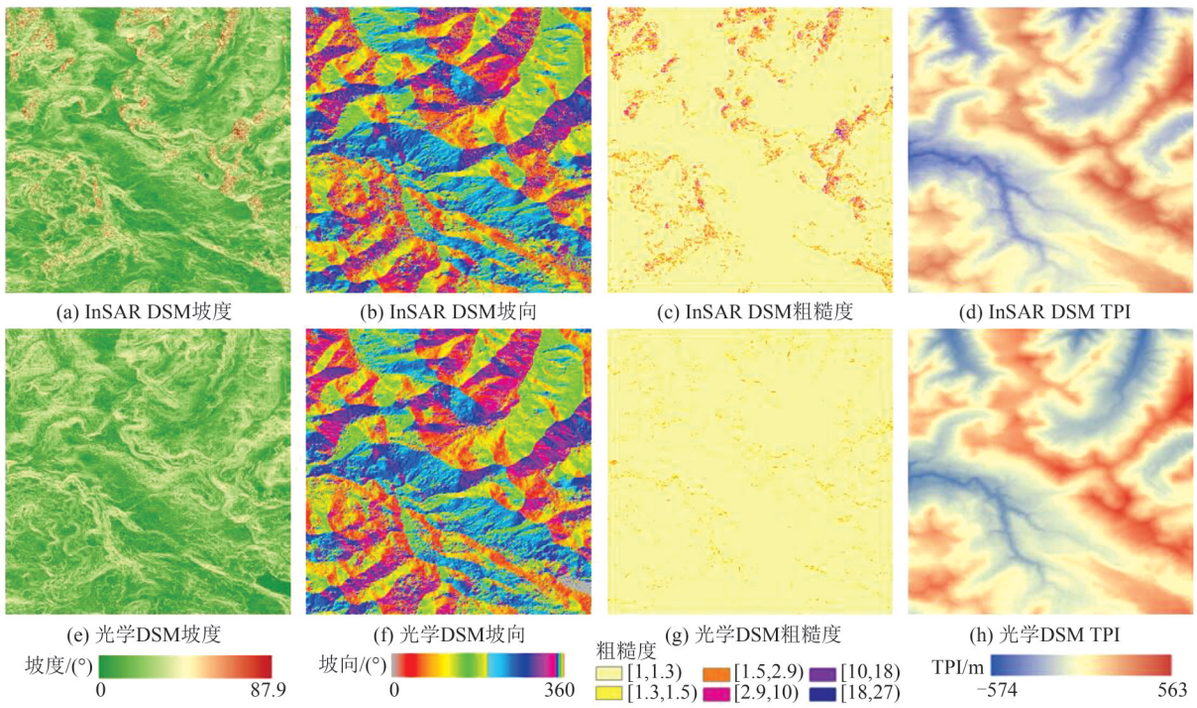


图4 DSM训练区域提取的地形因子

Fig. 4 Terrain Factors Extracted from the DSM Training Area

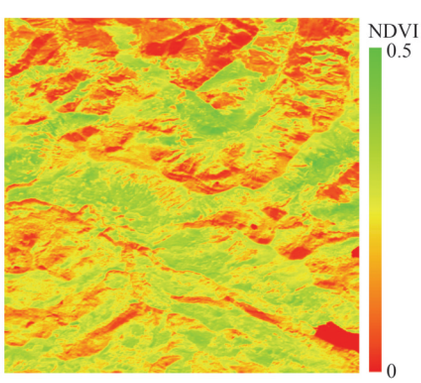


图5 DSM训练区域提取的NDVI

Fig. 5 NDVI Extracted from the DSM Training Area

从图9、图10中InSAR和光学DSM的误差分布来看,大部分区域的误差绝对值集中在20 m以内,部分区域误差超过50 m。对比图9、图10中各融合DSM与原始DSM的误差分布可知,所有融合方法均通过信息互补,在一定程度上降低了误差,提升了DSM的精度。具体而言,与融合前相比,等权的简单平均融合DSM和遍历权重寻优融合DSM的误差分布变化较小,其精度提升有限,随机森林融合DSM进一步降低了误差水平。在回归分析方法中,逐步回归方法表现最佳,线性回归方法次之,二次回归方法效果相对一般,但三者误差分布趋势相似,均显著改善了DSM的质量。值得注意的是,回归方法的误差分布与随机森林方法较为接近,且在某些区域误差更小,分布也更为均匀,优于随机森林方法。本文提出的顾及地形特征的逐步回归方法在两个测试区域均表现出较少的深色分布,表明其误差量级最小,更接近真实地形。

