remote sensing

MDPI

Article

Multidimensional Validation of FVC Products over Qinghai–Tibetan Plateau Alpine Grasslands: Integrating Spatial Representativeness Metrics with Machine Learning Optimization

Junji Li [1], Jianjun Chen [1,*], Xue Cheng [1], Jiayuan Yin [1], Qingmin Cheng [1], Haotian You [1], Xiaowen Han [1] and Xinhong Li [2]

[1] Guangxi Key Laboratory of Spatial Information and Geomatics, College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China; lijunji@glut.edu.cn (J.L.); cx12@glut.edu.cn (X.C.); yinjy@glut.edu.cn (J.Y.); chengqingmin123456@glut.edu.cn (Q.C.); youht@glut.edu.cn (H.Y.); xwhan@glut.edu.cn (X.H.)

[2] Guizhou Institution of Prataculture, Guizhou Academy of Agricultural Sciences, Guiyang 550006, China; lxh@glut.edu.cn

* Correspondence: chenjj@glut.edu.cn

Highlights

What are the main findings?

- A framework synergizing spatial representativeness, ML tuning, and multidimensional validation enhances FVC product validation robustness.
- GEOV3 demonstrates superior accuracy compared to GLASS and MuSyQ FVC products on alpine grasslands.

What are the implications of the main findings?

- The proposed framework provides a methodological basis for field sampling design and scale-matched validation of satellite FVC in heterogeneous landscapes.
- The comparative results can offer scientific guidance for selecting appropriate FVC products and improving retrieval algorithms for alpine ecosystem monitoring on the Qinghai–Tibetan Plateau.

Abstract

Fractional Vegetation Cover (FVC) dynamics on the Qinghai–Tibetan Plateau (QTP) are critical indicators for assessing ecosystem condition. However, uncertainties persist in the accuracy of existing FVC products over the QTP due to retrieval differences, scale effects, and limited validation data. This study utilized the Google Earth Engine platform to integrate unmanned aerial vehicle (UAV) observations, Sentinel-2, MODIS, climate, and topography datasets, and proposed a comprehensive framework incorporating dual-index screening, machine learning optimization, and multidimensional validation to systematically assess the accuracy of GEOV3, GLASS, and MuSyQ FVC products in the alpine grasslands. The dual-index screening reduced validation uncertainty by improving the spatial representativeness of ground data. To build a high-precision evaluation dataset with limited inter-class coverage, recursive feature elimination and grid search were applied to optimize five ML models, and CatBoost achieved the superior performance (R^2 = 0.880, RMSE = 0.122), followed by XGBoost, GBM, LightGBM, and RF models. Four validation scenarios were implemented, including direct validation using 250 m UAV plot FVC and multi-scale validation using a 10 m FVC reference aggregated to product grids. Results show that

Check for updates

Academic Editor: Kevin Tansey

Received: 12 December 2025
Revised: 6 January 2026
Accepted: 8 January 2026
Published: 10 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

GEOV3 (R^2 = 0.909–0.925, RMSE = 0.082–0.103) outperformed GLASS (R^2 = 0.742–0.771, RMSE = 0.138–0.175) and MuSyQ (R^2 = 0.739–0.746, RMSE = 0.138–0.181), both of which exhibited systematic underestimation. This framework significantly enhances FVC product validation reliability, providing a robust solution for remote sensing product validation in alpine grassland ecosystems.

Keywords: fractional vegetation cover; product validation; machine learning; Qinghai-Tibetan Plateau

1. Introduction

The Qinghai–Tibetan Plateau (QTP), known as the water tower of Asia and the third pole of the world, exerts a fundamental influence on downstream water security and the global climate. Alpine grasslands dominate about 70% of the region. Ongoing climate warming and increasing human pressures can alter vegetation structure and cover, making continuous monitoring of vegetation condition essential for detecting changes and supporting ecosystem management [1,2]. Fractional Vegetation Cover (FVC), defined as the ratio of the vertical projection area of vegetation (including roots, stems, and leaves) to unit ground area, is a core indicator of ecosystem condition and an essential input to hydrologic, climate, and ecological models [3–5]. Global FVC products such as GEOV, GLASS, CYCLOPES, and MuSyQ are widely used. However, on the QTP, harsh climatic conditions and rugged topography severely constrain in situ observations, hindering robust validation of these products [6–8]. Together with algorithm and scale differences, this has left regional reliability insufficiently quantified and product accuracy uncertain, which constrains applications [9]. Therefore, region-specific validation is needed to quantify product performance and support reliable product selection and use over the QTP.

Validation is a systematic comparison of remote sensing products with reference data that represent true surface conditions and enables quantitative assessment of accuracy and reliability [10]. The main approaches are direct validation, multi-scale validation, and cross-validation. Direct validation compares ground measurements with satellite pixels and is most suitable where vegetation is uniform and surface heterogeneity is low [11]. Multi-scale validation uses high-resolution imagery as a bridge linking ground observations to medium and low resolution products; it reduces scale mismatch but introduces retrieval and aggregation uncertainties during scale transitions [12,13]. Cross-validation refers to inter-product comparison among spatiotemporally consistent FVC datasets. Because no in situ measurements are used for calibration in this step, it cannot determine an absolute truth, but it can quantify relative agreement and highlight systematic differences among products [14]. Recent advances include multi-source spatiotemporal harmonization, relative error decomposition such as triple collocation and error in variables regression, hierarchical Bayesian uncertainty modeling, and regional evaluations of updated FVC products. For the highly heterogeneous QTP, a multitier validation system that integrates field observations with high-resolution imagery is needed to improve reliability.

Validation errors arise mainly from three sources: reference data errors due to suboptimal sampling and limited representativeness, scale effects from pixel size discrepancies, and uncertainties in retrieval models and parameterization [15,16]. Robust assessment requires a relative truth at the product pixel scale derived from ground measurements or high-resolution references. Widespread heterogeneity often makes point data unrepresentative, creating discrepancies between site observations and product values, especially in complex terrain or fragmented vegetation [17]. Multi-scale validation alleviates scale mismatch but remains constrained by uncertainties in generating the high-resolution ref-

erence and by amplification of representativeness deficits during scaling [18]. If plots do not capture the within-pixel distribution of key attributes, nonlinear aggregation during scaling can invalidate locally precise measurements and reduce accuracy [10]. Hence representativeness evaluation and careful selection and tuning of retrieval models are essential for reliable validation.

Therefore, this study conducts a context-specific accuracy assessment of FVC data products over alpine grasslands, compares mainstream products across multiple spatial scales and years, identifies error characteristics and spatial patterns, and provides a basis for product selection and application. Specifically, the objectives of this study were: (1) To develop and apply representativeness metrics for ground measurements, namely the Relative Spatial Sampling Error (RSSE) and an NDVI-based Spatial Heterogeneity Index (NDVI-SHI), in order to optimize their suitability for FVC validation; (2) To comparatively analyze the performance of RF, GBM, LightGBM, XGBoost, and CatBoost models for FVC retrieval and identify the optimal algorithm for generating reference validation datasets; and (3) To produce a high-spatial-resolution validation dataset and establish a multidimensional validation framework to evaluate the applicability of GEOV3, GLASS and MuSyQ FVC products over alpine grasslands on the QTP, thereby providing a robust foundation for product utilization.

2. Materials and Methods

2.1. Study Area

The QTP (26°00′–39°46′N, 73°18′–104°46′E) is located in western China and covers approximately 2.57 × 10^6 km^2 (Figure 1). With a mean elevation of about 4000 m, it has a typical plateau mountain climate with a mean annual temperature of 1.6 °C and mean annual precipitation of 413.6 mm. Alpine grasslands, including alpine meadows and steppes, are the dominant ecosystem and provide key ecosystem services such as water regulation, soil conservation, carbon sequestration, and biodiversity maintenance. The surface is highly heterogeneous, with vegetation and bare-ground patches forming a complex mosaic and strong illumination contrasts driven by rugged terrain (Figure 2). Such fine-scale heterogeneity can be smaller than the pixel sizes of commonly used FVC products and may increase uncertainty in product application. Therefore, validating FVC products over the QTP is necessary to quantify product uncertainty and support reliable product selection and use in this plateau environment.

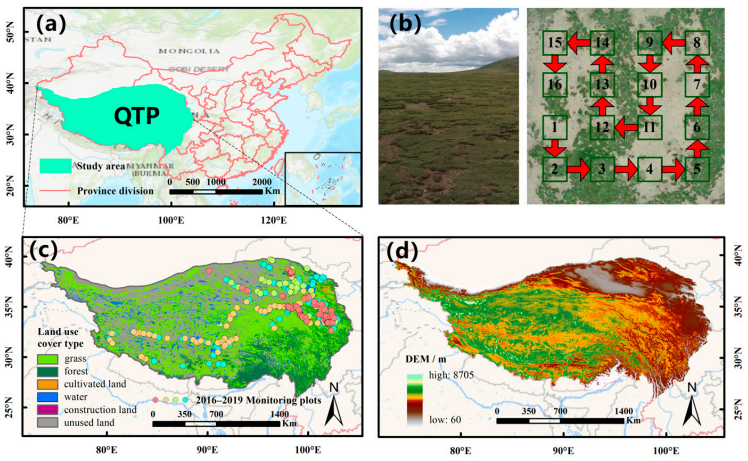


Figure 1. Study area and monitoring plot distribution on the Qinghai–Tibetan Plateau (QTP). (**a**) Location of the QTP in China. (**b**) Alpine grassland field photo and schematic of within-site plot layout. (**c**) Land use-cover types and the locations of monitoring plots (2016–2019). (**d**) Digital elevation model (DEM) of the QTP. All monitoring plots are located within alpine grassland coverage.

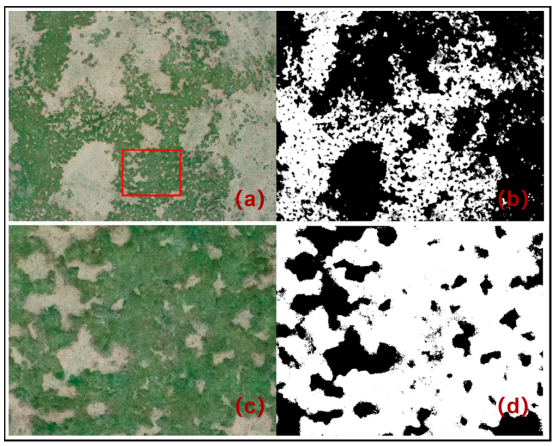


Figure 2. Aerial imagery and processing results: (**a**) original aerial image acquired by UAV; (**b**) EGI segmentation result of the aerial image; (**c**) region within the red rectangle in (**a**); (**d**) EGI segmentation result of the region shown in (**c**), where white indicates vegetation areas and black indicates non-vegetation areas.

