第 XX 卷第 XX 期
XXXX 年 XX 月

红 外 与 毫 米 波 学 报
J. Infrared Millim. Waves

Vol. XX, No. XX
XX,XXXX

文章编号：1001-9014(XXXX)XX-0001-14

DOI:10. 11972/j. issn. 1001-9014. XXXX. XX. 001

Understory terrain estimation using multi-source remote sensing data under different forest-type conditions

HUANG Jia-Peng*，FAN Qing-Nan，ZHANG Yue

（School of Geomatics，Liaoning Technical University，Fuxin 123000，China）

Abstract：Accurate estimation of understory terrain has significant scientific importance for maintaining ecosystem balance and biodiversity conservation. Addressing the issue of inadequate representation of spatial heterogeneity when traditional forest topographic inversion methods consider the entire forest as the inversion unit, this study proposes a differentiated modeling approach to forest types based on refined land cover classification. Taking Puerto Rico and Maryland as study areas, a multi-dimensional feature system is constructed by integrating multi-source remote sensing data：ICESat-2 spaceborne LiDAR is used to obtain benchmark values for understory terrain, topographic factors such as slope and aspect are extracted based on SRTM data, and vegetation cover characteristics are analyzed using Landsat-8 multispectral imagery. This study incorporates forest type as a classification modeling condition and applies the random forest algorithm to build differentiated topographic inversion models. Experimental results indicate that, compared to traditional whole-area modeling methods (RMSE=5. 06 m), forest type-based classification modeling significantly improves the accuracy of understory terrain estimation (RMSE=2. 94 m), validating the effectiveness of spatial heterogeneity modeling. Further sensitivity analysis reveals that canopy structure parameters (with RMSE variation reaching 4. 11 m) exert a stronger regulatory effect on estimation accuracy compared to forest cover, providing important theoretical support for optimizing remote sensing models of forest topography.

Key words：understory terrain，forest type，multi-source remote sensing data，random forest model

针对不同森林类型条件下协同多源遥感数据估测林下地形

黄佳鹏*，樊庆南，张　玥

（辽宁工程技术大学 测绘与地理科学学院，辽宁 阜新 123000）

摘要：精确估算林下地形对维持生态系统平衡和生物多样性保护具有重要科学意义。针对传统森林地形反演方法以整体森林作为反演单元时存在的空间异质性表征不足问题，本研究提出基于精细化土地覆盖分类的森林类型差异化建模方法。以波多黎各地区和马里兰州为研究区，通过整合多源遥感数据构建多维特征体系：采用ICESat-2星载激光雷达获取林下地形基准值；基于SRTM数据提取坡度和坡向等地形因子；结合Landsat-8多光谱影像解析植被覆盖特征。本研究将森林类型作为分类建模条件，运用随机森林算法构建差异化地形反演模型。实验结果表明，相较于传统全域建模方法(RMSE=5.06 m)，本研究提出的森林类型分类建模使林下地形估算精度显著提升(RMSE=2.94 m)，验证了空间异质性建模的有效性。进一步敏感性分析发现，冠层结构参数(RMSE变异达4.11 m)相较于森林覆盖度对估算精度具有更强的调控作用，这为优化森林地形遥感模型提供了重要理论依据。

关　键　词：林下地形；森林类型；多源遥感数据；随机森林模型

中图分类号：S757　　　　**文献标识码**：A

Received date：2025-02-08，**revised date**：2025-02-20　　**收稿日期**：2025-02-08，**修回日期**：2025-02-20

Foundation items：Supported by the National Natural Science Foundation of China (42401488 and 42071351), the National Key Research and Development Program of China (2020YFA0608501 and 2017YFB0504204), the Liaoning Revitalization Talents Program (XLYC1802027), the Talent Recruited Program of the Chinese Academy of Science (Y938091), the Project Supported Discipline Innovation Team of the Liaoning Technical University (LNTU20TD-23), the Liaoning Province Doctoral Research Initiation Fund Program (2023-BS-202), and Basic Research Projects of Liaoning Department of Education (JYTQN2023202)

Biography：HUANG Jia-Peng (1993-), male, Heilongjiang, doctorate. Research area involves 3-D positioning of quantitative remote sensing. E-mail：huangjp@nefu. edu. cn

* **Corresponding author**：E-mail：huangjp@nefu. edu. cn

Introduction

As the largest terrestrial ecosystem on Earth, forests are the birthplace and habitat for the reproduction and development of many animals and plants, and thus occupy an irreplaceable position in the ecosystem. Therefore, the modeling of forest ecosystems and the investigation of important resources are urgently needed, but they represent challenges in the current research[1]. Understory terrain is an important basis for estimating forest structural parameters, and its impact on forest ecosystems is extremely multifaceted and far-reaching. Understory terrain can have a significant effect on water circulation, soil conservation, climate environment, and forest structure in forests[2]. Therefore, how to obtain understory terrain rapidly and accurately has become a challenge.

Measured data can be used for understory terrain measurement. However, in the understory environment, where vegetation is lush, the usage of measuring instruments is greatly restricted in aspects of placement and normal operation[3]. For instance, for complex understory terrains, such as areas with dense vegetation, it is difficult for surveyors to reach all the locations to be measured, which can easily cause blank areas in data collection and thus affect the reliability of terrain estimation. At the same time, measured data can be obtained only at limited sampling points, making it difficult to estimate understory terrain from these data. However, remote sensing data can also be used to estimate understory terrain. These data can cover a large area, which is highly beneficial for understory terrain estimation. Compared with traditional ground-based measurement methods, remote sensing can rapidly obtain information over a large range, saving much manpower, time cost, and many material resources[4]. Nevertheless, using only a single type of remote sensing data poses certain limitations in reflecting the complex characteristics of understory terrain. These limitations can be overcome using collaborative multi-source remote sensing data, which provides the possibility for a more accurate estimation of understory terrain.

Currently, optical remote sensing technology plays an important role in forest surveys and forest parameter estimation. The Landsat-8 data have certain advantages in estimating understory terrain; it can provide multi-spectral information, including band data such as visible light and infrared data[5]. The texture and other characteristics of vegetation shown in an image are related to the topography to a certain extent. However, optical remote sensing is severely affected by factors such as forest shielding and weather, which further affect the estimation result of forest structure parameters. Optical remote sensing can perform well in estimating horizontal structure parameters, but its ability to estimate forest vertical structure parameters is relatively weak. Therefore, understory terrain estimation cannot completely rely on optical remote sensing data. However, the synthetic aperture radar (SAR) data can be used for understory terrain estimation because SAR uses microwave frequency band detection, and microwaves have certain penetration characteristics. Although microwaves are affected by vegetation scattering and absorption in the forest, they can still penetrate the vegetation layer, reach the understory ground, and reflect. In addition, microwaves with different wavelengths have different penetration capabilities[6]. Liu et al. [7] proposed an understory terrain estimation method based on dual-polarization PolInSAR data. They concluded that the accuracy of dual-polarization PolInSAR in estimating understory terrain was close to that of full-polarization PolInSAR and was better than the estimation accuracy of traditional InSAR technology; however, its penetration ability was limited. Namely, the complex branch structure and dense leaf layer of trees in the forest prevent the SAR signals from penetrating the understory terrain [8]. Therefore, the effect of SAR data on understory terrain estimation is limited.

In vegetation-covered areas, airborne light detection and ranging (LiDAR) can penetrate the vegetation canopy and obtain topographic information under the vegetation, ensuring high-precision understory terrain estimation[9]. However, although airborne LiDAR has a certain ability to penetrate vegetation, it is suitable only for estimation on a small regional scale and is slightly insufficient for large areas.