图11和图12分别为测试区域$B1$、$B2$融合前DSM与不同方法融合后DSM的误差直方图,展示了不同方法处理得到的DSM误差量级分布情况。由图11可知,融合后DSM的误差更集中于0值附近。相比于其他方法,本文方法融合后的DSM误差十分接近正态分布,证明了本文方法的有效性。由图12可知,原始DSM误差在该区域呈现出双峰分布,这是因为该区域为混合地形,主要包含了平缓的裸地和植被茂密的陡峭山地,这两种地形存在不同的误差特性。在平地部分,误差分布集中在0值周围,数据精度有较好的一致性;而在高山地森林区域,由于不同技术手段测高表面的差异性,所以误差较大,且分布相对分散。在这种情况下,用本文方法对DSM进行了加权平均融合处理后,误差分布在整体上更趋向于0,且相对集中。这一结果表明,本文方法不仅能够提升DSM的精度,且在多样地形条件下的表现具有一定的稳定性。

3.3 整体精度定量分析

为评估构建的误差模型在预测高程误差方面的性能,本文对随机森林、线性回归、二次回归和逐步回归4种方法训练得到的模型进行了精度

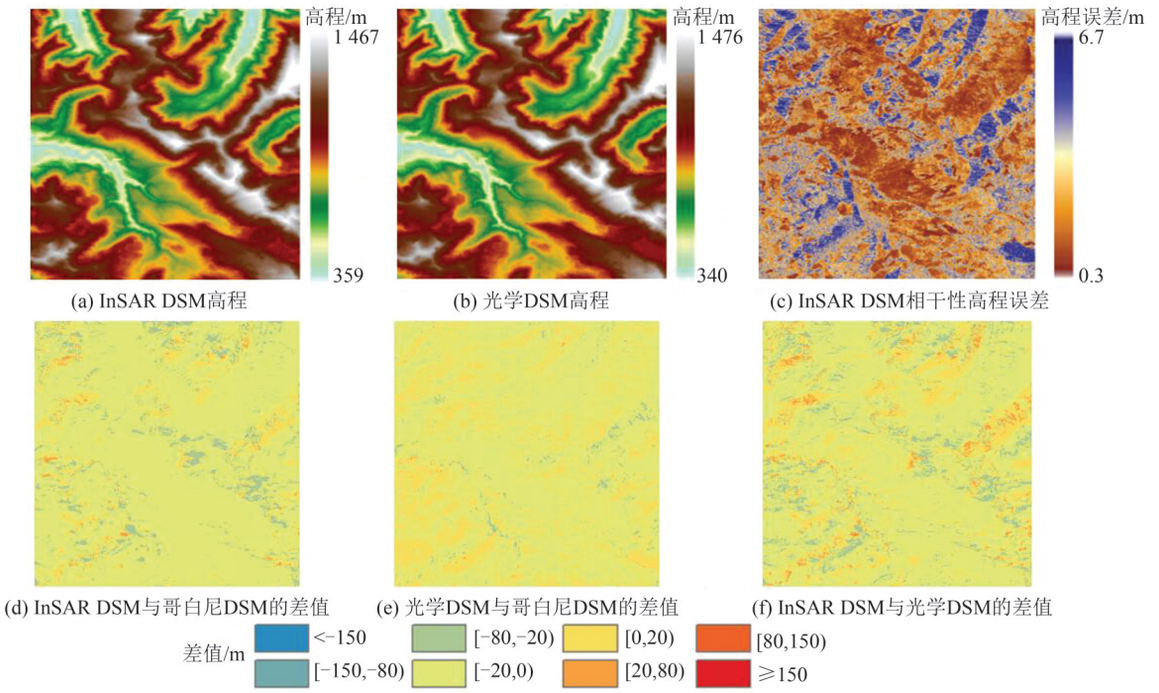


图 6　DSM 训练区域的高程信息与特异性误差

Fig. 6　Elevation Information and Specific Errors of the DSM Training Area

表 2　融合实验所用对比方法的原理

Table 2　Principle of Comparison Methods Used in Fusion Experiment

方法类型	具体方法	原理
传统加权平均融合	简单平均法	对多源DSM进行算术平均
	遍历权重寻优法	遍历不同权重组合,寻找最优的权重分配方案
机器学习	随机森林	利用随机森林方法构建误差模型,进行加权平均融合
回归分析	线性回归法	利用线性回归模型构建误差模型,进行加权平均融合
	二次回归法	利用二次回归模型构建误差模型,进行加权平均融合
	逐步回归法	利用逐步回归筛选特征变量以构建误差模型,进行加权平均融合

评定。首先,在两个测试数据集上,分别对InSAR DSM和光学DSM的误差预测结果进行了评估,将预测误差与真实误差作差;然后,采用差值的均值和RMSE作为评价指标来评估模型精度。这4种方法下误差模型的预测精度如表3所示,由表3可知,本文逐步回归模型在InSAR DSM误差的预测中表现最佳,误差均值和RMSE均优于其他模型。在光学DSM误差的预测中,逐步回归方法同样表现良好,在测试区域*B*2的预测效果最好。虽然在测试区域*B*1,逐步回归方法在光学DSM误差的预测精度略低于线性回归方法,但综合两个数据集的结果来看,逐步回归方法在预测高程误差方面的表现最为稳定和准确,能够为加权融合过程提供可靠的误差估计。