2.2. Data Sources and Preprocessing

2.2.1. UAV Data Collection and Preprocessing

Over the July–August periods spanning 2016–2019, our team established 623 remote sensing monitoring plots across the QTP and acquired UAV imagery, including 113 plots in 2016, 119 in 2017, 95 in 2018, and 296 in 2019. Each plot measured 250 m × 250 m to match the MOD13Q1 pixel. Before fieldwork, MOD13Q1 pixel centers were extracted and plots positioned around them. The plots were deployed across the QTP to span a wide range of FVC conditions and surface settings, forming a plateau-scale observation network. Spatial heterogeneity within each plot was captured through a standardized within-plot sampling layout, allowing sub-pixel variability to be characterized consistently across sites. Autonomous flights were planned with the Fragment Monitoring and Aerial Photography Analysis System (FragMAP) [19]. Each route adopted a concave layout with 16 survey points uniformly arranged within a 30 m × 30 m subplot (Figure 1b) to capture fine-scale FVC variability. A square 2 km × 2 km buffer centered on each plot was created in GEE using meter-based geometry: the center was buffered by 1 km, then the minimum bounding rectangle was used as the extraction window for the FVC validation dataset.

The UAV was a DJI Phantom 4 Pro V2.0 (DJI Technology Co., Ltd., Shenzhen, China) with GPS/GLONASS positioning, providing horizontal accuracy of ±1.5 m and vertical accuracy of ±0.5 m. Flights at 20 m altitude produced RGB imagery at approximately 2 cm resolution. To automate FVC extraction, we developed software integrating Excess Green Index thresholding with OpenCV [20,21]. For each image, per-pixel EGI was computed, an initial threshold within a reference range produced a vegetation or non-vegetation mask, the mask was overlaid for visual inspection, and the threshold was iteratively adjusted until segmentation stabilized; FVC was then calculated as the fraction of vegetation pixels. To reduce subjectivity, two analysts processed each image independently; if their FVC estimates differed by more than 0.05, a third analyst adjudicated until agreement. Representative aerial images and EGI segmentation results are shown in Figure 2.

2.2.2. Satellite Data

The Sentinel 2 Multispectral Instrument acquires 13 spectral bands, with visible bands at 10 m and near-infrared and shortwave infrared bands at 20 m. Its five day revisit enables high-frequency observation [22]. We obtained Sentinel 2 surface reflectance over July–August periods spanning 2016–2019 via Google Earth Engine (GEE; https://earthengine.google.com/). Scenes with less than 10% cloud were retained, then

a quality assurance band mask removed remaining low-quality pixels. NIR and SWIR bands were resampled to 10 m by bilinear interpolation to ensure spatial consistency. The Moderate Resolution Imaging Spectroradiometer (MODIS) provides 36 bands from 0.4 to 14 μm and products at 250 m, 500 m, and 1000 m [23]. We used the MOD13Q1 vegetation index product containing NDVI and EVI at 250 m with a 16 day interval. MOD13Q1 data over the July–August periods spanning 2016–2019 were extracted from GEE. Cloud masking with the DetailedQA band removed low-quality pixels. Annual composites for Sentinel 2 and MODIS were generated using the median composite function in GEE, then mosaicked and clipped to the study area.

2.2.3. FVC Product Data

This study used three FVC products and harmonized them for joint analysis: Copernicus GEOV FCover (GEOV3), GLASS FVC, and MuSyQ FVC. GEOV3 spans 2014 to the present with 300 m spatial and 10 day temporal resolution, is generated with neural networks, stored in NetCDF, and is available from the Copernicus Land Monitoring Service [24]. GEOV3 data were processed in ArcGIS 10.8 for extraction, mosaicking, resampling, and reprojection. GLASS provides global FVC at 500 m and 8 day resolution for 2000–2020; early products used a GRNN algorithm, and later versions adopted MARS to improve accuracy and stability. GLASS data, distributed in HDF by the National Earth System Science Data Center, were processed with the MODIS Reprojection Tool for extraction, mosaicking, resampling, and projection transformation [25]. The MuSyQ FVC product and its quality control layer were obtained from the same center; the QC masked water, urban areas, clouds, shadows, snow, and other low-quality pixels, retaining only high-quality pixels. Maximum Value Compositing was applied to GEOV3, GLASS, and MuSyQ datasets over July–August periods spanning 2016–2019 to generate annual FVC composites. For validation, comparisons were conducted at scale-matched resolutions: GEOV3 was evaluated at 300 m, while GLASS and MuSyQ were evaluated at 500 m by aggregating the reference data to the corresponding grids.

2.2.4. Auxiliary Data

Auxiliary datasets in this study include topographic and climatic information. Topography was characterized using the 30 m Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) from NASA, generated by InSAR and providing high horizontal and vertical accuracy. Meteorological variables were taken from the ERA5 reanalysis produced by ECMWF, which combines model simulations and global observations to provide a continuous climate field with 0.1° spatial resolution and hourly temporal resolution since 1979. To ensure spatial consistency, the SRTM DEM and ERA-5 data were resampled to a uniform 10 m resolution using bilinear interpolation. Detailed specifications of all datasets are presented in Table 1.

Table 1. Data sources information.

Data Type	Specific Data	Resolution	Data Source
Measured data	FVC ground data	2 cm	UAV data
Satellite data	Sentinel-2 data	10 m	GEE platform
	MODIS data	250 m	GEE platform
FVC product data	GEOV3 FVC data	300 m	Copernicus Land Monitoring Service
	GLASS FVC data	500 m	National Earth System Science Data Center
	MuSyQ FVC data	500 m	National Earth System Science Data Center
Auxiliary data	Topographic data	30 m	SRTM-DEM data (NASA)
	Climatic data	0.1°	ERA-5 reanalysis data (ECMWF)

2.3. Methods

This study focused on assessing the reliability of three global FVC products over the QTP under conditions of strong heterogeneity and sparse observations. To obtain a relative truth at the pixel scale, we integrated multi-source remote sensing and ground observations on the GEE platform, first performing spatial representativeness diagnostics to mitigate point-pixel mismatch, and then applying feature selection and parameter optimization to construct a multi-model ensemble that produced a robust reference FVC. Under a unified spatiotemporal baseline, four validation scenarios were subsequently designed according to validation strategy and reference-data form to conduct multidimensional assessments of the GEOV3, GLASS, and MuSyQ FVC products. The research workflow is illustrated in Figure 3.

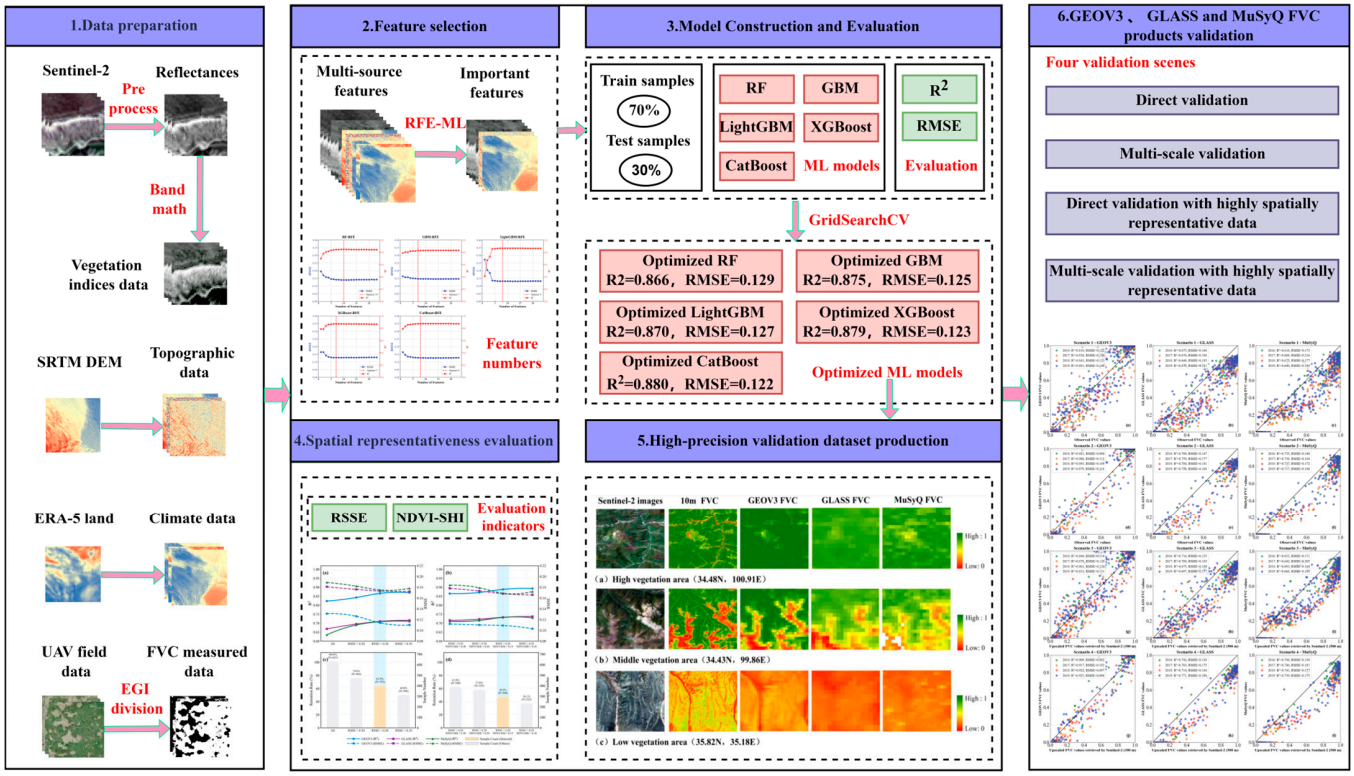


Figure 3. Flowchart of this study.