Spaceborne LiDAR has good accuracy in estimating understory terrain. Namely, this technology enables satellites to obtain multiple observation footprints at the same time and penetrate vegetation more effectively for surface measurement. The ICESat-2 has a horizontal accuracy of approximately 6.5 m and a nominal vertical accuracy of about 0.1 m, providing reliable data on forest areas under various terrain conditions[10]. In Ref. [11], the vertical accuracy of ASTER, SRTM, GLO-30, and ATLAS was evaluated using the digital terrain model (DTM) of the reference data set provided by the G-LiHT. The results showed that in the forest environment, the ICESat-2's ATL03 had the highest accuracy on the footprint scale, with a correlation coefficient (R^2) close to 1 and a root mean square error (RMSE) of 1.96 m. In Ref. [12], the DTM of the airborne G-LiHT system was used to evaluate the ICESat-2/ATLAS data, AW3D30 DEM data, and TanDEM-X data. The results showed that at the footprint scale, the ATLAS data provided more accurate evaluation indicators than the AW3D30 DEM and TanDEM-X data products. Although spaceborne LiDAR can penetrate the canopy and obtain understory terrain data, the ICESat-2 data start with spot-like data and can be discretely sampled only in space but can neither obtain planar understory terrain information nor fully cover the study area. Therefore, how to use spaceborne LiDAR data to estimate high-precision planar understory terrain has become an important problem. Zhang et al. [13] analyzed the differences in estimation accuracy between ground elevation and vegetation canopy height under different beam intensities, times, slopes, and vegetation coverage using high-precision airborne LiDAR data and plot measured data as references. The result indicated that the root mean square error (RMSE) of the estimation accuracy of the strong beam of ATL08 was

1.9 m, and the mean absolute error (MAE) was 1.1 m. Zhang et al. [14] developed a method for generating understory terrain using the ICESat-2/ATLAS data and TanDEM-X DEM. The experimental results showed that the accuracy of inverted understory terrain was 9.14 m, which was 21.2% higher than the accuracy of the original Tandem-X DEM. In Ref. [15], a mathematical modeling method based on the ICESat-2 data was proposed to interpolate and correct the SRTM DEM of understory terrain. This method could significantly improve the accuracy of SRTM DEM in understory terrain and reduce the root mean square error from 13.40 m to 6.74 m. Although this method could effectively improve the estimation accuracy of understory terrain, relying only on mathematical principles was not sufficient to adapt to different forest environments. Namely, using only mathematical methods cannot ensure high estimation accuracy of understory terrain for different forest types.

Aiming to overcome the aforementioned limitations, this study uses the understory terrain data extracted from the ICESat-2's ATL03 data for modeling and combines the ATL03 understory terrain data of ICESat-2 with the Landsat-8 OLI remote sensing data and SRTM topographic factors. In addition, multiple regression models and random forest models are employed to construct accurate understory terrain elevation data models. Moreover, the land cover type data are used to obtain data on different forest types and establish single-forest-type understory terrain estimation models to judge the accuracy differences between different forest types when estimating understory terrain.

The main innovations of this study are as follows:

(1) This study uses the forest type as a distinguishing condition and the ICESat-2/ATLAS ATL03 product as a research object. Also, it combines the slope aspect parameters and relevant parameter data extracted from the Landsat-8 optical remote sensing images to construct an understory terrain estimation model for the estimation of different forest types in the study area, which compensates for the deficiency of the ICESat-2 data that cannot achieve accurate understory terrain estimation at the wall-to-wall scale and realizes a high-precision understory terrain elevation model in the study area.

(2) The impacts of forest canopy height and forest coverage on understory terrain estimation are evaluated.

1 Study material

1.1 Study area

The research area in Puerto Rico has a latitude between 17.97° N and 18.34° N and a longitude between 66.31° W and 66.83° W and experiences a tropical marine climate. The research area's topography includes hills and mountains, having mainly sloping understory terrain. The mountain slopes are relatively steep, and the plains are nearly flat [16]. The slopes of forest areas are mainly concentrated between 8° and 45°, and the elevation ranges from 106 m to 744 m. The vegetation coverage ranges from 25% to 66%, and the main vegetation types include tropical broad-leaved forests and shrubs. The main forest types are closed evergreen broad-leaved forests and closed mixed forests, and the tree species mostly include *T*abebuia heterophylla, palm trees, and kapok trees; the canopy height ranges from zero to 34 m.

The research area in Maryland has a latitude between 37°53′N and 39°43′N and a longitude between 75°03′W and 79°31′W. It is located on the east coast of the United States, bordering the Atlantic Ocean, and has a unique geographical location. In the western and northern parts of the research area, there is a part of the Appalachian Mountains, presenting undulating hills and mountainous terrain. The terrain undulates severely; the mountains are steep, and the slopes reach an inclination of 30°-60° or more. The mountainous areas are relatively high in elevation and have complex terrain with many steep slopes and valleys. The Appalachian Mountains in the west have relatively high elevations, with the highest point exceeding 1,000 m. In contrast, the Atlantic Coastal Plain in the east has a lower elevation, close to the sea level; in some places, the elevation value is from a few meters to several tens of meters above the sea level. Maryland has vast forests and a wide variety of vegetation. The canopy height is mainly concentrated between 4 m and 35 m. The main forest types include open deciduous broad-leaved forests, closed deciduous broad-leaved forests, open evergreen coniferous forests, closed evergreen coniferous forests, and closed mixed forests. In addition, the main tree species are white oak, American beech, and eastern white pine[17]. The schematic diagram of the study area is shown in Fig. 1, (a) is the fine land types of Maryland, (b) is the fine land types of Puerto Rico.

1.2 ICESat-2/ATLAS data product

The ICESat-2/ATLAS was launched at Vandenberg Air Force Base in the United States for the purpose of continuously measuring changes in the ice sheets, land and sea ice, collecting information on sea and land elevations. The ICESat-2 system is equipped with a new generation of single-photon, multi-beam laser altimeter system, whose ranging principle is defined by the round-trip time of photons. The ATLAS system uses single-photon detection technology and emits green laser pulses with a wavelength of 532 nm at a frequency of 10 kHz. Each laser pulse contains more than 100 trillion photons. The pulses are divided into three pairs and six beams by a spectroscope to obtain photon point cloud data [10]. This system has two lasers, but they generally do not operate simultaneously [18]. This study used the ATL03 and ATL08 data for analysis. The ATL03 data are also known as global geolocated photon data and provide information on five different surface types: land, ocean, sea ice, inland water, and land ice. These data include detailed information on the longitude, latitude, and surface elevation of each photon event. In addition, the ATL03 data provide a rough distinction between potential signals and background events, as well as other parameters that are helpful for advanced processing. The ATL08 data are the level 3A data product of ICESat-2, and these data contain the elevation information of each photon classifica-

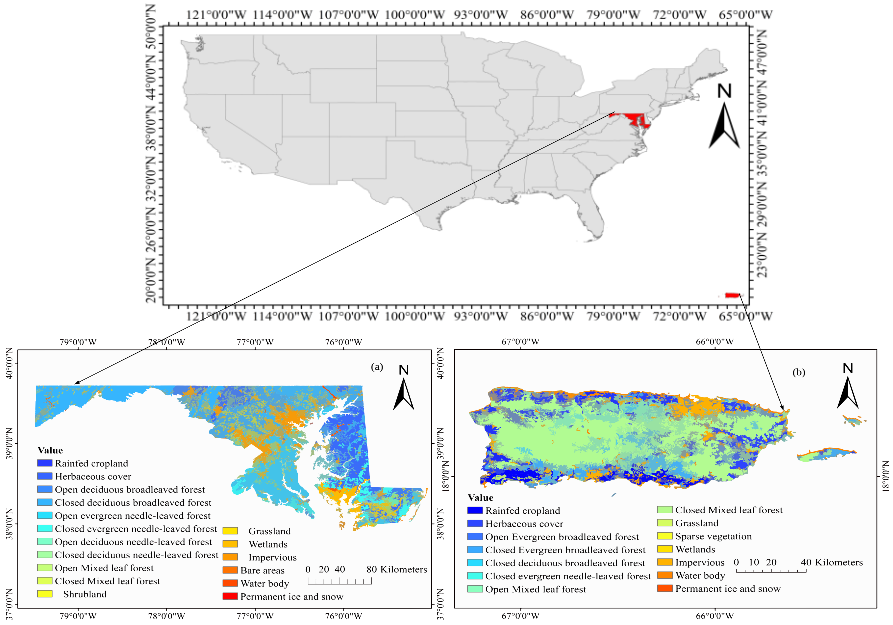

Fig. 1 The schematic diagram of the study area: (a) the fine land types of Maryland; (b) the fine land types of Puerto Rico
图1 研究区域示意图:(a)马里兰州的精细土地类型;(b)波多黎各的精细土地类型

tion and 100-m data along the track section, as well as the estimation data about understory terrain and canopy height[19].