InSAR DSM、光学DSM以及6种融合方法生成的DSM在测试区域*B*1、*B*2的定量评价结果如表4所示。由表4可知,融合前的InSAR DSM在测试区域*B*1、*B*2的RMSE值和NMAD值均最大,表明其整体误差较大,尤其是在测试区域*B*1,表明其在复杂高山地形区域的精度较低。相比之下,光学DSM在测试区域*B*1、*B*2的RMSE值和NMAD值均低于InSAR DSM,表明其精度显著高于InSAR DSM。

融合后DSM的精度均有所提升。将对应方法融合的DSM与输入的DSM中精度最高的数据进行对比,结果表明,简单平均融合方法在测试区域*B*1、*B*2中的数据精度均有一定改善,说明它在一定程度上能够提升精度,但效果有限。遍历权重寻优方法融合得到的DSM的RMSE优于简单平均方法。相较于简单平均、遍历权重寻优方法,随机森林在两个测试区域有更好的提升效果,但其对地形的泛化能力不够稳定。线性回归方法在测试区域*B*1的表现优于随机森林方法,但在测试区域*B*2的NMAD表现略逊于随机森林方法。二次回归融合方法在测试区域*B*2的表现优于线性回归融合方法,但在测试区域*B*1的精度略

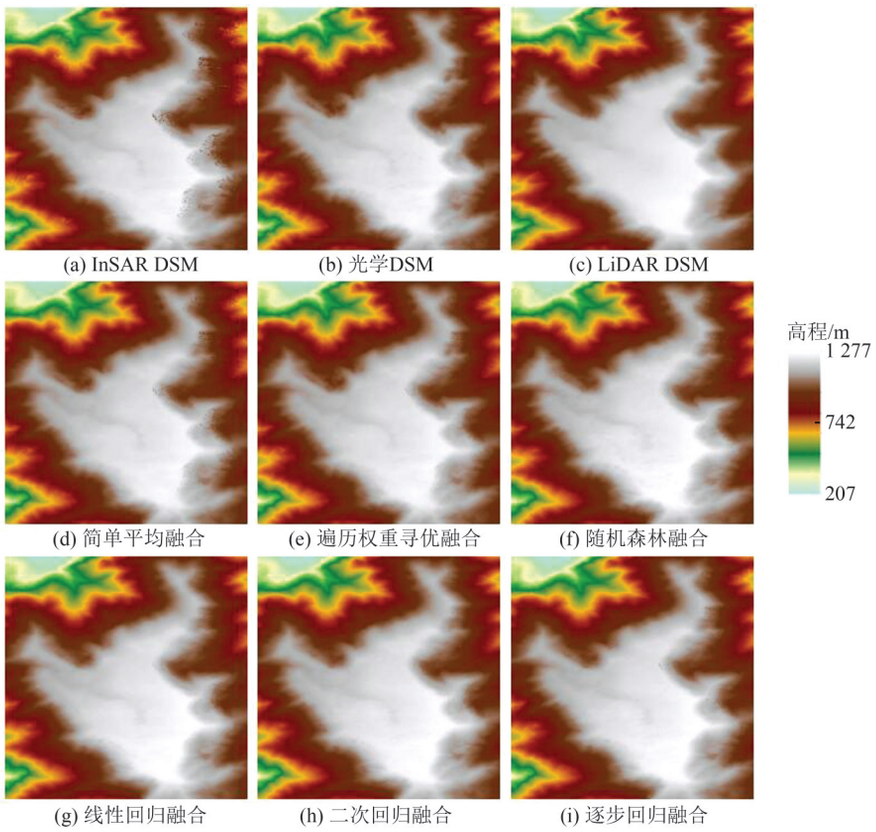


图7 测试区域*B*1的DSM及融合结果

Fig. 7 DSM and Fusion Results of Test Area *B*1

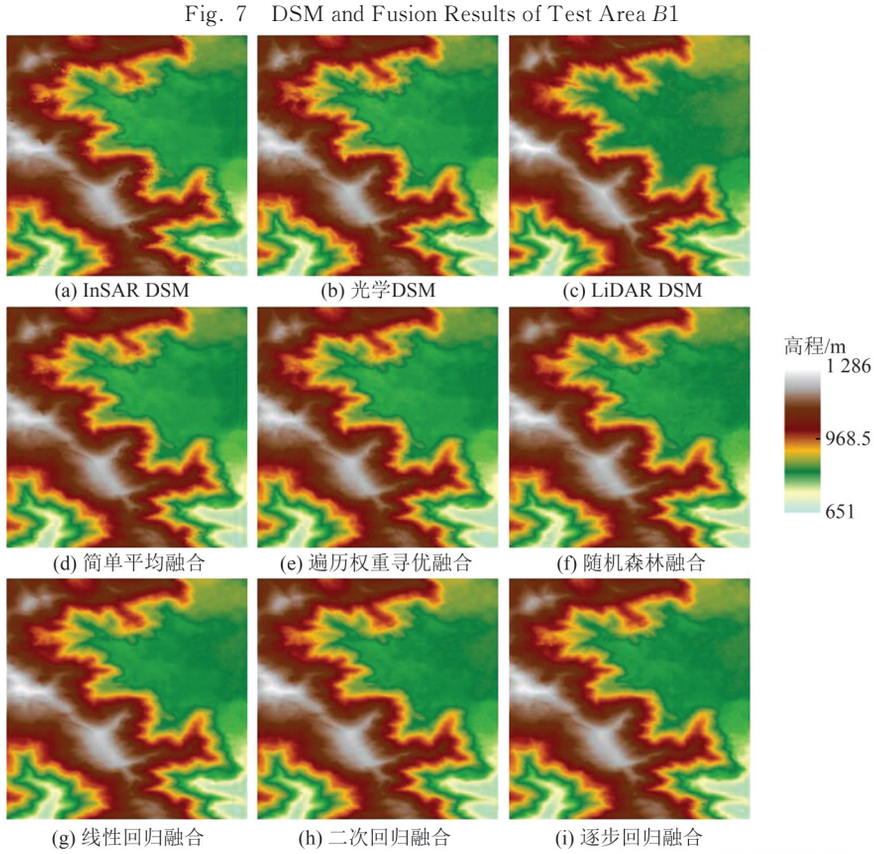


图8 测试区域*B*2的DSM及融合结果

Fig. 8 DSM and Fusion Results of Test Area *B*2

低于线性回归融合方法。逐步回归融合方法在所有融合方法中表现最佳。与融合前精度最高的光学DSM相比,逐步回归融合DSM在测试区域*B*1的RMSE和NMAD值分别降低了14.3%