2.3.1. Spatial Representativeness Evaluation

Centered on harmonizing point to area characteristics, the point to area comparison evaluates representativeness by analyzing discrepancies between site measurements and regional statistics (for example, the mean). Using 250 m inversion-based imagery, this study defined the Spatial Sampling Error (SSE) and the Relative Spatial Sampling Error (RSSE) to quantify deviations between site measurements and co-located pixels. RSSE normalizes the difference by the measured plot-scale FVC to enable comparison across FVC levels and to avoid permissive absolute thresholds under low FVC conditions [26]. When the discrepancy is below a preset threshold, the site measurement is deemed representative at the target scale. The SSE and RSSE are computed as follows:

$$SSE = /FVC_R - FVC_M/ \quad (1)$$

$$RSSE = \frac{SSE}{FVC_R} \quad (2)$$

where FVC_R denotes the measured FVC at 250 m plot scale, and FVC_M represents the co-located FVC estimated from MODIS data inversion.

Surface spatial heterogeneity assessment, rooted in landscape ecology, characterizes spatial variability within the sampling window. As NDVI is closely related to vegetation cover and canopy density, an NDVI-based Spatial Heterogeneity Index (NDVI-SHI) was used to evaluate site-level heterogeneity. SHI characterizes the dispersion of surface cover within the window; smaller values indicate a more homogeneous window and better site representativeness [27]. Based on MOD13Q1, NDVI-SHI was computed by taking the root mean square of NDVI differences between the central pixel and its neighbors [28,29]. Here, NDVI-SHI is calculated on the MOD13Q1 250 m grid using the central pixel and its eight neighbors, so it describes within-window heterogeneity at the same spatial support as the 250 m plot-scale validation data. This scale is also close to the native resolutions of GEOV3, GLASS, and MuSyQ, making NDVI-SHI suitable for screening samples for direct product validation. Measurements with NDVI-SHI below a threshold were retained as low heterogeneity, whereas those above the threshold were excluded. RSSE quantifies the point-to-pixel mismatch, whereas NDVI-SHI directly describes within-window spatial heterogeneity. By jointly applying RSSE and NDVI-SHI thresholds, we excluded plots with strong sub-pixel variability, thereby reducing heterogeneity-driven representativeness errors in product validation. The formulas for NDVI and NDVI-SHI are as follows:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{3}$$

where *NIR* and *RED* represent near-infrared and red band reflectance, respectively.

$$NDVI - SHI = \sqrt{\frac{1}{8}\sum_{i=1}^{8}(NDVI_i - NDVI_0)^2} \tag{4}$$

where $NDVI_0$ denotes the NDVI value of the central pixel and $NDVI_i$ (i = 1 to 8) corresponds to the NDVI values of the eight surrounding pixels.

To accommodate different cover conditions, representativeness thresholds were set in a segmented manner: the RSSE threshold was stratified by the reference FVC level, with stricter thresholds for low FVC and more relaxed thresholds for medium to high FVC; the breakpoint values were determined by sensitivity analysis using the responses of Root Mean Square Error (RMSE) and Coefficient of Determination (R^2) together with a trade-off in sample size.

2.3.2. Multi-Source Heterogeneous Feature Dataset Construction

A multidimensional feature ensemble was constructed by integrating Sentinel-2 data, SRTM digital elevation data, and ERA-5 reanalysis climate products accessed through GEE. Annual median-composed Sentinel-2 imageries overlapping ground sampling plots over July–August periods spanning 2016–2019 were initially generated using GEE's median composite function. Subsequent feature extraction encompassed FVC-sensitive multispectral bands (B2-B8, B8A, B9, B11, B12) [30] and vegetation indices capturing phenological dynamics, including the NDVI, EVI, Soil-Adjusted Vegetation Index (SAVI), Kernel NDVI (KNDVI), Ratio Vegetation Index (RVI), Extended Ratio Index (EXR), and Difference Vegetation Index (DVI) [18]. Concurrently, topographic parameters (elevation, slope, aspect) were derived from SRTM DEM via terrain analysis, while key bioclimatic variables, namely Mean Annual Temperature (MAT), Mean Annual Precipitation (MAP), and Mean Wind Speed (MWS), were extracted from ERA-5 datasets to comprehensively characterize environmental conditions. Previous studies have demonstrated that incorporating meteorological

and other ancillary variables as covariates in ecological parameter retrieval can effectively reduce uncertainty and improve the accuracy of the inversion results [31,32]. Detailed specifications of the multidimensional feature dataset are presented in Table 2.

Table 2. Multidimensional feature dataset in this study.

Date Types	Names	Spatial Resolution	Calculation Formula
Reflectance	B2-Blue	10 m	*Sentinel-2 data*
	B3-Green	10 m	
	B4-Red	10 m	
	B5-Vegetation red edge1	20 m	
	B6-Vegetation red edge2	20 m	
	B7-Vegetation red edge3	20 m	
	B8-NIR	10 m	
	B8A-Narrow NIR	20 m	
	B11-SWIR1	20 m	
	B12-SWIR2	20 m	
Vegetation index	NDVI	10 m	$NDVI = (NIR - R) / (NIR + R)$
	EVI	10 m	$EVI = 2.5 \times (NIR - R)/(NIR + 6 \times R - 7.5 \times B + 1)$
	SAVI	10 m	$SAVI = (1 + L1) \times (NIR - R)/(NIR + R + L1)$
	KNDVI	10 m	$KNDVI = ((NIR - R)/(NIR + R))/\sqrt{NIR^2 + R^2}$
	RVI	10 m	$RVI = NIR/R$
	EXR	10 m	$2.5 \times (NIR - R)/(NIR + 5 \times R + 1)$
	DVI	10 m	$DVI = NIR - R$
Climate	MAT	0.1°	*ERA5-Land*
	MAP	0.1°	
	MWS	0.1°	
Topographic	Elevation	30 m	*SRTM DEM*
	Aspect	30 m	
	Slope	30 m	

2.3.3. ML Models

Five widely used tree-based ensemble regressors, which are standard choices for remote-sensing parameter inversion, were employed to generate the FVC validation dataset: Random Forest (RF), Gradient Boosting Machine (GBM), Light Gradient Boosting Machine (LightGBM), eXtreme Gradient Boosting (XGBoost), and Categorical Boosting (CatBoost). RF is a bagging-based ensemble of decision trees that improves robustness and reduces overfitting, while GBM, XGBoost, LightGBM, and CatBoost are gradient-boosting frameworks that iteratively build decision trees to minimize a loss function; LightGBM uses histogram-based, leaf-wise growth for efficient training on large datasets, XGBoost incorporates second-order optimization and regularization, and CatBoost employs ordered boosting and symmetric trees. These models have been extensively applied to non-linear regression and the inversion of biophysical parameters such as LAI and FVC [33–37].

2.3.4. Feature Engineering and Hyperparameter Optimization

Model optimization comprised feature selection and hyperparameter tuning. RFE, a sophisticated wrapper-based technique, commenced with the establishment of feature importance rankings using machine learning estimators, proceeded through iterative deletion of the least consequential feature, and concluded with the ordering of features based on their sequential elimination [18]. Feature importance rankings from each ML model were evaluated using RMSE and R^2 to ensure selection validity. Hyperparameter optimization employed grid search with five-fold cross-validation [28]. Superior parameter configurations odel gewere determined by evaluating performance across predefined search

spaces, enhancing mneralization capability. Hyperparameter search ranges, optimization processes, and final values for all ML models are detailed in Table 3.

Table 3. Parameter grid search ranges and results for each ML model.

Model	Parameters	Scales	Step Size	Best Parameters
RF	n_neighnbors	(100, 1000)	50	800
	max_depth	(1, 20)	2	17
	min_samples_split	(1, 20)	4	5
	max_features	(1, 20)	4	1
	min_samples_leaf	(1, 20)	2	1
GBM	n_neighnbors	(100, 1000)	100	700
	max_depth	(1, 20)	2	7
	learning_rate	(0.01, 0.20)	0.02	0.05
	subsample	(0.1, 1.0)	0.2	0.9
	min_samples_split	(1, 20)	2	15
LGBM	num_iterations	(0, 1000)	50	400
	max_depth	(1, 20)	2	15
	num_leaves	(5, 20)	5	20
	learning_rate	(0.01, 0.20)	0.02	0.07
	subsample	(0.1, 1.0)	0.2	0.7
XGBoost	n_estimators	(0, 1000)	50	800
	max_depth	(1, 20)	2	7
	learning_rate	(0.01, 0.20)	0.02	0.03
	subsample	(0.1, 1.0)	0.2	0.7
	colsample_bytree	(0.1, 1.0)	0.2	0.7
CatBoost	iterations	(0, 2000)	50	2000
	learning_rate	(0.01, 0.20)	0.02	0.03
	depth	(1, 10)	2	9
	bagging_temperature	(0.1, 1)	0.2	0.1

2.3.5. Model Performance Evaluation

The R^2 and RMSE were utilized to evaluate ML model performance. Higher R^2 values (closer to 1) indicate superior model fitting accuracy, while lower RMSE values (closer to 0) reflect smaller discrepancies between inverted and measured values, denoting enhanced model performance. The computational formulas for metrics are presented as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(S_i - S_i^{'})^2}{\sum_{i=1}^{n}(S_i - \overline{S_i})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - S_i^{'})^2} \quad (6)$$

where n signifies the sample size, S_i denotes the measured values, $S_i^{'}$ represents the inverted values, and $\overline{S_i}$ is the mean of the measured values.