1.3 SRTM data product

The shuttle radar topography mission (SRTM) dataset represents the result of a collaborative effort between the National Aeronautics and Space Administration (NASA), the National Geospatial-Intelligence Agency (NGA), and the space agencies of Germany and Italy. This international space cooperation has yielded an almost global digital elevation model of the Earth, obtained using radar interferometry techniques [20]. The SRTM instrument consists of a modified manned spaceborne imaging radar (SIR-C) hardware suite, a space station-derived mast with a 60 m baseline and additional antennas used to form an interferometer. In addition, SAR is used as a side-looking instrument to acquire data along continuous swaths. The SRTM swath extends from about 30° from the off-track point to about 58° from the off-track point and has a height of 233 km; thus, it is about 225 km wide. During the data flight, the instrument was always running when the orbit flew over land and acquired about 1,000 individual swaths within ten days of mapping operations [21]. The length of the acquired swaths varies from a few hundred kilometers to several thousand kilometers. Each individual data acquisition is called a "data take." This international space cooperation generated an almost global digital elevation model of the Earth.

1.4 Landsat-8 data product

The Landsat-8 was developed and constructed by NASA in cooperation with the U. S. Geological Survey. It represents the eighth satellite in the U. S. Landsat satellite program. It carries two sensors, namely the operational land imager (OLI) and the thermal infrared sensor (TIRS) [22]. The Landsat-8 is mostly consistent with Landsat 1-7 in terms of spatial resolution and spectral characteristics. The satellite has a total of 11 bands. The spatial resolution of bands 1-7 and 9-11 is 30 m, and band 8 is a pan-chromatic band with a resolution of 15 m. The satellite can achieve global coverage once every 16 days. This study used the Landsat-OLI satellite data covering the study area from June 1, 2020 to October 1, 2020[23]. Considering the spectral characteristics of vegetation, different bands have different absorption and reflection capabilities. Therefore, this study aimed to use the corresponding spectral factors related to understory vegetation.

1.5 GLC_FCS30 data product

The GLC_FCS30 data include indispensable and important basic information for climate change research, ecological environment assessment, and geographical national condition monitoring. It is generated from the Landsat satellite data (Landsat TM, ETM+, and OLI) collected from 1984 to 2020. Covering the period from 2015 to 2020, the spatial scope is global. These data contain a total of 29 surface cover types, including cultivated land, forest, grassland, shrubland, wetland, water body, artificial surface, and bare land[24]. Based on the global 30-m fine surface cover classification product in 2020, an automated surface cover dynamic monitoring scheme combining change detection and dynamic update is designed. In addition, using the full-time series Landsat satellite data from 1984 to 2020, independent modeling and dynamic update are performed region by region to obtain the global 30-m fine surface cover dynamic monitoring product from 1985 to 2020. The resolution of 30 m is selected because it can provide relatively fine surface classification results and can accurately present the detailed characteristics of surface cover.

1.6 Canopy height data product

Canopy height data mainly originate from high-resolution LiDAR measurements and provide global information on tree canopy height and forest structure. By analyzing laser echoes from different ecosystems, these data can achieve accurate estimations of the height and coverage of trees, providing important support for forest ecology, carbon cycle research, and biodiversity protection [25]. The generation process of this dataset includes the processing and analysis of laser echoes and the combination of various remote sensing data, such as terrain and vegetation index to ensure the accuracy and reliability of data. The NASA's canopy height data provides an important tool for researchers to monitor forest changes, assess ecosystem health, and support sustainable management.

1.7 Forest cover data product

In NASA's GFC product, vegetation coverage is a key indicator that reflects the vegetation cover status of the Earth's surface and is of great significance for understanding the global ecosystem. NASA mainly uses satellite remote sensing data to obtain relevant information. Sensors such as the moderate-resolution imaging spectroradiometer (MODIS) use the reflection and absorption characteristics of vegetation in different spectral bands to infer vegetation coverage. It also integrates multi-source satellite data to improve accuracy. In the GFC product, it plays a central role in the forest definition. According to the International Geosphere-Biosphere Programme (IGBP) standard, land with an area of at least 0.27 hectares and a tree vegetation coverage of at least 30% is defined as forest. Changes in vegetation coverage denote the key factors for determining forest increase and decrease and can reflect the health status and changes in the forest ecosystem. It should be noted that higher coverage indicates a healthier and more stable ecosystem [26].

2 Method

The accuracy of understory terrain estimation has always been an important factor but is challenging to achieve. In this study, the ICESat-2 data are used as a dependent variable, and the elevation, slope, and aspect values in the SRTM data, as well as different bands, combined bands, and different vegetation indices in the Landsat-8 data are used as independent variables. In addition, the land cover data are used to distinguish and select different forest types. Moreover, different forest types are selected to estimate the understory terrain maps for different forest types by multiple linear regression models and RF random forest models. This study aims to analyze differences in influencing factors in the understory terrain estimation process for different forest types while improving the overall accuracy of understory terrain estimation. The entire analysis process conducted in this study is presented in Fig. 2, where specific data extraction and modeling processes are shown.

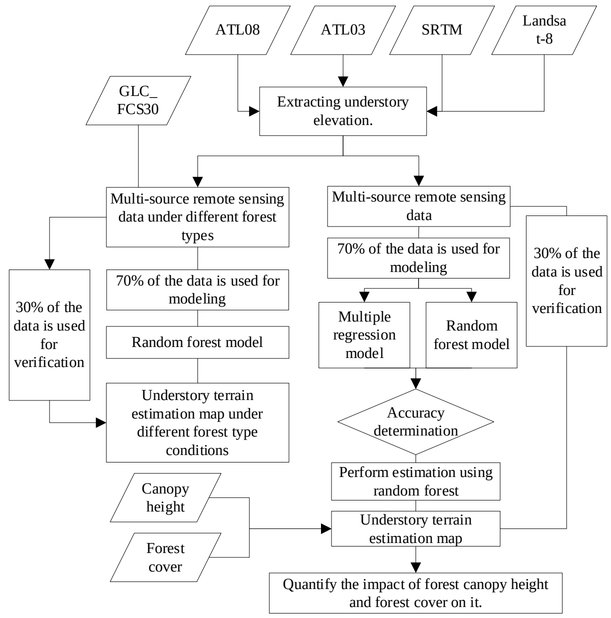

Fig. 2 Illustration of the research process conducted
图2 研究过程示意图

2.1 Data preprocessing

The data preprocessing conducted in this study includes the following steps:

1. The ATL03 and ATL08 products provided by ICESat-2/ATLAS were used in this study, where ATL08 data were employed to provide corresponding label information for ATL03 data[27]. ATL03 data offers high-precision photon detection information, enabling accurate identification of ground reflection signals. ATL08, after processing the original photon data, generates surface elevation data at a specific geospatial resolution. By combining them, using ATL03's high-precision photon detection information to calibrate and optimize ATL08's elevation calculation can reduce errors and enhance terrain measurement accuracy. Using the PhoREAL software, relevant data segments were extracted to associate these datasets. This process enabled the extraction of various parameters from ATL03 data products, including latitude ("lat_ph"), longitude ("lon_ph"), elevation ("h_ph"), and geoid correction information ("geoid"). Additionally, based on ATL08's classification parameter "Classed_pc_flag," ground photons were filtered from ATL03's photon cloud data (Classed_pc_flag=1) to represent ground photons in the understory terrain.

2. The spatial resolution of ATL03 photons is 17 m, while the other three datasets have a spatial resolution of 30 m. Therefore[28], we utilized ArcGIS software Nearest Neighbor Search (NNS) method to resample the spatial resolution of ATL03 data to 30 m. Subsequently, using the latitude and longitude information extracted from ATL03, it was matched with the corresponding footprint's latitude and longitude from the data set, generating understory terrain data. Subsequently, the obtained understory terrain data is divided into two parts: one part is applied to the accuracy evaluation of the regression model, the other part is used as validation data to verify the accuracy of the estimation results.

3. In this study, Landsat 8 OLI band combinations were utilized to quantify and maximize the contrast in the

selected spectral regions, enhancing the correlation with vegetation characteristics. This approach aids in improving pixel contrast and the ability to estimate forest parameters. Extracted vegetation indices in the study include: NDVI, GRVI, RVI, DVI, EVI, SAVI, OSAVI. Additionally, to enhance the experiment's accuracy, the following band combinations were introduced: a: SWIR 1 + NIR + red, b: green + NIR + red, c: NIR + SWIR1 + blue, d: SWIR 2 + NIR + green, e: SWIR 2 + NIR + coastal. Furthermore, using ArcGIS software, factors such as elevation, slope, and aspect of the SRTM DEM product within the study area were considered.