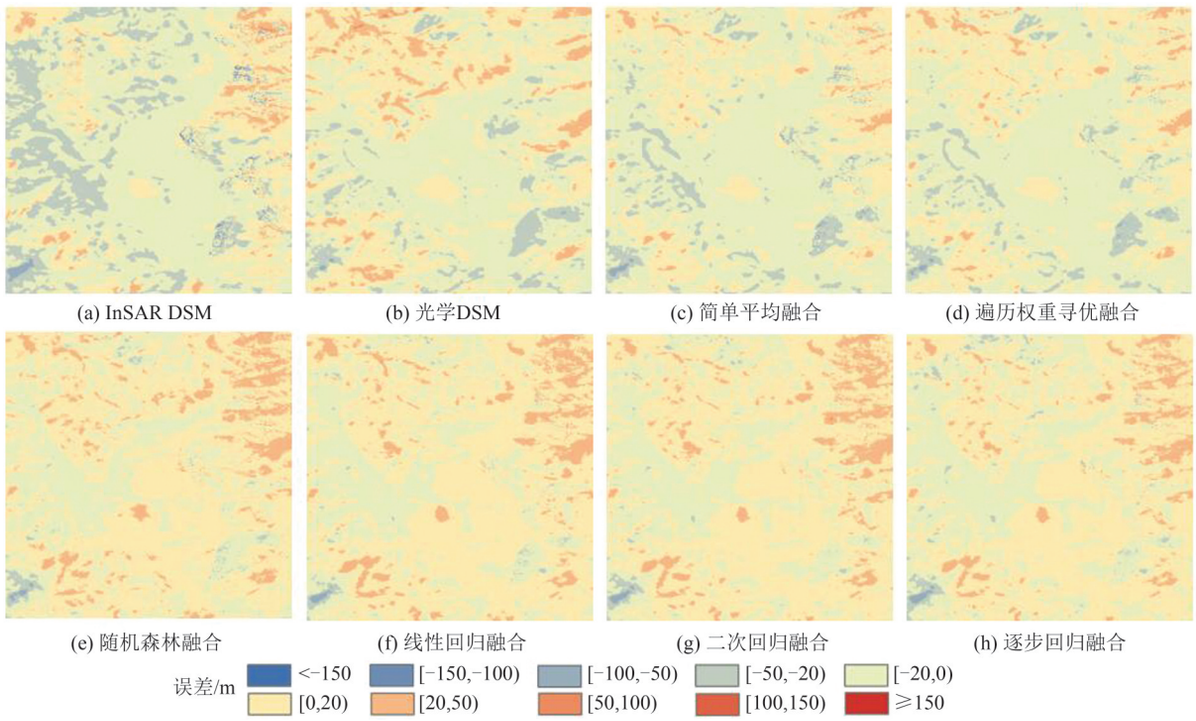


图 9　测试区域 *B*1 的 DSM 误差分布

Fig. 9　Distribution of DSM Error in Test Area *B*1

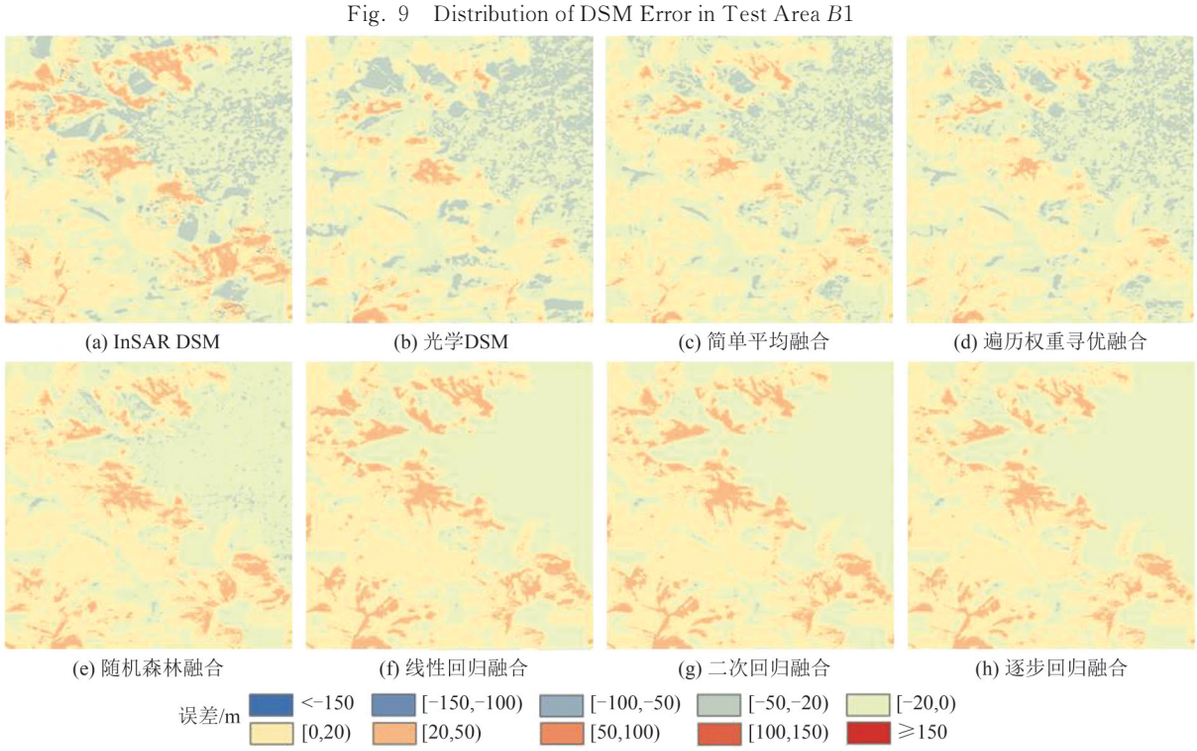


图 10　测试区域 *B*2 的 DSM 误差分布

Fig. 10　Distribution of DSM Error in Test Area *B*2

和 20%,在测试区域 *B*2 的 RMSE 和 NMAD 值分别降低了 18% 和 10.3%。总体而言,逐步回归方法能够有效提取多源 DSM 的互补信息,尤其是在地形复杂区域,表现出较强的精度提升能力。随机森林和另外两种回归方法在部分区域表现良好,但整体精度仍不及逐步回归方法。

与机器学习方法的黑盒特性不同,逐步回归方法建立的是一种透明且可解释的统计模型,其建模过程和结果都可以通过明确的统计标准进行解释,适用于需要高可解释性的应用场景。然而,其局限性也需要注意,特别是在处理非线性问题或高维数据时,可能需要结合其他方法进行优化。