2.3.6. Product Validation Scenarios

FVC product accuracy was assessed using direct validation and multi-scale validation. Direct validation compares ground-referenced FVC with the corresponding product pixels to quantify accuracy at a scale close to the product spatial support. In the direct-validation scenarios, the 250 m UAV-derived plot FVC is compared with each product value extracted at the plot center, because direct validation aims to assess product performance at a scale close to the product pixel support. NDVI-SHI is computed at 250 m for the same purpose, to identify plots where sub-pixel heterogeneity would otherwise bias the point-to-pixel comparison. We used UAV-derived plot FVC at 250 m × 250 m as the ground reference, which approximates the native supports of GEOV3 (300 m) and GLASS/MuSyQ (500 m) and helps reduce scale-mismatch effects in direct comparisons. Within each 250 m plot,

30 m × 30 m subplots were surveyed to characterize fine-scale variability, and subplots with obvious internal differences were excluded from subsequent validation analyses. Residual heterogeneity at the 250 m support was further controlled using RSSE and NDVI-SHI to select more representative 250 m samples. Multi-scale validation used a 10 m FVC reference retrieved from Sentinel-2 using the selected ML model. Specifically, the 10 m reference FVC was aggregated to 300 m for GEOV3 and to 500 m for GLASS and MuSyQ to ensure consistent spatial support in the comparisons. Four validation scenarios were designed: (1) direct validation with raw 250 m ground data; (2) direct validation with 250 m data filtered by RSSE and NDVI-SHI; (3) multi-scale validation with high-precision 10 m FVC reference; and (4) multi-scale validation with high-precision 10 m FVC reference filtered by RSSE and NDVI-SHI.

3. Results

3.1. Evaluation of Spatial Representativeness for Ground Measurements

Spatial representativeness constraints were progressively tightened (with RSSE thresholds sequentially set to no limit, 0.30, 0.20, and 0.10) to determine the optimal balance between validation accuracy and sample size (Figure 4). The analysis revealed that overall correlation increased while RMSE decreased as constraints became stricter. The most significant accuracy gain was observed when the RSSE threshold was tightened from 0.30 to 0.20; in contrast, further tightening to 0.10 yielded only marginal improvements while causing a sharp reduction in sample size. Consequently, using RSSE < 0.20 as the baseline, an NDVI-SHI filter was introduced. Reducing the NDVI-SHI threshold from 0.20 to 0.15 provided further benefits in accuracy. However, a threshold of 0.10 reached a performance plateau with substantial loss of valid samples. Notably, under this strictest threshold (NDVI-SHI < 0.10), the validation accuracy for the MuSyQ product actually decreased slightly compared to the 0.15 threshold. Therefore, balancing the trade-off between accuracy gains and data retention, the final selection criteria were determined as RSSE < 0.20 and NDVI-SHI < 0.15. Under these criteria, the number of high-representativity samples was: 62 of 113 in 2016 (54.9%), 46 of 119 in 2017 (38.7%), 40 of 95 in 2018 (42.1%), and 138 of 296 in 2019 (46.6%).

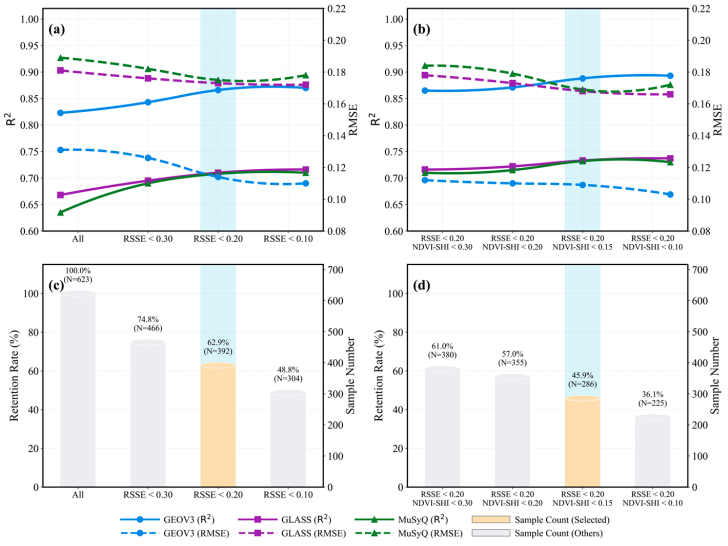


Figure 4. Threshold screening results for representativeness metrics. (**a**) Product validation results under RSSE thresholds. (**b**) Product validation results under NDVI-SHI thresholds (RSSE fixed at 0.20). (**c**) Retained plots under RSSE thresholds. (**d**) Retained plots under NDVI-SHI thresholds (RSSE fixed at 0.20).

3.2. Feature Selection and Performance Evaluation of ML Models

Across the five RF-ML models, performance showed a consistent pattern: as the number of features increased, R^2 rose, and RMSE decreased, then both metrics plateaued (Figure 5). RF-RFE improved strongly from 1 to 10 features, with only marginal gains thereafter. GBM-RFE and XGBoost-RFE stabilized at 7 features, and LightGBM-RFE and CatBoost-RFE at 8 features. Accordingly, the optimal feature subsets were: RF-RFE, top 10 features (EVI, RVI, Elevation, B5, B4, MWS, MAT, MAP, B12, B11); GBM-RFE, top 7 (EVI, RVI, B4, MAT, Elevation, B5, MWS); LightGBM-RFE, top 8 (MWS, Elevation, MAT, EVI, MAP, B12, B4, B2); XGBoost-RFE, top 7 (EVI, MWS, MAT, MAP, KNDVI, Elevation, B5); CatBoost-RFE, top 8 (EVI, Elevation, RVI, MWS, MAT, MAP, B12, B5).

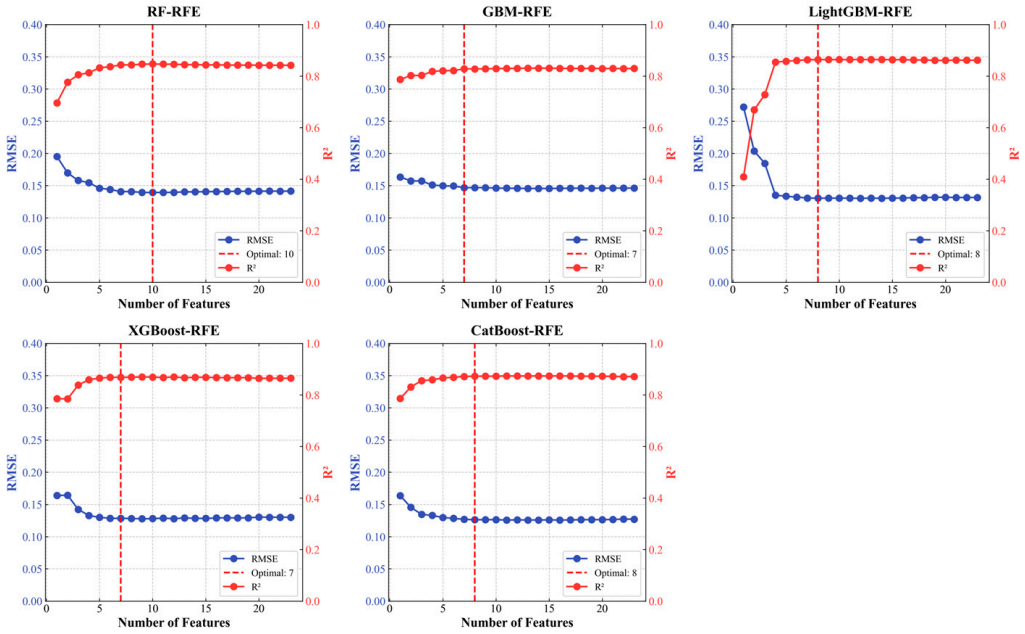


Figure 5. Relationship between RFE model performance (R^2, RMSE) and the number of features using five ML models as estimators.

The contributions of different features varied markedly across models (Figure 6). Vegetation indices dominated, with EVI showing the highest importance in RF-RFE (0.413), GBM-RFE (0.425), XGBoost-RFE (0.392), and CatBoost-RFE (0.278), confirming its key predictive role; in LightGBM-RFE, its importance was much lower (0.134, ranked 4th). Reflectance bands (B2, B4, B5, B11, B12) consistently had low importance (0.053–0.123) and acted mainly as auxiliary predictors. Climate variables showed model-dependent patterns: in LightGBM-RFE, MWS (0.165), MAT (0.141), and MAP (0.122) all ranked within the top five, whereas in XGBoost-RFE, CatBoost-RFE, and GBM-RFE they were of moderate but secondary importance relative to EVI. For topography, Elevation was consistently informative, ranking 2nd in LightGBM-RFE (0.157) and CatBoost-RFE (0.120), but served as an auxiliary feature in the other models (0.064–0.091).

Following feature selection and parameter optimization, all ML models exhibited enhanced performance, characterized by universally elevated R^2 values and reduced RMSE. Before optimization, the model accuracy ranking was as follows (Figure 7): CatBoost (R^2 = 0.872, RMSE = 0.126) > XGBoost (R^2 = 0.863, RMSE = 0.131) = LightGBM (R^2 = 0.863, RMSE = 0.131) > RF (R^2 = 0.844, RMSE = 0.140) > GBM (R^2 = 0.829, RMSE = 0.146). After optimization (Figure 8), the model performance hierarchy evolved to: CatBoost (R^2 = 0.880, RMSE = 0.122) > XGBoost (R^2 = 0.879, RMSE = 0.123) > GBM (R^2 = 0.875, RMSE = 0.125) > LightGBM (R^2 = 0.870, RMSE = 0.127) > RF (R^2 = 0.867,

RMSE = 0.129). Based on these evaluations, the selected superior algorithm was deployed on preprocessed Sentinel-2 imagery to batch generate a 10 m FVC reference dataset. Following feature construction and model inference, pixel-level FVC rasters were produced, and quality control was implemented using RSSE and NDVI-SHI. This 10 m reference dataset served as the high-resolution benchmark for the subsequent multi-scale validation (Scenarios 3 and 4).