4. This study selected 7 vegetation indices, 5 band combinations, 7 individual bands, as well as elevation, slope, and aspect provided by SRTM, as factors. Before utilizing these 22 factors for estimating understory terrain, a T-test was conducted using SPSS software to determine the correlation between these 22 factors and the understory surface elevation provided by the processed ATL03 data. Meanwhile, to study the impact of various factors on the results under different forest type conditions, this study conducted separate correlation analyses for different forest types during the correlation analysis. In this study, factors that were not significantly correlated at the 0.05 significance level were excluded, retaining only those factors correlated at the 0.01 significance level.

2.2 Cooperative multi-source remote sensing data for understory terrain estimation

Using the preprocessed data, this study obtains the understory terrain modeling data. The modeling data are used to construct the multiple linear regression model and random forest model and improve the modeling accuracy.

2.2.1 Multiple linear regression model

The multiple linear regression model denotes a linear regression model containing multiple explanatory variables and represents a commonly used regression analysis method in the field of statistics. The overall idea of establishing a multiple regression model is to construct a dataset containing feature variables and target variables. Before model training, necessary preprocessing on the dataset, including separating feature and target variables and dividing the data into training and test sets, is performed. The training dataset is used to train a multiple linear regression model and find the linear relationship between the feature and target variables. The feature variable data of the test dataset are input to the model to estimate the target variable's value. In this study on estimating the understory terrain by integrating multi - source remote sensing data under different forest - type conditions, using the multiple linear regression model has certain feasibility.

The study integrates multi-source remote sensing data. For example, ICESat-2 is used to obtain the benchmark values of the understory terrain, SRTM is used to extract topographic factors, and Landsat-8 is used to analyze vegetation cover characteristics. These rich data provide sufficient independent variables for the multiple linear regression model. Moreover, there are certain linear correlation trends among some of the data. For instance, in some areas, there is an approximate linear feature between the slope in SRTM data and the elevation change of the understory terrain obtained from ICESat-2 data. Some vegetation indices in Landsat - 8 data also show a certain linear relationship with the degree to which the understory terrain is affected by vegetation. This meets the basic requirements of the multiple linear regression model for variable relationships and lays a data foundation for model construction.

2.2.2 Random forest model

With the development of remote sensing technology, multi-source remote sensing data can provide rich surface information from different perspectives, which creates conditions for accurately extracting the understory terrain. Different types of remote sensing data, such as the high-precision elevation information provided by ICESat-2, the terrain data from SRTM, and the vegetation cover information from Landsat-8, each have their unique advantages. Fusing these data can comprehensively utilize their strengths and describe the understory terrain and its surrounding environment more comprehensively.

The random forest model, due to its powerful characteristics, becomes an ideal choice for handling such complex data. The forest terrain is comprehensively affected by multiple factors, and the relationships among these factors are highly non-linear, with noise and uncertainties possibly existing in the data[29]. The random forest model has the ability to handle complex non-linear relationships. By constructing multiple decision trees and performing random sampling and feature selection for each decision tree, it can not only effectively capture the complex patterns in the data but also reduce the overfitting risk of the model and enhance its generalization ability. At the same time, it has good robustness to noise and outliers in the data and can maintain high accuracy without complex data pre-processing.

In this study, the understory terrain is complexly influenced by multiple factors such as forest type, topography, and vegetation cover. These factors are intertwined, forming highly non-linear relationships. For example, the differences in vegetation structures of different forest types can lead to different degrees of obstruction of terrain measurement signals, thus affecting the accurate acquisition of the understory terrain. Traditional methods have difficulty fully considering these complex factors and their interactions. The combination of multi-source remote sensing data fusion and the random forest model provides a possible solution to this problem.

The understory terrain elevation information provided by ICESat-2 data was used as the dependent variable, and the relevant factors obtained from SRTM and Landsat-8 data were used as independent variables. After removing the points with blank values, reliable modeling data were obtained. Random sampling with replacement was performed on the original data to construct multiple training subsets. The purpose of this operation was to increase the diversity of the data, enabling the model to learn more comprehensive features. For each training subset, some features were randomly selected to construct decision trees. Node splitting was carried out according to splitting criteria such as the Gini index until stop conditions such as reaching the tree depth limit or too few node samples were met, finally forming a random forest composed of multiple decision trees. This construction method allows the random forest model to automatically capture the complex non-linear relationships in the data, which is exactly the characteristic of the relationship between the understory terrain and various influencing factors.

Sampling was carried out using the land cover type data, with the center of each pixel set as a sampling point. Ensuring that the sampling points did not overlap and the resolution was fixed at 30 m was to guarantee the scientificity and consistency of sampling, enabling the model to accurately reflect the topographic features under different forest types. The corresponding values of each factor were extracted through the longitude and latitude coordinates of the sampling points and input into the random forest model for estimation, thus determining the estimated understory terrain value of each point. Finally, the point-to-raster tool was used to convert the point elevation values and their corresponding longitude and latitude coordinates into raster data to obtain an intuitive understory terrain estimation result. Compared with the multiple linear regression model, the random forest model performs better in handling complex relationships. Its estimation results can more accurately reflect the actual situation of the understory terrain, providing strong support for studying the understory terrain under different forest types.

2.3 Evaluation metrics

The modeling data corresponding to different forest types are separated to obtain the modeling data of the understory terrain estimation model under different forest-type conditions. To construct and verify the model, this study randomly extracts 30% of the data as verification data, and the remaining 70% as modeling data. The random forest model is used to establish understory terrain estimation models for different forest types.

The 30% randomly extracted verification data are used to calculate the accuracy of the original SRTM, the overall understory terrain modeling estimation result, and the understory terrain estimation result for different forest types relative to the ICESat-2 data to obtain propagation accuracy. This accuracy evaluation process can comprehensively evaluate the accuracy of different models and data in understory terrain estimation and provide a reliable evaluation basis for research.

Finally, 30% verification data are used to verify the understory terrain estimation result, thus determining the understory terrain estimation accuracy. In this study, the R^2 and RMSE metrics are selected to evaluate the accuracy and error distribution of the models in understory terrain estimation[30]. To investigate the estimation results for unclassified forest types and classified forest types, this study employs two methods, namely method 1 and method 2, respectively. Method 1 is the RF model without forest type classification, and method 2 is the RF model with forest type classification.

The R^2 metric, also known as the coefficient of determination, is a commonly used statistic in statistics. It represents an index used to evaluate the performance of regression models and represents the proportion of a dependent variable that can be explained by an independent variable. The calculation formula of R^2 is given by:

$$R^2 = 1 - \frac{\sum_i \left(\widehat{y_i} - y_i\right)^2}{\sum_i \left(y_i - \bar{y}\right)^2} \quad , \quad (1)$$

where y_i is the true value, $\widehat{y_i}$ is the predicted value corresponding to true value y_i, and $\bar{y}$ is the average value.

The RMSE metric is a commonly used error index to measure the error between the estimated and true values, and it is defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{m} \left(y_i - \widehat{y_i}\right)^2}{m}} \quad , \quad (2)$$

where m is the number of samples, y_i is the true value, and $\widehat{y_i}$ is the estimated value corresponding to true value y_i.

The smaller the RMSE value is, the better the estimated ability of the model and the smaller the error. The RMSE metric is sensitive to errors and can better reflect the degree of deviation between estimated values and true values and can intuitively show the average error size of predicted values.

3 Results and discussions

3.1 Results of understory terrain estimation using multi-source remote sensing data

The modeling accuracy results of the multiple linear regression model and random forest model are shown in Table 1.

Although the multiple linear regression model could improve the accuracy of understory terrain estimation, its improvement range was limited. This was because the multiple linear regression model assumed a linear relationship between the dependent and independent variables. However, the relationship between understory terrain and influencing factors may be nonlinear. For instance, the influence of factors such as canopy height and forest cover on the understory terrain estimation result might not be a simple linear superposition but a complex, nonlinear relationship. In this case, linear models could not accurately capture these complex relationships, resulting in inaccurate estimations. In addition, the results indicated that compared to the multiple linear regression model, the modeling accuracy of the random forest model was significantly improved. This was because the random forest model was composed of multiple decision trees and could automatically capture nonlinear relationships in the data. The relationship between the understory terrain and various influencing factors could be complex and nonlinear, so the random forest could

Table 1 The results of modeling accuracy of the three methods
表1 三种方法建模精度的结果

Study area	Forest type	Multiple linear regression model		Method 1		Method 2	
		R^2	RMSE/m	R^2	RMSE/m	R^2	RMSE/m
Puerto Rico	Closed evergreen broad-leaved forest	0.99	6.81	0.99	2.28	0.99	1.64
	Closed mixed forest	0.99	8.04			0.99	3.64
Maryland	Open deciduous broad-leaved forest	0.99	2.21	0.99	1.66	0.99	1.47
	Closed deciduous broad-leaved forest	0.99	6.18			0.99	0.82
	Evergreen coniferous forest	0.99	4.23			0.99	0.54
	Closed mixed forest	0.98	5.29			0.99	1.26

better adapt to this complexity than the multiple linear regression model. Therefore, the random forest model was selected as an understory terrain estimation model.