实验验证结果表明,本文提出的基于逐步回归方法的 InSAR 和光学 DSM 融合策略能够有效

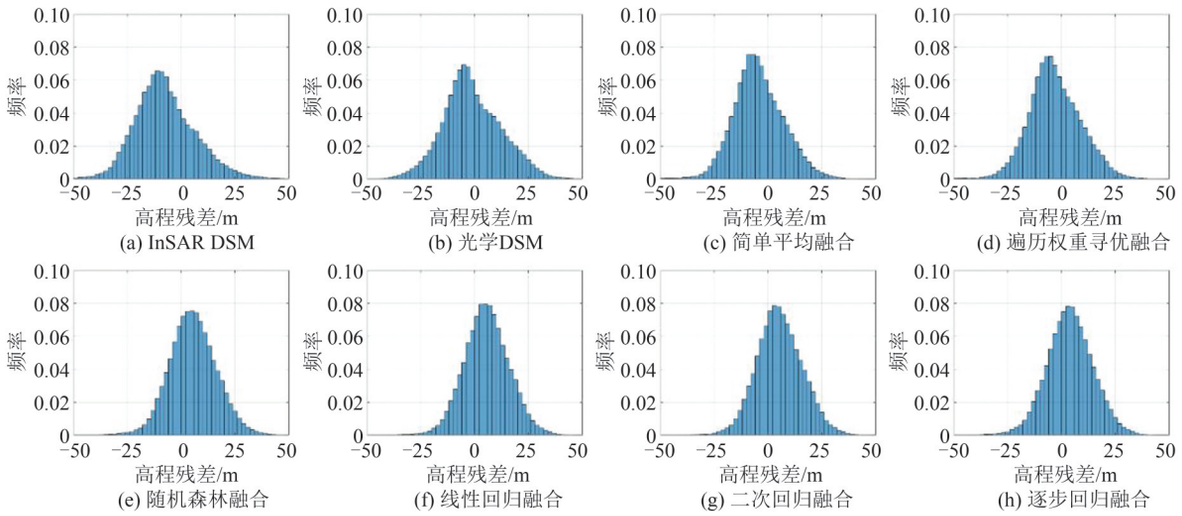


图11 测试区域*B*1的DSM误差直方图

Fig. 11 DSM Error Histogram of Test Area *B*1

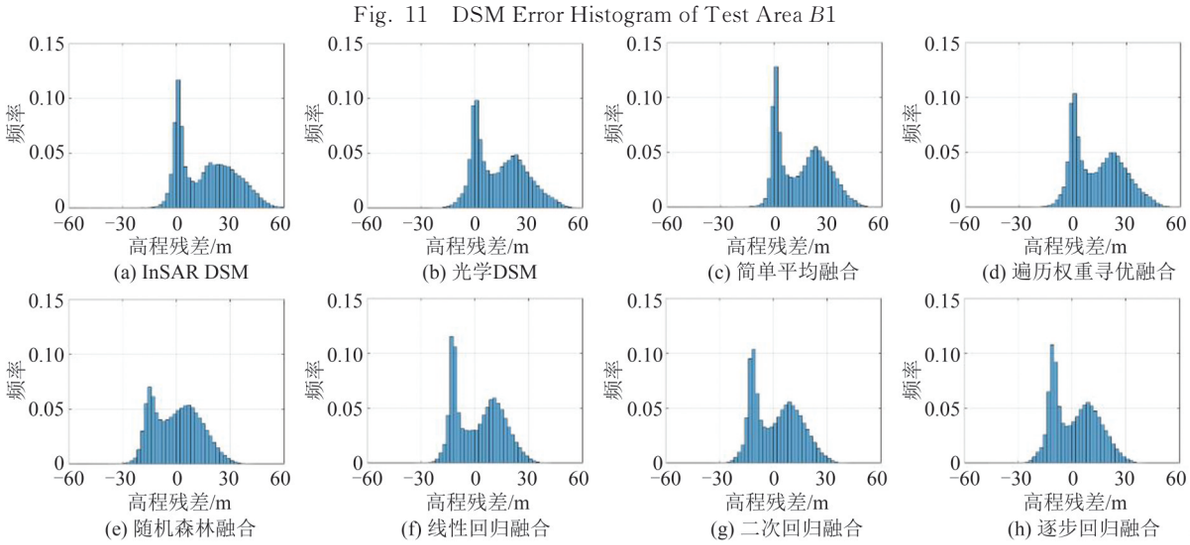


图12 测试区域*B*2的DSM误差直方图

Fig. 12 DSM Error Histogram of Test Area *B*2

表3 不同误差模型的预测精度/m

Table 3 Prediction Accuracy of Different Error Models/m

方法	测试区域*B*1				测试区域*B*2			
	InSAR		光学		InSAR		光学	
	均值	RMSE	均值	RMSE	均值	RMSE	均值	RMSE
随机森林	9.66	13.17	9.97	12.74	10.96	13.10	11.75	13.61
线性回归	9.43	12.05	9.63	12.34	11.41	13.07	11.71	13.36
二次回归	9.54	12.23	9.79	12.65	10.88	12.72	11.40	13.22
逐步回归	9.05	11.79	9.76	12.78	10.46	12.24	10.83	12.66