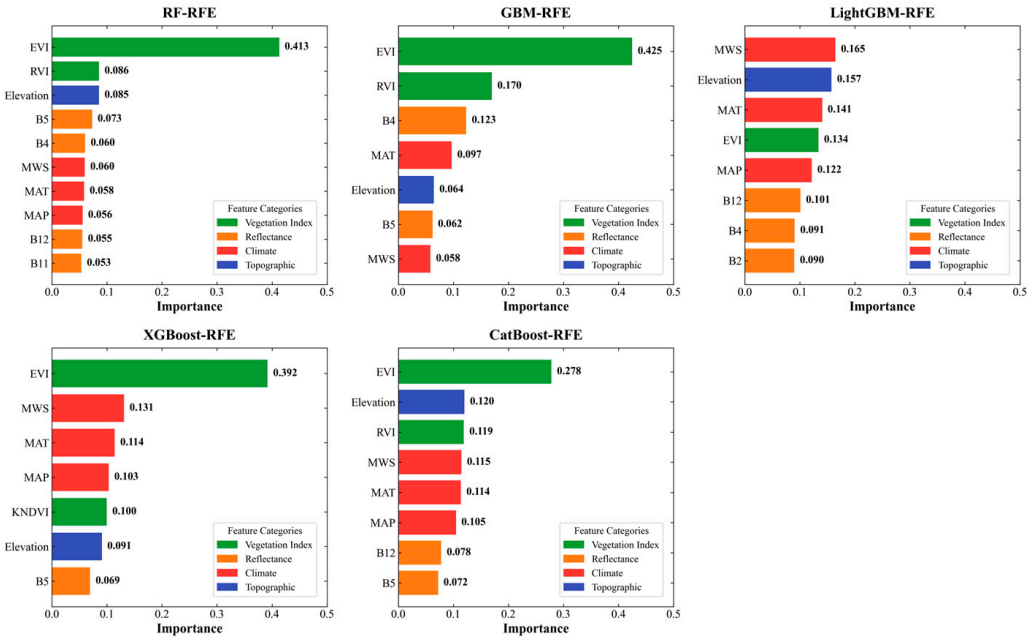


Figure 6. Feature importance of selected features by RFE models using five ML models as estimators.

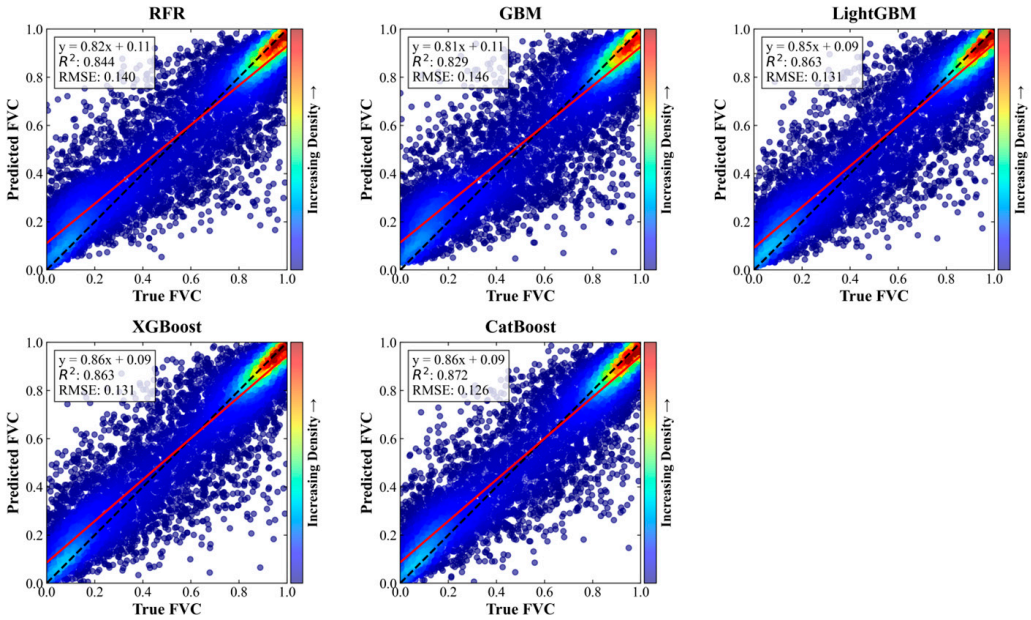


Figure 7. Linear regression plot of FVC inversion results by ML models before feature selection and parameter tuning.

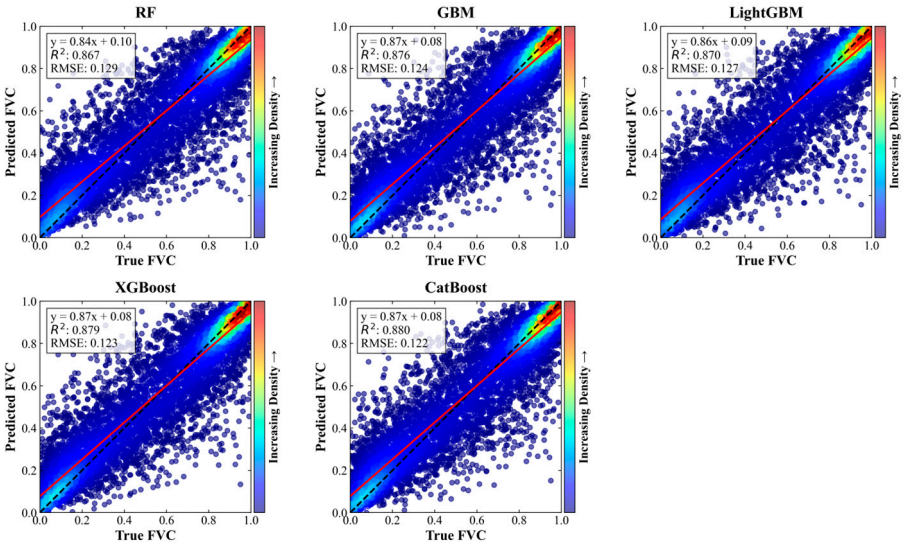


Figure 8. Linear regression plot of FVC inversion results by ML models after feature selection and parameter tuning.

3.3. Spatial Consistency and Accuracy Comparison of GEOV3, GLASS, and MuSyQ FVC Products over the QTP

The difference maps of the three FVC products across the QTP show that the plateau interior and contiguous alpine grassland areas are dominated by near-zero light tones, indicating generally small differences. In contrast, contiguous belts of positive and negative differences appear in the mountains and valleys along the eastern and southeastern margins. The two GEOV3-based difference layers (Figure 9a,b) exhibit highly consistent spatial patterns, with a continuous positive difference belt along the southeastern margin and scattered negative patches locally; both the magnitude and areal extent are notably greater than in Figure 9c. The overall differences between GLASS and MuSyQ are the smallest, with most regions near zero and only narrow or scattered difference zones in the southeast and parts of the east.

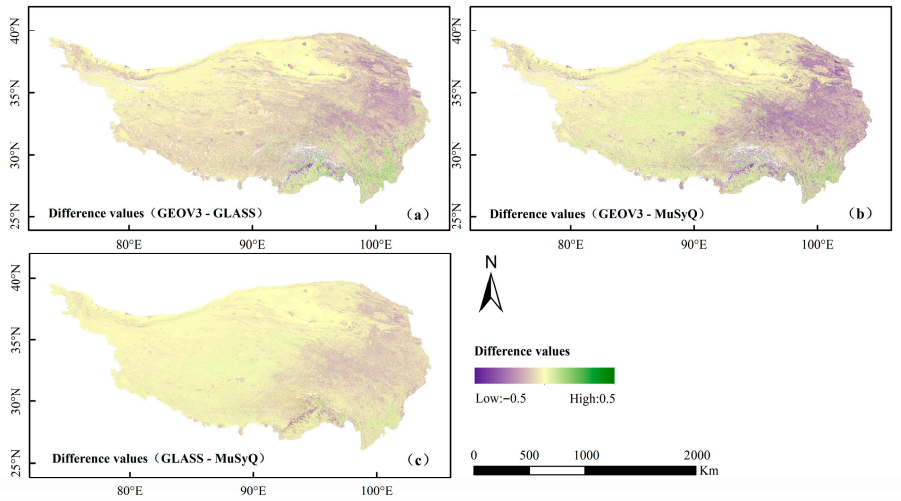


Figure 9. Spatial difference maps of three FVC products over the QTP. (**a**) GEOV3-GLASS; (**b**) GEOV3-MuSyQ; (**c**) GLASS-MuSyQ.

Scenario-based validation results further quantify these differences and reveal the effects of representativeness screening and scale matching (Figure 10). Under direct validation with raw 250 m UAV plot data (Scenario 1), GEOV3 showed the highest agreement and the

most concentrated distribution around the 1:1 line (R^2 = 0.801–0.843; RMSE = 0.122–0.140; Figure 10a), with slight underestimation at low cover (FVC < 0.2) and mild saturation or overestimation at high cover (FVC > 0.8). GLASS exhibited systematic underestimation across the full FVC range (R^2 = 0.648–0.675; RMSE = 0.160–0.196; Figure 10e). MuSyQ also underestimated, with larger dispersion and a stronger low-cover bias (R^2 = 0.608–0.675; RMSE = 0.173–0.216; Figure 10i). These results indicate that, without screening, heterogeneity and point-to-pixel mismatch can substantially inflate validation uncertainty and obscure product performance.

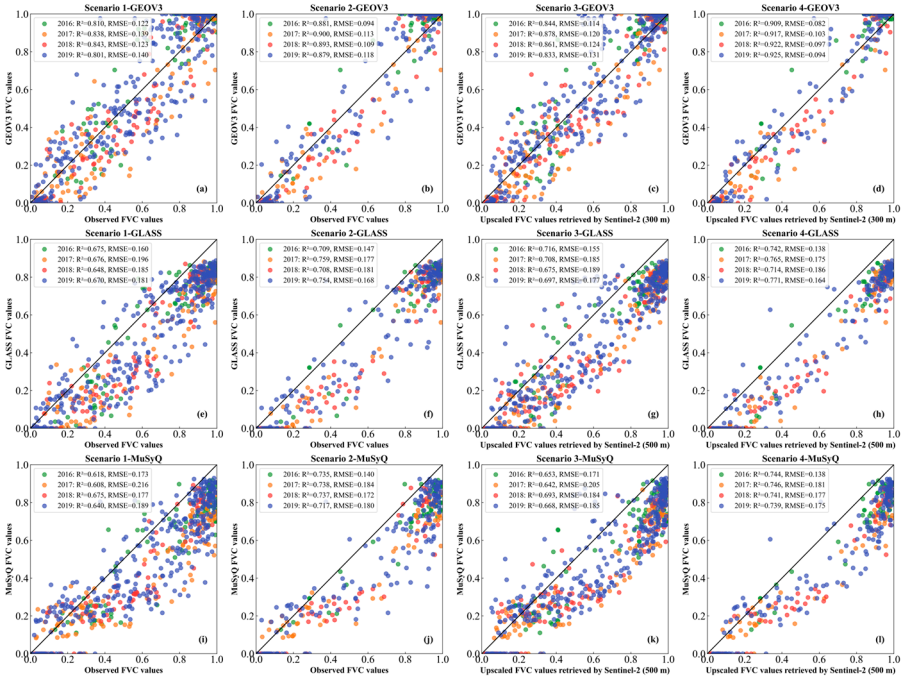


Figure 10. Scatter plots of FVC product validation results under different scenarios, where panels (**a**–**d**), (**e**–**h**), and (**i**–**l**) correspond to GEOV3, GLASS, and MuSyQ FVC products under Scenarios 1–4, respectively.