3.2 The estimation results of understory terrain under the random forest model

The estimation map of the understory terrain obtained through the random forest method in this paper is shown in the Fig. 3.

The comparison of the verification accuracy results obtained when the understory terrain was estimated as a whole, without classification, using the random forest model and when it was estimated with classification of different forest types also using the random forest model is presented in Table 2.

The experimental results show that in the study area, the understory terrain estimation relying on multi-source remote sensing data was more accurate than the traditional SRTM in terms of the understory terrain type. R^2 was increased from 0.98 to 0.99, and RMSE was increased from 10.86 m to 5.06 m. This indicates that

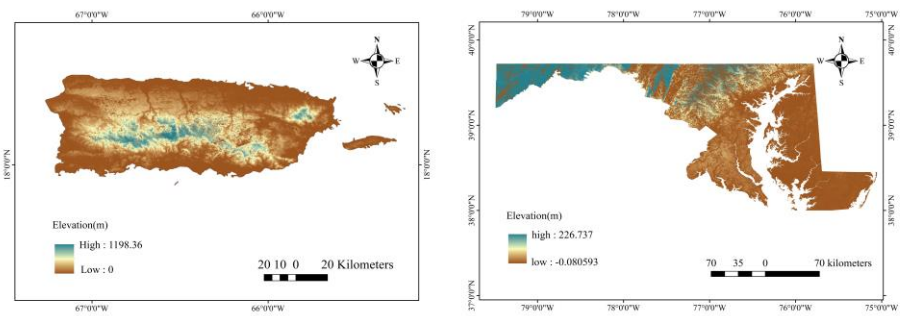

Fig. 3 Understory terrain estimation map
图3 林下地形估算图

Table 2 The estimation accuracy of different methods using the random forest model
表2 基于随机森林模型的不同方法的估算精度

Study area	Method	Vegetation type	SRTM		This study	
			R^2	RMSE/m	R^2	RMSE/m
Puerto Rico	Method 1		0.98*	13.06*	0.99	4.58
	Method 2	Closed evergreen broad-leaved forest	0.99**	13.29**	0.99	2.91
		Closed mixed forest	0.99**	12.87**	0.99	5.59
		overall accuracy	0.99**	13.06**	0.99	3.58
Maryland	Method 1		0.98*	8.67*	0.99	5.55
	Method 2	Open deciduous broad-leaved forest	0.99**	2.78**	0.99	1.20
		Closed deciduous broad-leaved forest	0.99**	9.50**	0.99	2.53
		Closed Evergreen coniferous forest	0.75**	7.17**	0.99	1.86
		Closed mixed forest	0.98**	9.37**	0.99	1.99
		overall accuracy	0.98**	8.67**	0.99	2.29
Average value	Method 1		0.98*	10.86*	0.99	5.06
	Method 2				0.99	2.94

* is accuracy evaluation indicators of SRTM data in different study areas

** is accuracy evaluation indicators of SRTM data in different study areas under vegetation type combining the ICESat-2, SRTM, and Landsat-8 data in understory terrain estimation could provide high credibility, and this method could be used for understory terrain research.

Further, based on the results, when the understory terrain was estimated considering a single forest type, the accuracy of modeling estimation was higher than the overall modeling estimation accuracy without classification according to the forest types. R^2 still remained at 0.99, but RMSE decreased from 4.58 m to 3.58 m in the Puerto Rico research area, and from 5.55 m to 2.29 m in the Maryland research area. This indicated that the classification according to the forest type could increase the accuracy of understory terrain estimation. The average value of the root mean square error (RMSE) for understory terrain estimation in the two study areas has increased from 5.06 m (Method 2) to 2.94 m (Method 1). The research results indicate that modeling by distinguishing forest types contributes to improve the accuracy of understory terrain estimation.

In Puerto Rico, Method 1 showed significant improvements for individual forest types. For example, the Closed Evergreen Broad-Leaved Forest RMSE decreased from 13.29 m to 2.91 m, and the Closed Mixed Forest RMSE reduced from 12.87 m to 5.59 m. The overall accuracy without classification was notably higher than the classified results, further supporting the necessity of forest type-specific modeling. Similarly, in Maryland, Method 1 demonstrated enhanced precision across all vegetation types. The Open Deciduous Broad-Leaved Forest RMSE dropped from 2.78 m to 1.20 m, while the Closed Deciduous Broad-Leaved Forest RMSE improved from 9.50 m to 2.53 m. Even the Closed Mixed Forest RMSE showed a marked reduction from 9.37 m to 1.99 m. The overall accuracy RMSE was higher than most classified forest types, reinforcing the advantage of forest type distinction. The RMSE differences between classified and unclassified modelings highlight that forest type-specific approaches reduce terrain estimation errors more effectively than generalized models.

The comparison between Method 1 and Method 2 reveals a significant improvement in understory terrain estimation accuracy when using Method 2. The average RMSE for Method 1 across both study areas is 5.06 m, while the overall accuracy RMSE for Method 2 is 2.94 m. This indicates that Method 2 reduces estimation errors by approximately 42% compared to Method 1. This substantial improvement can be attributed to the forest type-specific modeling approach employed in Method 2. By distinguishing between different forest types, Method 2 captures the unique structural and ecological characteristics of each vegetation type, leading to more precise terrain estimation. In contrast, Method 1, which does not classify forest types, likely struggles to account for the variability in canopy structure and terrain complexity, resulting in higher RMSE values.

The consistent reduction in RMSE across both study areas (Puerto Rico and Maryland) further validates the robustness of Method 2. For example, in Puerto Rico, Method 2 achieved an overall RMSE of 3.58 m, while in Maryland, it reached an even lower RMSE of 2.29 m. These results demonstrate that Method 2 not only outperforms Method 1 but also provides reliable and accurate terrain estimation across diverse forest ecosystems. In conclusion, the lower average RMSE of Method 2 highlights the importance of incorporating forest type classification into modeling approaches. This strategy enhances the precision of understory terrain estimation and provides a more reliable foundation for ecological and environmental applications.

The scatter plots of understory terrain for different forest types obtained by the original SRTM digital elevation model and the understory terrain estimation model developed using multi-source remote sensing data are presented in Fig. 4.

The accuracy of the original SRTM data was significantly lower, whereas for the estimation with multi-source remote sensing data, the accuracy was significantly improved. In general, it showed good consistency with the verification data. The possible reasons could be as follows.

The ICESat-2, SRTM, and Landsat-8 data were obtained using sensors of different types, and these types of data can provide multiple-aspect information about the understory terrain. The ICESat-2 data directly measured the three-dimensional structure of the ground and vegetation and were highly accurate in capturing the details of the understory terrain. The optical images provided rich information on surface texture and color, which was helpful for identifying different geomorphic features. In contrast, the SRTM data mainly relied on a single radar measurement technology, and these data were relatively single, making it difficult to fully reflect the complex characteristics of the understory terrain.

Multi-source remote sensing data could be integrated using data fusion technology, thus fully leveraging the advantages of different data sources while compensating for their respective deficiencies. Namely, the SRTM and Landsat-8 data could compensate for the defect of the ICESat-2 data that could not be promoted in a planar manner. At the same time, the ICESat-2 data had a stronger penetrating effect compared to the SRTM and Landsat-8 data.

Figure 5 presents the comparison of the estimation results obtained with the classification according to the forest type, independently modeling and inverting each type, and those of the overall terrain evaluation without the forest type classification. The comparison results indicated that the accuracy was improved when the classification was used, and good consistency with the verification data was achieved.