降低误差,提升DSM的精度。与简单平均方法相比,逐步回归方法能够更好地利用地形特征和数据质量信息,从而实现更精准的误差估计和加权平均融合。融合后的DSM在低山区域和高山区域均表现出较高的精度,误差量级显著降低。

4 结 语

本文提出了一种顾及地形特征的加权平均融合方法用于InSAR和光学DSM的融合。该方法通过构建误差模型,并利用地形特征和数据质量信息来估计误差,进而实现加权平均融合。为验证本文方法的有效性,开展了InSAR与光学DSM数据融合实验。结果表明,本文方法能够显著提升DSM的精度,为提高融合DSM的数据质量提供了一种新的思路和解决方案。与其他方法相比,本文方法在精度、稳定性和可解释性方面均展现出显著优势,在一定程度上揭示了地形

表4　DSM精度评定/m

Table 4　DSM Accuracy Evaluation/m

融合前后	DSM和融合DSM	测试区域$B1$		测试区域$B2$	
		RMSE	NMAD	RMSE	NMAD
融合前	InSAR DSM	18.09	13.02	16.74	20.09
	光学DSM	13.82	12.94	14.94	17.14
融合后	简单平均	13.17	11.08	14.37	18.68
	遍历权重寻优	12.65	11.40	14.21	18.10
	随机森林	12.45	10.50	13.58	17.75
	线性回归	12.12	10.11	13.14	17.93
	二次回归	12.17	10.35	12.78	16.63
	逐步回归	11.85	10.35	12.25	15.37