After representativeness screening, Scenario 2 conducted direct validation using the filtered 250 m UAV plot samples. Under this direct-validation setting, all products moved closer to the 1:1 line and the scatter became less dispersed. GEOV3 improved to R^2 = 0.879–0.900 and RMSE = 0.094–0.118 (Figure 10b). GLASS improved to R^2 = 0.708–0.754 and RMSE = 0.147–0.181, but underestimation remained evident (Figure 10f). MuSyQ improved to R^2 = 0.717–0.738 and RMSE = 0.140–0.184 (Figure 10j), with underestimation persisting. The consistent improvement from Scenario 1 to Scenario 2 demonstrates that representativeness screening effectively reduces mismatch-related error and stabilizes the evaluation, while the overall ranking among products remains unchanged.

Multi-scale validation using the 10 m reference aggregated to the product grids (Scenario 3) confirmed GEOV3's leading performance (R^2 = 0.833–0.878; RMSE = 0.114–0.131; Figure 10c). GLASS remained higher than MuSyQ but continued to underestimate (R^2 = 0.675–0.716; RMSE = 0.155–0.189; Figure 10g), whereas MuSyQ showed lower agreement and larger errors (R^2 = 0.642–0.693; RMSE = 0.171–0.205; Figure 10k). Compared with Scenario 1, the multi-scale approach reduces part of the scale mismatch by using an intermediate high-resolution reference, but without representativeness screening, residual heterogeneity still propagates into the aggregated comparison and limits accuracy gains.

When multi-scale validation is combined with representativeness screening (Scenario 4), all products reach their best performance and the comparisons are most reliable. GEOV3 is closest to the 1:1 line (R^2 = 0.909–0.925; RMSE = 0.082–0.103; Figure 10d). GLASS improves further (R^2 = 0.742–0.771; RMSE = 0.138–0.175; Figure 10h) but still lies mostly below the 1:1 line, indicating persistent underestimation. MuSyQ reaches its optimal range (R^2 = 0.739–0.746; RMSE = 0.138–0.181; Figure 10l), with underestimation weaker than in earlier scenarios but still present. Overall, cross-scenario comparisons show that representativeness screening and multi-scale alignment both contribute to reducing validation uncertainty, and their combination provides the most robust basis for product assessment over the highly heterogeneous QTP. At the same time, the stable ranking across scenarios, with GEOV3 consistently outperforming GLASS and MuSyQ and the latter two retaining an underestimation tendency, suggests that part of the differences reflects product-specific characteristics beyond scale mismatch alone.

4. Discussion

4.1. Impact of Ground Measurements Spatial Representativeness on FVC Product Validation

Spatial representativeness of ground measurements is central to FVC product validation but is limited by scale mismatch between point observations and coarse satellite pixels. FVC products depict surfaces at hundreds of meters to kilometers, whereas field surveys sample discrete plots, creating coverage gaps that are most problematic in heterogeneous terrain. For QTP alpine meadows, heterogeneity is mainly expressed as patchy vegetation-bare soil mixtures and terrain-driven illumination contrasts within a coarse pixel. This violates the implicit homogeneity assumption behind direct point-to-pixel validation and can inflate apparent product errors. Vegetation spatial variability increases nonlinearly with observation scale, and plot layout and density distribution of quadrats bias validation [38]. Multiscale heterogeneity also makes the ratio of vegetation spatial autocorrelation length to pixel size a key control on upscaling accuracy. These scale effects inflate validation errors and mask true accuracy, especially in ecotones and fragmented landscapes [39]. Accordingly, we used RSSE and NDVI-SHI to screen out heterogeneous plots and evaluate products under more representative conditions (RSSE < 0.2; NDVI-SHI < 0.15). Applied to QTP ground data from 2016 to 2019, it removed on average 54.1% of low-representativeness plots and substantially improved validation reliability for FVC products.

4.2. Feature-Model Interactions and Optimization

Feature selection and hyperparameter optimization jointly determine the effectiveness of FVC inversion models. In this heterogeneous alpine setting, vegetation indices dominated across models, and EVI was consistently among the most informative predictors, which aligns with its spectral design that reduces background and atmospheric effects and maintains sensitivity at higher cover [40]. In contrast, NDVI tends to saturate at high cover and is strongly collinear with EVI; within tree-based models, shared information among correlated indices is often absorbed by a dominant feature, so a lower NDVI importance mainly reflects redundancy rather than irrelevance [41]. A similar rationale applies to single reflectance bands (e.g., B4), where terrain shadow and BRDF variability can increase noise relative to composite indices. Feature importance patterns also differed among boosting models. In LightGBM, EVI ranked below Mean Wind Speed (MWS), which is mainly related to gain-based importance and leaf-wise growth: variables that yield large loss reductions early can accumulate higher gain, while the contributions of correlated spectral predictors can be distributed across features, lowering the apparent gain for any single index. Climatic variables can also be spatially smoother and less affected by pixel-level back-

ground and terrain effects than spectral predictors, making them stable split candidates in complex scenes [35].

Model sensitivity to optimization reflected both learning mechanisms and the error structure of alpine FVC inversion. The largest gain observed for GBM suggests that its baseline configuration did not fully exploit nonlinear relationships; after selecting key predictors and tuning boosting parameters, the model better separated vegetation signals from soil and illumination-related background effects [34]. CatBoost and XGBoost improved less after tuning, consistent with their stronger regularization and robust handling of feature interactions; once the main mapping between spectral indices and FVC is captured, further gains are constrained by mixed pixels, topographic shading, and background heterogeneity typical of the QTP [28]. LightGBM is efficient but can under-represent weaker yet physically meaningful spectral indicators when information is distributed across correlated features. RF remained stable due to averaging across many trees, but without sequential error-correction it is less effective at modeling subtle, interaction-driven variability in fragmented landscapes. Ancillary variables from SRTM and ERA-5 were included to represent terrain and background climatic conditions. To build a unified feature stack, these layers were resampled to 10 m to align with the Sentinel-2 grid for consistent extraction, rather than to imply intrinsic 10 m information in the ancillary data. This alignment supports pixel-wise co-registration between optical, topographic, and climatic predictors, but it may introduce uncertainty because interpolation can smooth DEM extremes and downscaling a 0.1° climate field to 10 m does not add physically meaningful fine-scale variability. In this study, climatic variables generally acted as auxiliary predictors relative to spectral indices, reducing the likelihood that resampling effects dominate the retrieval [30]. Future work will explore scale-aware integration, such as using ancillary variables at their native resolution or an intermediate support.

Comparative analysis showed clear advantages of the Sentinel 2-derived 10 m FVC dataset over GEOV3, GLASS, and MuSyQ products (Figure 11). The CatBoost optimized 10 m FVC provided finer spatial detail, resolving riparian patches, heterogeneous grassland mosaics, and sparse vegetation to bare soil transitions. Coarser GEOV3, GLASS, and MuSyQ exhibited spatial averaging that underestimated local variability, missed transitional features, and blurred boundaries. The 10 m reference alleviated scale mismatch and agreed closely with ground measurements, especially in complex, fragmented terrain where conventional products could not resolve fine-scale patterns, thereby offering a reliable benchmark for validation.

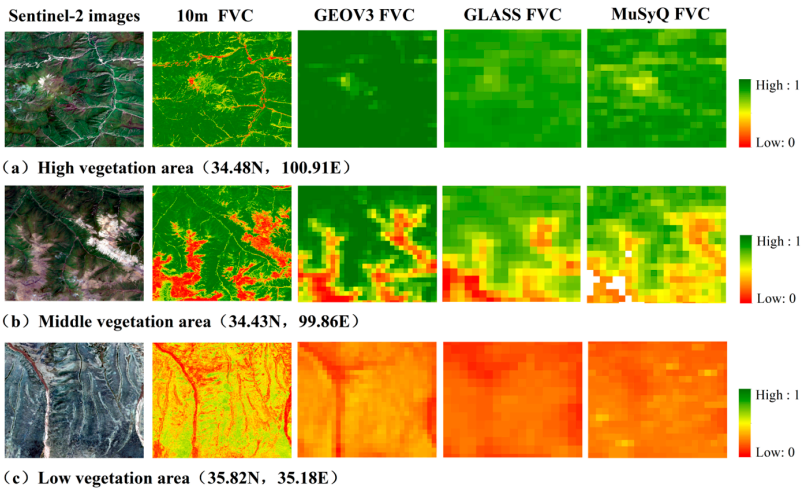


Figure 11. Comparison plot of the 10 m FVC product validation dataset with GEOV3, GLASS, and MuSyQ FVC products.

4.3. Validation Analysis of GEOV3, GLASS and MuSyQ FVC Products

Direct validation and multiscale validation characterize, respectively, absolute accuracy near the native scale and robustness after explicit control of scale effects. Across four scenarios, GEOV3 outperformed GLASS and MuSyQ, and this ranking remained after representativeness screening and multiscale alignment (Figure 10). This indicates that inter-product differences are driven by multiple factors acting together, rather than by nominal resolution or algorithm category alone [42]. In this study, the "scale issue" mainly refers to point-to-pixel mismatch and within-pixel heterogeneity under rugged terrain and fragmented vegetation, which cannot be fully represented by pixel size alone. Even after reducing point-to-pixel mismatch, stable bias patterns persisted. GEOV3 showed slight saturation and overestimation at high coverage, whereas GLASS and MuSyQ underestimated from low coverage through a broader coverage range. Under stricter screening combined with multiscale alignment, MuSyQ converged toward GLASS, indicating some calibratability and practical utility [25]. Representativeness screening and multiscale validation mitigated errors from scale mismatch and surface heterogeneity and improved reliability for GEOV3 and GLASS [43]. This accords with the consensus that ignoring heterogeneity and scale effects is a primary source of bias. In Scenario 1, unscreened heterogeneity prevented ground measurements from representing pixel-level conditions, amplifying mismatch errors and yielding dispersed outcomes (Figure 10a,e,i). Scenario 2 applied RSSE and NDVI-SHI to remove highly heterogeneous plots and markedly improved reliability (Figure 10b,f,j). Scenarios 3 and 4 integrated 10 m Sentinel-2 data to link ground observations to high-resolution imagery and then to medium and low resolution products, further reducing scale mismatch. Scenario 3 improved accuracy relative to Scenario 1 by better characterizing intra pixel structure through aggregation (Figure 10c,g,k). Scenario 4 combined screening with multiscale validation, removed highly heterogeneous plots, and achieved the highest reliability (Figure 10e,h,l).