There could be many reasons for such results. Among all forest types, the improvement of the RMSE metric was most obvious for closed mixed forests, rising from 9.37 m to 1.99 m. Closed mixed forests usually have a complex canopy structure. In addition, the heights and foliage distributions of different tree species are diverse, enabling the canopy to form multi-level shad-

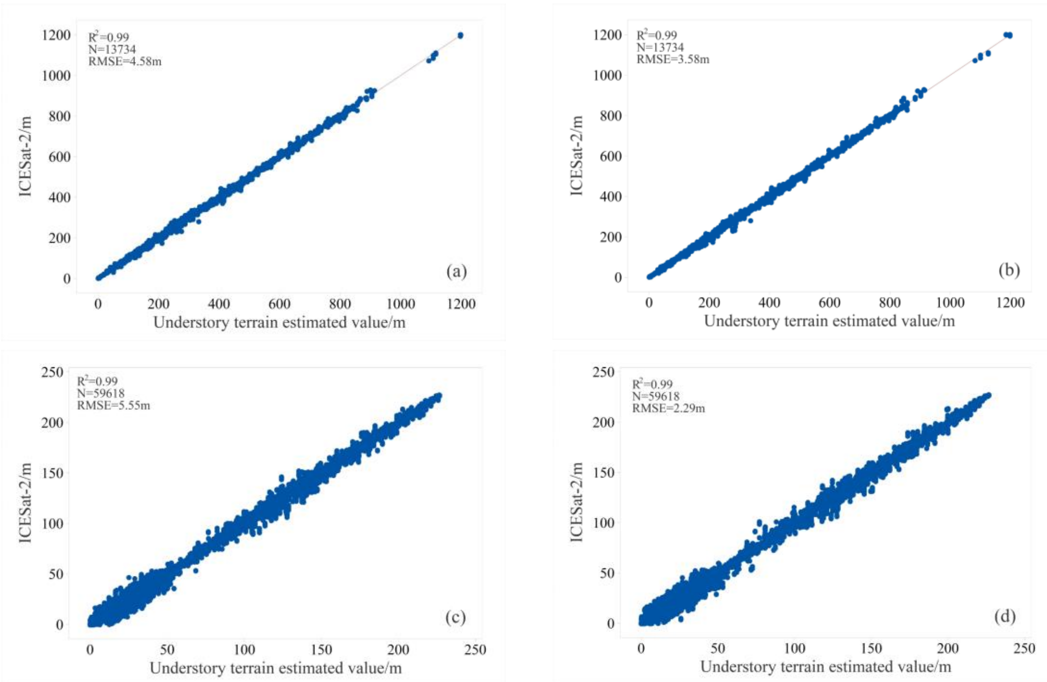

Fig. 4 The classified and unclassified results of the two study areas:(a) Method 1, Puerto Rico;(b) Method 2, Puerto Rico;(c) Method 1, Maryland;(d) Method 2, Maryland

图4 两个研究区域的分类和未分类结果:(a)方法1,波多黎各;(b)方法2,波多黎各;(c)方法1,马里兰州;(d)方法2,马里兰州

ing. Such a complex structure can reduce the interference of external factors (e. g., direct sunlight and atmospheric scattering) on the understory terrain measurement and make the lighting conditions under the forest more diverse. Moreover, the shadows and reflections formed by light with different intensities and incident angles on the understory terrain are also different. This information can provide more reference bases for terrain estimation. At the same time, mixed forests contain multiple tree species, which bring multiple advantages. The root distributions of different tree species are different; some tree species have deep and widely distributed roots, whereas other tree species have shallow roots but large lateral expansions. This diversity in terms of root distribution can result in different degrees of influence on the understory terrain, posing more challenges to terrain estimation. In addition, the growth rates and life cycles of different tree species are also different, making the changes in understory terrain more complex and diverse. Therefore, when estimating terrain, by analyzing the growth characteristics of different tree species and their influence on terrain, the changing trend of understory terrain can be more accurately determined.

The vegetation coverage and growth characteristics of different forest types are different. For instance, broad-leaved forests have dense shrub and herb layers, which will interfere with the signals of terrain measurement equipment; coniferous forests have relatively small leaves and are relatively sparse, so their impact on terrain measurements is relatively small. Hence, distinguishing different forest types and considering vegetation influence factors in a targeted way could improve estimation accuracy. The influence of vegetation on different forest types is also different. In addition, different forest types are often associated with specific terrain conditions. Namely, broad-leaved forests are more inclined to grow in mountainous areas, whereas coniferous forests and mixed forests are suitable for growing in plains or hilly areas. Thus, understanding this correlation when estimating understory terrain can ensure that the terrain characteristics corresponding to certain forest types are combined and reasonable terrain models and parameter settings are selected. Finally, in the correlation analysis, the results of different forest types in screening modeling factors can differ. Furthermore, different forest types might face different interference factors. For instance, broad-leaved forests might be more susceptible to natural disturbances, such as fires, pests, and diseases, whereas coniferous forests might be more affected by human activities, such as logging and reclamation. Consequently, interference factors will have different degrees of influence on understory terrain.

3.3 Analysis of canopy height impact on estimation result

To explore the impact of canopy height on the accuracy of understory terrain estimation, this study used the ICESat-2 data to extract values for canopy height data. In this analysis, areas of 0-10 m, 10-20 m, and greater than 20 m were distinguished, and different canopy height areas were verified separately. The verification results are shown in Fig. 6.

The results showed that as the canopy height increased, the estimation accuracy of understory terrain decreased. R^2 remained at 0.99, and RMSE decreased from 3.35 m to 7.46 m. It becomes evident that during the process where changes in canopy height influence the

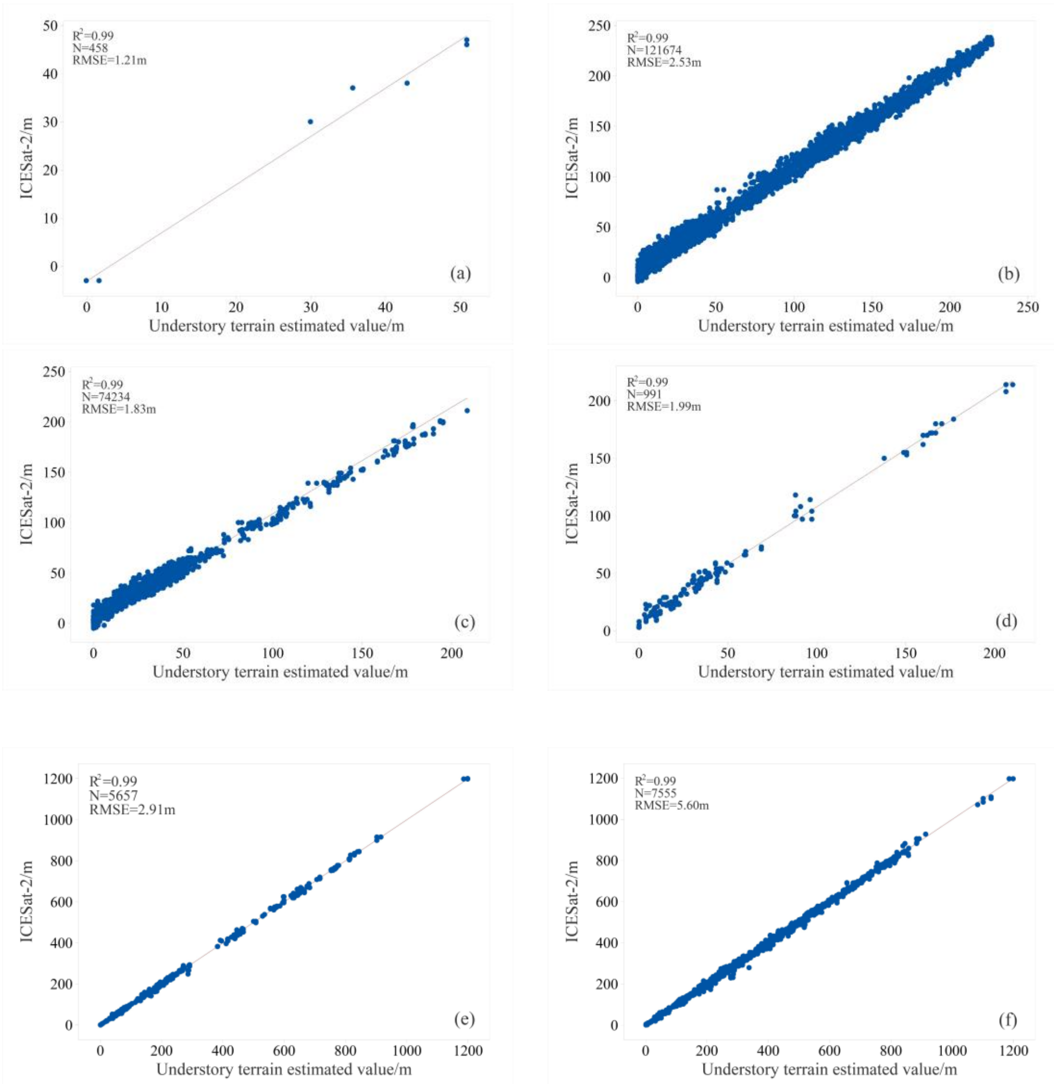

Fig. 5 The scatter plots of the estimation results for different forest types: (a) open deciduous broad-leaved forest; (b) closed deciduous broad-leaved forest; (c) closed evergreen coniferous forest; (d) closed mixed forest; (e) closed evergreen broad-leaved forest; (f) closed mixed forest