特征与DSM误差的非线性关系。本文提出的基于误差模型定权的加权融合方法，弥补了传统加权平均融合方法中权重依赖外部先验信息的不足。这种方法不仅提高了融合后DSM的精度，还增强了融合的适应性和可靠性。未来的研究可以进一步优化误差模型，结合更多的地形特征和数据源，进一步提升融合DSM的精度和可靠性。此外，还可以探索本文方法在其他复杂地形条件下的适用性，进一步拓展其应用范围。

参考文献

[1] 雷秋佳，刘婧，曹新运. 利用机载LiDAR数据的开放DEM产品精度评估[J]. 武汉大学学报(信息科学版)，2025，50(1)：153-163.
LEI Qiujia，LIU Jing，CAO Xinyun. Accuracy Evaluation of Open DEM Products Based on Airborne LiDAR Data[J]. *Geomatics and Information Science of Wuhan University*，2025，50(1)：153-163.

[2] 张文婷. 基于InSAR技术的DEM重建理论与应用研究[D]. 西安：长安大学，2018.
ZHANG Wenting. Theory and Application of High Accuracy DEM Constructed by InSAR Technology [D]. Xi'an：Chang'an University，2018.

[3] 李振洪，李鹏，丁咚，等. 全球高分辨率数字高程模型研究进展与展望[J]. 武汉大学学报(信息科学版)，2018，43(12)：1927-1942.
LI Zhenhong，LI Peng，DING Dong，et al. Research Progress of Global High Resolution Digital Elevation Models [J]. *Geomatics and Information Science of Wuhan University*，2018，43(12)：1927-1942.

[4] YUE L W，SHEN H F，ZHANG L P，et al. High-Quality Seamless DEM Generation Blending SRTM-1，ASTER GDEM V2 and ICESat/GLAS Observations[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2017，123：20-34.

[5] TOUTIN T. Impact of Terrain Slope and Aspect on Radargrammetric DEM Accuracy [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2002，57(3)：228-240.

[6] 杜亚男，冯光财，李志伟，等. TerraSAR-X/TanDEM-X获取高精度数字高程模型技术研究[J]. 地球物理学报，2015，58(9)：3089-3102.
DU Yanan，FENG Guangcai，LI Zhiwei，et al. Generation of High Precision DEM from TerraSAR-X/TanDEM-X[J]. *Chinese Journal of Geophysics*，2015，58(9)：3089-3102.

[7] 刘平芝，孙亮，张卫柱，等. 多源DEM数据融合[J]. 测绘科学与工程，2014(3)：14-20.
LIU Pingzhi，SUN Liang，ZHANG Weizhu，et al. Multi-source DEM Data Fusion[J]. *Geomatics Science and Engineering*，2014(3)：14-20.

[8] OKOLIE C J，SMIT J L. A Systematic Review and Meta-Analysis of Digital Elevation Model (DEM) Fusion：Pre-processing，Methods and Applications [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2022，188：1-29.

[9] BAGHERI H，SCHMITT M，ZHU X X. Fusion of TanDEM-X and Cartosat-1 Elevation Data Supported by Neural Network-Predicted Weight Maps [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*，2018，144：285-297.

[10] GIROHI P，BHARDWAJ A. ANN-Based DEM Fusion and DEM Improvement Frameworks in Regions of Assam and Meghalaya Using Remote Sensing Datasets[J]. *European Journal of Environment and Earth Sciences*，2022，3(4)：79-89.

[11] GUI R，QIN Y J，HU Z，et al. Neural Network-Based Fusion of InSAR and Optical Digital Elevation Models with Consideration of Local Terrain Features[J]. *Remote Sensing*，2024，16(19)：3567.

[12] 岳林蔚，沈焕锋，袁强强，等. 基于深度置信网络的多源DEM点面融合模型[J]. 武汉大学学报(信息科学版)，2021，46(7)：1090-1097.
YUE Linwei，SHEN Huanfeng，YUAN Qiangqiang，et al. A Multi-source DEM Point-Surface Fusion Model Based on Deep Belief Network[J]. *Geomatics and Information Science of Wuhan University*，2021，46(7)：1090-1097.

[13] 黄帅元，董有福，李海鹏. 黄土高原区SRTM1 DEM高程误差校正模型构建及对比分析[J]. 地球信息科学学报，2023，25(3)：669-681.
HUANG Shuaiyuan，DONG Youfu，LI Haipeng. Establishment and Comparative Analysis of SRTM1 DEM Error Correction Model in the Loess Plateau[J]. *Journal of Geo-Information Science*，2023，25(3)：669-681.

[14] HOJA D，D'ANGELO P. Analysis of DEM Combination Methods Using High Resolution Optical Stereo Imagery and Interferometric SAR Data[C]//ISPRS Commission I Symposium "From Sensors to Imagery"，Marne-la-Vallée，France，2006.

[15] LEITÃO J P, DE SOUSA L M. Towards the Optimal Fusion of High-Resolution Digital Elevation Models for Detailed Urban Flood Assessment[J]. *Journal of Hydrology*, 2018, 561: 651-661.