GEOV3 consistently performed better over the QTP, which is likely related to both its retrieval design and spatial support. GEOV3 provides FVC at 300 m and uses a neural-network-based retrieval [24], which is well suited for capturing nonlinear reflectance–cover relationships over heterogeneous surfaces, and its finer support than 500 m products helps mitigate mixed-pixel effects in fragmented landscapes [44]. This interpretation is consistent with previous evaluations reporting moderate-to-high accuracy for GEOV FVC products across biomes and with regional multi-scale assessments over QTP alpine grasslands showing better performance of GEOV3 relative to MuSyQ. By contrast, the persistent underestimation of GLASS and MuSyQ can be attributed to algorithmic and scale-related factors. GLASS relies on MARS-type regression that can shrink predictions toward the mean, which may lead to underestimation on endpoint-rich, heterogeneous surfaces, while MuSyQ is more sensitive to NDVI saturation and background brightness under strong soil background and residual snow/shadow [42]. In addition, their predominant 500 m spatial support increases mixed-pixel effects [44], and conservative masking/compositing together with terrain and BRDF effects under frequent cloud/snow can further bias seasonal composites toward lower values.

Compared with similar regional and global validation studies, the accuracy levels obtained in this study are broadly consistent with the typical ranges reported for FVC products. Evaluations of GEOV1/GEOV3 and related FVC products generally indicate moderate to high accuracy across diverse biomes [45,46]. Over China, GLASS and MuSyQ FVC products usually exhibit uncertainties in grassland and heterogeneous areas that are comparable to or slightly larger than those of other mainstream products [42]. Our best-performing scenario (Scenario 4) places GEOV3 over QTP alpine grasslands at the upper end of these qualitative accuracy levels, while GLASS and MuSyQ show error

patterns similar to those previously reported for complex surfaces. This suggests that, even under the challenging high-elevation and highly heterogeneous conditions of the QTP, the proposed validation framework can achieve product performance that is generally in line with domestic and international benchmarks.

4.4. Limitations and Prospects

This study advances FVC product validation over QTP alpine grasslands, but has several limitations. First, observations are limited to July–August, so the results mainly represent peak-season conditions and do not fully characterize performance during green-up or senescence; future work will extend validation to the full growing season with phenology-stratified sampling. Second, the UAV-derived FVC relies on EGI thresholding. Even with independent processing by multiple analysts and adjudication, threshold selection remains partly subjective, especially in sparse vegetation, shadowed areas, and soil–vegetation transition zones. This may introduce uncertainty in the UAV reference and can propagate into the reported agreement between satellite products and UAV observations. Future improvements will include more objective image classification approaches, such as deep learning-based semantic segmentation trained with manually labeled samples, and the use of ortho-mosaics with consistent radiometric normalization to reduce scene-to-scene threshold variability. Third, averaging sixteen 30 m × 30 m subplots to approximate 250 m resolution may still miss local heterogeneity in fragmented landscapes; future campaigns will use full-coverage UAV ortho-mosaics to better represent coarse pixels. In addition, while representativeness screening is applied at the 250 m plot support, within-pixel heterogeneity at the native product scales is not explicitly quantified. Future work will quantify within-pixel heterogeneity at 300 m/500 m using independent high-resolution reference maps and corresponding variability metrics, and then stratify product errors under different heterogeneity conditions. Building on these improvements, future work will extend this framework to produce and assess plateau-wide high-resolution FVC datasets with consistent quality.

5. Conclusions

This study proposes a multidimensional FVC product evaluation framework for the QTP that integrates spatial representativeness screening, ML-based reference construction, and scenario-based validation across scales. The framework is designed to reduce validation uncertainty caused by surface heterogeneity and point-to-pixel mismatch and to enable more reliable comparisons among GEOV3, GLASS, and MuSyQ. Four validation scenarios were implemented to quantify the effects of representativeness and scale matching: Scenario 1 performed direct validation using raw 250 m UAV plot FVC; Scenario 2 performed direct validation using representativeness-screened 250 m samples based on RSSE and NDVI-SHI; Scenario 3 performed multi-scale validation using an aggregated 10 m reference; and Scenario 4 performed multi-scale validation using the aggregated 10 m reference after representativeness screening based on RSSE and NDVI-SHI.

Across all scenarios, GEOV3 consistently outperformed GLASS and MuSyQ. Representativeness screening improved the robustness of direct validation by excluding low-representativeness plots. For GEOV3, R^2 increased from 0.801–0.843 to 0.879–0.900 and RMSE decreased from 0.122–0.140 to 0.094–0.118; GLASS improved from R^2 = 0.648–0.675 and RMSE = 0.160–0.196 to R^2 = 0.708–0.754 and RMSE = 0.147–0.181; MuSyQ improved from R^2 = 0.608–0.675 and RMSE = 0.173–0.216 to R^2 = 0.717–0.738 and RMSE = 0.140–0.184. A 10 m FVC reference was generated from Sentinel-2 using optimized ML models, where CatBoost achieved the best retrieval performance (R^2 = 0.880, RMSE = 0.122), and the reference was used for multi-scale validation after aggregation to product grids. Under the

most robust setting (Scenario 4), GEOV3 achieved R^2 = 0.909–0.925 and RMSE = 0.082–0.103, while GLASS achieved R^2 = 0.742–0.771 and RMSE = 0.138–0.175 and MuSyQ achieved R^2 = 0.739–0.746 and RMSE = 0.138–0.181, with overall underestimation for GLASS and MuSyQ. Future work will extend validation to full growing-season periods, further reduce uncertainty in UAV-based reference extraction, and develop long-term multi-scale benchmark datasets to support plateau vegetation monitoring and product improvement.

Author Contributions: J.L., J.C. and X.C.: Conceptualization. J.L., J.C. and X.H.: data curation. J.Y., X.C. and Q.C.: formal analysis. J.L., J.C. and X.L.: methodology. J.Y., Q.C., H.Y. and X.H.: supervision. X.L., J.L. and J.C.: validation. J.L., Q.C. and H.Y.: writing—original draft preparation. J.L., Q.C., J.Y. and X.C.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by grants from the National Natural Science Foundation of China (42361023), Natural Science Foundation of Guangxi Province (2024GXNSFAA010268), Guangxi Key Laboratory of Spatial Information and Geomatics under Grant (21-238-21-31), BaGui Young Outstanding Talents Cultivation Program of the provincial government of Guangxi.

Data Availability Statement: The data presented in this study are not available due to privacy restrictions, but are available from the corresponding author on reasonable request.

Acknowledgments: The authors sincerely thank the editors and the anonymous reviewers for their constructive feedback and valuable suggestions, which have greatly contributed to the improvement of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gou, Y.; Hao, L.; Huang, Y.; Jin, X.; Zhang, A.; Ma, H. The Changes in Grassland Animal Husbandry and Herdsmen's Life in the Qinghai Pastoral Area of China Based on the Perspective of Changes in the Grassland Property Rights System. *Sustainability* **2025**, *17*, 1262. [CrossRef]
2. Liu, Q.; Zhang, S.; Ma, Z.; Zhou, X.; Wang, T. A novel approach to retrieving the surface soil freeze/thaw state in the Qinghai-Tibetan Plateau using the seasonality of CYGNSS time series. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *137*, 104428. [CrossRef]
3. Hu, T.; Cao, M.; Zhao, X.; Liu, X.; Liu, Z.; Liu, L.; Huang, Z.; Tao, S.; Tang, Z.; Guo, Y. High-resolution mapping of grassland canopy cover in China through the integration of extensive drone imagery and satellite data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *218*, 69–83. [CrossRef]
4. Bai, Y.; Cotrufo, M.F. Grassland soil carbon sequestration: Current understanding, challenges, and solutions. *Science* **2022**, *377*, 603–608. [CrossRef]
5. Hao, X.; Yang, J.; Dong, S.; He, F.; Zhang, Y. The influence of grazing intensity on soil organic carbon storage in grassland of China: A meta-analysis. *Sci. Total Environ.* **2024**, *924*, 171439. [CrossRef]
6. He, B.; Zhu, W.; Zhao, C.; Xie, Z.; Zhuang, H. A novel index for directly indicating fractional vegetation cover based on spectral differences between vegetation and soil. *Remote Sens. Environ.* **2025**, *331*, 115056. [CrossRef]
7. Song, D.-X.; Wang, Z.; He, T.; Wang, H.; Liang, S. Estimation and validation of 30 m fractional vegetation cover over China through integrated use of Landsat 8 and Gaofen 2 data. *Sci. Remote Sens.* **2022**, *6*, 100058. [CrossRef]
8. An, R.; Jin, H.; Zhao, H.; Wei, D.; Zhao, W.; Wang, X. Productivity experienced a more rapid enhancement trend than greenness across the Tibetan Plateau. *Sci. Total Environ.* **2024**, *954*, 176666. [CrossRef]
9. Zeng, Q.; Wu, X.; Tang, R.; Wang, J.; Lin, X.; Wen, J.; Xiao, Q. Quantification of uncertainty caused by geometric location mismatch in the validation of TROPOMI solar-induced chlorophyll fluorescence product. *Int. J. Digit. Earth* **2024**, *17*, 2330688. [CrossRef]
10. Wu, X.; Wen, J.; Tang, R.; Wang, J.; Zeng, Q.; Li, Z.; You, D.; Lin, X.; Gong, B.; Xiao, Q. Quantification of the uncertainty in multiscale validation of coarse-resolution satellite albedo products: A study based on airborne CASI data. *Remote Sens. Environ.* **2023**, *287*, 113465. [CrossRef]
11. Lv, T.; Zhou, X.; Tao, Z.; Sun, X.; Wang, J.; Li, R.; Xie, F. Remote sensing-guided spatial sampling strategy over heterogeneous surface ground for validation of vegetation indices products with medium and high spatial resolution. *Remote Sens.* **2021**, *13*, 2674. [CrossRef]