图5 不同森林类型估计结果的散点图:(a) 开放落叶阔叶林;(b) 封闭落叶阔叶林;(c) 封闭常绿针叶林;(d) 封闭混交林;(e) 封闭常绿阔叶林;(f) 封闭混交林

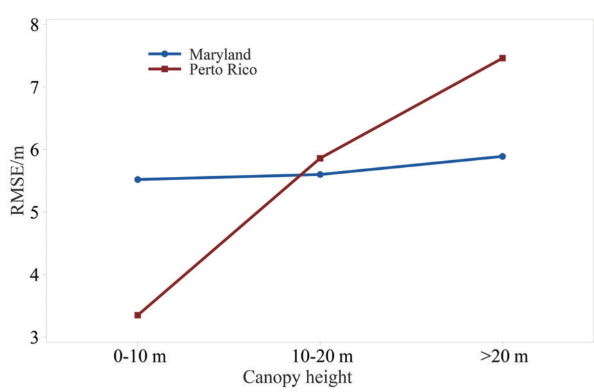

Fig. 6 The analysis results of the canopy height influence on the estimation result

图6 树冠高度对估算结果的影响分析结果

estimation accuracy, RMSE exhibits a remarkable variation of 4.11 m. This clearly indicates that canopy height, being a crucial element within the canopy structure parameters, has a substantial impact on the estimation accuracy.

The canopy height in the range of 0-10 m provided better accuracy for several reasons. First, the signal penetration for this height range was better than for the other ranges. Namely, for commonly used terrain estimation technologies, such as LiDAR, when the emitted laser pulses pass through the forest canopy of 0-10 m, there will be relatively less occlusion and attenuation. Therefore, most laser pulses can reach the forest floor and be reflected and received by the receiver. The abundant and accurate reflected signals provide sufficient data for accurately estimating the understory terrain. Second, vegetation interference is less than in other height ranges. Within this height range, the vegetation hierarchical structure is relatively simple, and there is no multi-layer complex canopy formed by tall trees, which reduces the interference to the terrain estimation signal. In taller forests, branches and leaves at different heights can cause multiple reflections and scatterings of the laser signal, reducing the accuracy of terrain estimation. In addition, data processing is relatively easy. The amount of data collected for forest heights from zero to 10 m is small and simple, making data processing less challenging and the processing process more efficient. Also, errors are less likely to occur, and the algorithm runs faster, enabling a more accurate terrain estimation result to be obtained more rapidly. Moreover, terrain features are more obvious. Namely, in shorter forests, there is less vegetation occlusion, and terrain undulations, slopes, and other features are more easily observed; also, small hills, gullies, and other terrains are easier to identify. In contrast, in tall and dense forests, the terrain may be covered by vegetation. Therefore, the forest height of 0-10 m can more accurately capture terrain details and improve estimation accuracy compared to the other height ranges.

The accuracy of understory terrain estimation in high canopy areas is poor for multiple reasons. First, signal occlusion is severe; the dense tree crowns in high canopy areas cause laser pulses to be reflected, refracted, and absorbed multiple times by layers of branches and leaves during signal propagation. Therefore, the effective signals reaching the forest floor are significantly reduced, and the received reflected signals of the understory terrain are weak and incomplete, making it challenging to reflect the true situation of the understory terrain accurately. In addition, there are complex multiple reflections in high-canopy forests. The lush foliage and complex hierarchy lead to multiple reflections of the measurement signal in the canopy layer. This makes the signal path difficult to determine and the intensity and time information distorted, generating a large amount of noise and interference and thus increasing the difficulty of data processing, which can easily lead to errors in terrain estimation; also, terrain features in high canopy areas are blurred. Tall trees and dense foliage cover the terrain undulations, gullies, and other features under the forest and may also cause shadow areas, affecting the accurate judgment of the understory terrain. Namely, even if understory signals are received, it can be difficult to extract terrain information accurately. Finally, the ecosystem is unstable. The ecosystem in high canopy areas is complex, and tree growth, lodging, and changes such as foliage shedding and renewal are frequent.

3.4 Analysis of forest coverage impact on estimation result

To explore the impact of forest coverage on the accuracy of understory terrain estimation, this paper used the ICESat-2 data to extract the forest coverage data. In this analysis, areas with forest coverage of 0-25%, 25%-50%, and greater than 50% were distinguished, and different forest coverages were verified separately. The verification results are shown in Fig. 7.

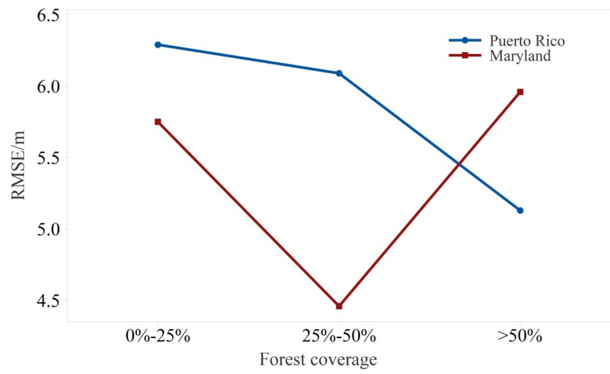

Fig. 7 The analysis results of forest coverage influence on the estimation accuracy
图7 森林覆盖率对估算精度的影响分析结果

The results in Fig. 7 show that in the same study area, the forest coverage had a slight effect on the accuracy of understory terrain estimation, and there was no regular change.

The LiDAR could accurately determine terrain height by emitting laser pulses and measuring their return time. This was because this technology mainly relied on physical principles and was not sensitive to changes in forest coverage. In both sparse and dense forests, as long as the laser could penetrate a certain degree of canopy to reach the ground, relatively accurate terrain data could be obtained.

However, considering the terrain characteristics, the understory terrain was mainly formed by long-term natural processes, such as geological structure, soil erosion, and water flow. These factors had no direct causal relationship with the forest coverage. Even if the forest coverage changed, the basic terrain structure under the forest would not change significantly in a short period of time. For instance, terrain features, such as ridges and valleys of a mountain, would not change fundamentally if the forest became denser or sparser.

From the perspective of the time scale of terrain changes, changes in forest coverage are usually relatively slow and may take several years or even decades to show obvious changes. The changes in understory terrain are often caused by longer term geological actions or sudden natural disasters (e. g., earthquakes and landslides). Therefore, in a shorter time range, the impact of changes in forest coverage on understory terrain is minimal; thus, when estimating the terrain, the relationship of forest coverage compared with canopy height is not significant for estimated accuracy.

4 Conclusions

This study aims to address the shortcomings of the ICESat-2 data of not being able to perform planar understory terrain estimation and the problem of a low accuracy of the Landsat-8 data and SRTM data in understory terrain estimation. To this end, it conducts research on understory terrain estimation using multi-source remote sensing data, combining the ICESat-2, Landsat-8, SRTM, and land cover data. The study areas include Puerto Rico and Maryland. The multiple linear regression model and random forest model are used for modeling to realize understory terrain estimation.

Based on the obtained results, the following conclusions are drawn:

1. Compared to the traditional SRTM, using multi-source remote sensing data in understory terrain estimation, the RMSE metric is improved from 10.86 m to 5.06 m, and the R^2 value is improved from 0.98 to 0.99. This indicates that the proposed method shows less fluctuation when facing understory terrain changes, increasing model stability. Thus, it is more suitable for accurate prediction of understory terrain than traditional SRTM.