[16] KAUR H, KOUNDAL D, KADYAN V. Image Fusion Techniques: A Survey[J]. *Archives of Computational Methods in Engineering*, 2021, 28(7): 4425-4447.

[17] CHU T, DAS A, LINDENSCHMIDT K E. RADARSAT-2-Based Digital Elevation Models Derived from InSAR for High Latitudes of Northern Canada[J]. *Journal of Applied Remote Sensing*, 2017, 11(3): 035013.

[18] 刘娇, 赵尚民. 结合NASA DEM和AW3D30 DEM的太原市DEM数据融合[J]. 测绘通报, 2022(11): 90-95.
LIU Jiao, ZHAO Shangmin. DEM Fusion Based on NASA DEM and AW3D30 DEM in Taiyuan[J]. *Bulletin of Surveying and Mapping*, 2022(11): 90-95.

[19] ZHAO S M, LIU J, CHENG W M, et al. Fusion Scheme and Implementation Based on SRTM1, ASTER GDEM V3, and AW3D30[J]. *ISPRS International Journal of Geo-Information*, 2022, 11(3): 207.

[20] SCHINDLER K, PAPASAIKA-HANUSCH H, SCHÜTZ S A, et al. Improving Wide-Area DEMs Through Data Fusion: Chances and Limits[C]// The 53rd Photogrammetric Week, Stuttgart, Germany, 2011.

[21] DEO R, ROSSI C, EINEDER M, et al. Framework for Fusion of Ascending and Descending Pass TanDEM-X Raw DEMs[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(7): 3347-3355.

[22] GRUBER A, WESSEL B, MARTONE M, et al. The TanDEM-X DEM Mosaicking: Fusion of Multiple Acquisitions Using InSAR Quality Parameters[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(3): 1047-1057.

[23] GROHMANN C H. Evaluation of TanDEM-X DEMs on Selected Brazilian Sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30[J]. *Remote Sensing of Environment*, 2018, 212: 121-133.

[24] HIRT C, FILMER M S, FEATHERSTONE W E. Comparison and Validation of the Recent Freely Available ASTER-GDEM Ver1, SRTM Ver4.1 and GEODATA DEM-9S Ver3 Digital Elevation Models over Australia[J]. *Australian Journal of Earth Sciences*, 2010, 57(3): 337-347.

[25] BHANG K J, SCHWARTZ F W, BRAUN A. Verification of the Vertical Error in C-Band SRTM DEM Using ICESat and Landsat-7, Otter Tail County, MN[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(1): 36-44.

[26] 蒋广鑫, 谢元礼, 高志远, 等. 中国5类典型区域常用DEM高程精度评价[J]. 水土保持研究, 2020, 27(5): 72-80.
JIANG Guangxin, XIE Yuanli, GAO Zhiyuan, et al. Evaluation on Elevation Accuracy of Commonly Used DEM in Five Typical Areas of China[J]. *Research of Soil and Water Conservation*, 2020, 27(5): 72-80.

[27] TRAN T A, RAGHAVAN V, MASUMOTO S, et al. A Geomorphology-Based Approach for Digital Elevation Model Fusion: Case Study in Danang City, Vietnam[J]. *Earth Surface Dynamics*, 2014, 2(2): 403-417.

[28] FU C Y, TSAY J R. Statistic Tests Aided Multi-source Dem Fusion[J]. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, XLI-B6: 227-233.

[29] BAGHERI H, SCHMITT M, ZHU X X. Fusion of Urban TanDEM-X Raw DEMs Using Variational Models[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(12): 4761-4774.

[30] 黎伟, 焦健, 韩海姣, 等. 双站与重轨InSAR数据融合获取DEM的方法[J]. 遥感信息, 2023, 38(2): 10-17.
LI Wei, JIAO Jian, HAN Haijiao, et al. Method of Generating DEM Through Fusion of Bistatic and Repeat-Pass InSAR Data[J]. *Remote Sensing Information*, 2023, 38(2): 10-17.

[31] KRIEGER G, MOREIRA A, FIEDLER H, et al. TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(11): 3317-3341.

[32] HÖHLE J, HÖHLE M. Accuracy Assessment of Digital Elevation Models by Means of Robust Statistical Methods[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2009, 64(4): 398-406.

[33] 游士兵, 严研. 逐步回归分析法及其应用[J]. 统计与决策, 2017, 33(14): 31-35.
YOU Shibing, YAN Yan. Stepwise Regression Analysis and Its Application[J]. *Statistics & Decision*, 2017, 33(14): 31-35.

[34] HU Z, GUI R, HU J, et al. InSAR Digital Elevation Model Void-Filling Method Based on Incorporating Elevation Outlier Detection[J]. *Remote Sensing*, 2024, 16(8): 1452.

[35] SHI W Z, WANG B, TIAN Y. Accuracy Analysis of Digital Elevation Model Relating to Spatial Resolution and Terrain Slope by Bilinear Interpolation[J]. *Mathematical Geosciences*, 2014, 46(4): 445-481.