12. Geng, L.; Che, T.; Tan, J.; Wang, H.; Wang, X. Evaluation of Sentinel-2 and MODIS NDVI Accuracy in Heterogeneous Agricultural Landscapes Using UAV Observations: A Case Study in the Heihe River Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 14928–14941. [CrossRef]
13. Xie, J.; Zhang, H.; Pu, J.; Yan, Z.; Cai, Y.; Cui, L.; Cao, H. Multi-Scale BRDF Modeling and Scale Effect Analysis Using UAV Multi-Angle Remote Sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 23090–23106. [CrossRef]
14. Otto, P.; Fassò, A.; Maranzano, P. A review of regularised estimation methods and cross-validation in spatiotemporal statistics. *Stat. Surv.* **2024**, *18*, 299–340. [CrossRef]
15. Zhou, P.; Geng, L.; Li, J. Enhancing Uncertainty Assessment of LAI Remote Sensing Products Through Multiple Collocation Method and Error Separation Techniques. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 22573–22588. [CrossRef]
16. Wu, X.; Xiao, Q.; Wen, J.; You, D. Advances and challenges in the validation of remote sensing albedo products. *J. Remote Sens.* **2019**, *23*, 11–23.
17. Fan, X.; Gao, P.; Zhang, M.; Cang, H.; Zhang, L.; Zhang, Z.; Wang, J.; Lv, X.; Zhang, Q.; Ma, L. The fusion of vegetation indices increases the accuracy of cotton leaf area prediction. *Front. Plant Sci.* **2024**, *15*, 1357193. [CrossRef]
18. Lin, X.; Chen, J.; Lou, P.; Yi, S.; Qin, Y.; You, H.; Han, X. Improving the estimation of alpine grassland fractional vegetation cover using optimized algorithms and multi-dimensional features. *Plant Methods* **2021**, *17*, 1–18. [CrossRef] [PubMed]
19. Yi, S. FragMAP: A tool for long-term and cooperative monitoring and analysis of small-scale habitat fragmentation using an unmanned aerial vehicle. *Int. J. Remote Sens.* **2017**, *38*, 2686–2697. [CrossRef]
20. Yi, S.; Chen, J.; Qin, Y.; Xu, G. The burying and grazing effects of plateau pika on alpine grassland are small: A pilot study in a semiarid basin on the Qinghai-Tibet Plateau. *Biogeosciences* **2016**, *13*, 6273–6284. [CrossRef]
21. Lin, X.; Chen, J.; Wu, T.; Yi, S.; Chen, J.; Han, X. Time-series simulation of alpine grassland cover using transferable stacking deep learning and multisource remote sensing data in the Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *131*, 103964. [CrossRef]
22. Moraes, D.; Barbosa, B.; Costa, H.; Moreira, F.D.; Benevides, P.; Caetano, M.; Campagnolo, M. Continuous forest loss monitoring in a dynamic landscape of Central Portugal with Sentinel-2 data. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *130*, 103913. [CrossRef]
23. Xu, J.; Liu, Z. Enhanced all-weather precipitable water vapor retrieval from MODIS near-infrared bands using machine learning. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *114*, 103050. [CrossRef]
24. Verger, A.; Sánchez-Zapero, J.; Weiss, M.; Descals, A.; Camacho, F.; Lacaze, R.; Baret, F. GEOV2: Improved smoothed and gap filled time series of LAI, FAPAR and FCover 1 km Copernicus Global Land products. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *123*, 103479. [CrossRef]
25. He, Q.; Lu, H.; Yang, K.; Zhao, L.; Zou, M. Benefit of incorporating GLASS remote sensing vegetation products in improving Noah-MP land surface temperature simulations on the Tibetan Plateau. *Sci. Remote Sens.* **2024**, *9*, 100115. [CrossRef]
26. Xu, B.; Li, J.; Liu, Q.; Xin, X.; Zeng, Y.; Yin, G. Review of methods for evaluating representativeness of ground station observations. *J. Remote Sens.* **2015**, *19*, 703–718. [CrossRef]
27. Xu, B.; Li, J.; Liu, Q.; Huete, A.R.; Yu, Q.; Zeng, Y.; Yin, G.; Zhao, J.; Yang, L. Evaluating spatial representativeness of station observations for remotely sensed leaf area index products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3267–3282. [CrossRef]
28. Li, X.; Chen, J.; Chen, Z.; Lan, Y.; Ling, M.; Huang, Q.; Li, H.; Han, X.; Yi, S. Explainable machine learning-based fractional vegetation cover inversion and performance optimization—A case study of an alpine grassland on the Qinghai-Tibet plateau. *Ecol. Inform.* **2024**, *82*, 102768. [CrossRef]
29. Anees, S.A.; Mehmood, K.; Rehman, A.; Rehman, N.U.; Muhammad, S.; Shahzad, F.; Hussain, K.; Luo, M.; Alarfaj, A.A.; Alharbi, S.A. Unveiling fractional vegetation cover dynamics: A spatiotemporal analysis using modis ndvi and machine learning. *Environ. Sustain. Indic.* **2024**, *24*, 100485. [CrossRef]
30. Zhong, G.; Chen, J.; Huang, R.; Yi, S.; Qin, Y.; You, H.; Han, X.; Zhou, G. High spatial resolution fractional vegetation coverage inversion based on UAV and Sentinel-2 data: A case study of alpine grassland. *Remote Sens.* **2023**, *15*, 4266. [CrossRef]
31. Shi, X.; Yang, D.; Zhou, S.; Li, H.; Zeng, S.; Yin, C.; Yang, M. Analysis of spatiotemporal variation characteristics and influencing factors of grassland vegetation coverage in the Qinghai–Tibet Plateau from 2000 to 2023 based on MODIS data. *Land* **2024**, *13*, 2127. [CrossRef]
32. Xu, Z.; Shen, X.; Ge, S.; Sun, Q.; Yang, Y.; Cao, L. An advanced TSMK-FVC approach combined with Landsat 5/8 imagery for assessing the long-term effects of terrain and climate on vegetation growth. *Front. Plant Sci.* **2024**, *15*, 1363690. [CrossRef]
33. Desai, S.; Ouarda, T.B. Regional hydrological frequency analysis at ungauged sites with random forest regression. *J. Hydrol.* **2021**, *594*, 125861. [CrossRef]
34. Konstantinov, A.V.; Utkin, L.V. Interpretable machine learning with an ensemble of gradient boosting machines. *Knowl.-Based Syst.* **2021**, *222*, 106993. [CrossRef]

35. Zhou, S.; Zhang, D.; Wang, M.; Liu, Z.; Gan, W.; Zhao, Z.; Xue, S.; Müller, B.; Zhou, M.; Ni, X. Risk-driven composition decoupling analysis for urban flooding prediction in high-density urban areas using Bayesian-Optimized LightGBM. *J. Clean. Prod.* **2024**, *457*, 142286. [CrossRef]
36. Zhang, J.; Ma, X.; Zhang, J.; Sun, D.; Zhou, X.; Mi, C.; Wen, H. Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model. *J. Environ. Manag.* **2023**, *332*, 117357. [CrossRef] [PubMed]
37. Guo, Z.; Wang, X.; Ge, L. Classification prediction model of indoor PM2.5 concentration using CatBoost algorithm. *Front. Built Environ.* **2023**, *9*, 1207193. [CrossRef]
38. Wu, X.; Wen, J.; Xiao, Q.; You, D.; Lin, X.; Wu, S.; Zhong, S. Impacts and Contributors of Representativeness Errors of In Situ Albedo Measurements for the Validation of Remote Sensing Products. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9740–9755. [CrossRef]
39. Tang, R.; Wu, X.; Wang, J.; Zeng, Q.; Li, Z.; Wen, J. Evaluation of several recently developed sampling strategies within the coarse pixel scale for validation of coarse-resolution satellite albedo products. *Int. J. Digit. Earth* **2022**, *15*, 2319–2334. [CrossRef]
40. Shi, S.; Wang, P.; Zhan, X.; Han, J.; Guo, M.; Wang, F. Warming and increasing precipitation induced greening on the northern Qinghai-Tibet Plateau. *Catena* **2023**, *233*, 107483. [CrossRef]
41. Wang, R.; Wang, Y.; Yan, F. Vegetation growth status and topographic effects in frozen soil regions on the Qinghai–Tibet plateau. *Remote Sens.* **2022**, *14*, 4830. [CrossRef]
42. Chen, J.; Huang, R.; Yang, Y.; Feng, Z.; You, H.; Han, X.; Yi, S.; Qin, Y.; Wang, Z.; Zhou, G. Multi-scale validation and uncertainty analysis of GEOV3 and MuSyQ FVC products: A case study of an alpine grassland ecosystem. *Remote Sens.* **2022**, *14*, 5800. [CrossRef]
43. Huang, G.; Li, X.; Huang, C.; Liu, S.; Ma, Y.; Chen, H. Representativeness errors of point-scale ground-based solar radiation measurements in the validation of remote sensing products. *Remote Sens. Environ.* **2016**, *181*, 198–206. [CrossRef]
44. Cao, C.; Lam, N.S.-N. Understanding the scale and resolution effects in remote sensing and GIS. In *Scale in Remote Sensing and GIS*; Routledge: London, UK, 2023; pp. 57–72.
45. Liu, D.; Jia, K.; Wei, X.; Xia, M.; Zhang, X.; Yao, Y.; Zhang, X.; Wang, B. Spatiotemporal comparison and validation of three global-scale fractional vegetation cover products. *Remote Sens.* **2019**, *11*, 2524. [CrossRef]
46. Chen, J.; Huang, R.; Luo, L.; Yi, S.; Qin, Y.; Qi, W.; You, H.; Han, X.; Zhou, G. A multi-tiered representativeness framework mitigating spatial scale effects in FVC validation: UAV-based assessment of global products in Qinghai-Tibetan Plateau alpine grasslands. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *143*, 104794. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.