2. By regarding forest types as distinguishing conditions, this study constructs understory terrain estimation models for different forest types in the study area and finds that the understory terrain estimation accuracy under the condition of a single forest type is better than that without classification. The RMSE value of the understory terrain estimation is improved from 5.06 m to 2.94 m, and the R^2 value remains at 0.99. This indicates good reliability in understory terrain estimation under various forest types.

3. The analysis of the impact of canopy height and forest coverage on the understory terrain estimation performance shows that as the canopy height increases, the accuracy of understory terrain estimation decreases. The R^2 value remains at 0.99, and the RMSE value of the understory terrain estimation is 3.35 m in the low canopy area, 5.58 m in the medium canopy area, 7.46 m in the high canopy area and the RMSE value exhibits a remarkable variation of 4.11 m. The results indicate that canopy height has more impact than forest coverage on understory terrain estimation.

Currently, although we have completed the research on this method, the study area is only a partial region of the United States. This to some extent limits the practicability of the method. In the follow-up, the study area should be expanded. Regions with different geographical environments and climatic conditions can be selected, such as countries and regions in different continents such as Europe, Asia, and Africa. This can test the applicability of the method in different natural environments. At the same time, it can also better understand the characteristics and needs of different regions. Then, the method can be optimized and adjusted to make it more universal and give full play to the value of this method.

References

[1] Wang M, Ma X, Zheng T, et al. MSMTRIU-Net: Deep learning-based method for identifying rice cultivation areas using multi-source and multi-temporal remote sensing images[J]. Sensors, 2024, 24(21): 6915.

[2] Huang J, Wang Y, Yu Y. Multi-criteria filtration and extraction strategy for understory elevation control points using ICESat-2 ATL08 product[J]. Forests, 2024, 15(12): 2064.

[3] Li H, Liu S, Cai T. Measurement of forest ecological benefits based on big data[J]. Sustainability, 2022, 14(12): 7248.

[4] Wan H, Tang Y, Jing L, et al. Tree species classification of forest stands using multisource remote sensing data[J]. Remote Sensing, 2021, 13(1): 144.

[5] Hogland J, Anderson N, Affleck D, et al. Using forest inventory data with Landsat 8 imagery to map longleaf pine forest characteristics in Georgia, USA[J]. Remote Sensing, 2019, 11(15): 1803.

[6] Yu Y, Li M, Fu Y. Forest type identification by random forest classification combined with SPOT and multitemporal SAR data[J]. Journal of Forestry Research, 2018, 29(5): 1407-1414.

[7] Liu Ya-Jia, Zhu Jian-Jun, Fu Hai-Qiang. Inversion of underforest terrain by dual-polarization interferometric SAR[J]. Engineering of Surveying and Mapping, 2020, 29(5): 20-26.
刘雅佳,朱建军,付海强.双极化干涉SAR林下地形反演[J].测绘工程,2020,29(5):20-26.

[8] Zhang Jun-Xiang. Research on the inversion of forest aboveground biomass by combining optical and SAR data[D]. Central South University of Forestry and Technology, 2024.
张浚翔.联合光学和SAR数据的森林地上生物量反演研究[D].中南林业科技大学,2024.

[9] Choi H, Song Y, Jang Y. Urban forest growth and gap dynamics detected by yearly repeated airborne light detection and ranging (LiDAR): a case study of Cheonan, South Korea[J]. Remote Sensing, 2019, 11(13): 1551.

[10] Huang Jia-Peng, Li Guo-Yuan, Liu Zhao. Current status analysis and prospect of forest structure parameter estimation using spaceborne LiDAR[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 18-33.
黄佳鹏,李国元,刘诏.星载激光雷达估测森林结构参数研究现状分析与展望[J].农业机械学报,2024,55(6):18-33.

[11] Huang J, Yu Y. Vertical accuracy assessment of the ASTER, SRTM, GLO-30, and ATLAS in a forested environment[J]. Forests, 2024, 15(3): 426.

[12] Huang J, Yang Y, Yu Y, et al. Vertical accuracy of open-source remote sensing data (AW3D30, TanDEM-X, ATLAS) for understory terrain estimation [J]. Geocarto International, 2024, 39 (1): 2356855.

[13] Wang C, Zhu X, Nie S, et al. Ground elevation accuracy verification of ICESat-2 data: a case study in Alaska, USA[J]. Optics Express, 2019, 27(26): 38168-38179.

[14] Zhang Chen, Zhu Jian-Jun, Fu Hai-Qiang. Understory terrain inversion based on ICESat-2 and TanDEM-X DEM[J]. Engineering of Surveying and Mapping, 2021, 30(1): 60-65.
张晨,朱建军,付海强.基于ICESat-2数据及TanDEM-X DEM的林下地形反演[J].测绘工程,2021,30(1):60-65.

[15] Huang J, Zhang Y, Yu Y. Mathematical model guided interpolation for mapping SRTM understory terrain by integrating ICESat-2 data [J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 1-5.

[16] Brandeis T, Woodall C. Assessment of forest fuel loadings in Puerto Rico and the U.S. Virgin Islands[J]. Ambio, 2009, 37(7/8): 557-562.

[17] Hansen M C, Potapov P V, Moore R, et al. High-resolution global maps of 21st-Century forest cover change[J]. Science, 2013, 342(6160): 850-853.

[18] Neumann T A, Martino A J, Markus T, et al. The ice, cloud, and land elevation satellite - 2 mission: a global geolocated photon product derived from the advanced topographic laser altimeter system [J]. Remote Sensing of Environment, 2019, 233: 111325.

[19] Parrish C, Magruder L, Neuenschwander A, et al. Validation of ICESat-2 ATLAS bathymetry and analysis of ATLAS's bathymetric mapping performance[J]. Remote Sensing, 2019, 11(14): 1634.

[20] Zhu Xiao-Xiao, Wang Cheng, Xi Xiao-Huan, et al. Research progress on data processing and application of ICESat - 2 spaceborne photon counting LiDAR[J]. Infrared and Laser Engineering, 2020, 49(11): 76-85.
朱笑笑,王成,习晓环,等.ICESat-2星载光子计数激光雷达数据处理与应用研究进展[J].红外与激光工程,2020,49(11):76-85.

[21] Chen Jun-Yong. Assessment of the quality of SRTM3 and GTOPO30 terrain data[J]. Geomatics and Information Science of Wuhan University, 2005, 30(11): 4-7.
陈俊勇.对SRTM3和GTOPO30地形数据质量的评估[J].武汉大学学报(信息科学版),2005,30(11):4-7.

[22] Roy D P, Wulder M A, Loveland T R, et al. Landsat-8: Science

and product vision for terrestrial global change research[J]. Remote Sensing of Environment, 2014, 145: 154-172.

[23] Vermote E, Justice C, Claverie M, et al. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product [J]. Remote Sensing of Environment, 2016, 185: 46-56.

[24] Zhang X, Liu L, Chen X, et al. GLC_FCS30: global land-cover product with fine classification system at 30m using time-series Landsat imagery[J]. Earth System Science Data, 2021, 13(6): 2753-2776.

[25] Li Yi, Zhu Jian-Jun, Fu Hai-Qiang, et al. Research progress on the inversion of forest height and underlying terrain from spaceborne photon counting LiDAR data[J]. Journal of Central South University (Science and Technology), 2023, 54(11): 4380-4390.
李毅, 朱建军, 付海强, 等. 星载光子计数激光雷达数据森林高度及林下地形反演研究进展[J]. 中南大学学报(自然科学版), 2023, 54(11): 4380-4390.

[26] Li J L, Cheng W, Sheng N, et al. Analysis of the influence of different algorithms of GEDI L2A on the accuracy of ground elevation and forest canopy height[J]. Journal of University of Chinese Academy of Sciences, 2022, 39(4): 502.

[27] Popescu S C, Zhou T, Nelson R, et al. Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data[J]. Remote Sensing of Environment, 2018, 208: 154-170.

[28] Beretta L, Santaniello A. Nearest neighbor imputation algorithms: a critical evaluation[J]. BMC Medical Informatics and Decision Making, 2016, 16(S3): 74.

[29] Biau G, Scornet E. A random forest guided tour[J]. Test, 2016, 25(2): 197-227.

[30] Chai T, Draxler R R. Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature[J]. Geoscientific Model Development, 2014, 7(3): 1247-1250.