Acta Geodyn. Geomater., Vol. 23, No. 2 (222), 235–248, 2026
DOI: 10.13168/AGG.2026.0015

journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

IMPROVING ICESAT-2 ATL08 LAND ELEVATION ACCURACY IN COMPLEX HILLY TERRAIN USING MACHINE LEARNING MODELS AND SPATIAL ANALYSIS

Shengpeng ZHANG 1, 2, 3), Yongming ZHANG 2, 4) *, Yongying ZHANG 2, 3, 7),
Guangjing HAO 1,5) and Hanmei LI 1, 6)

1) College of Geographical Sciences, Qinghai Normal University, Xining 810008, China
2) Geomatics Technology and Application Key Laboratory of Qinghai Province, Xining, China
3) Qinghai Basic Surveying and Mapping Institute, Xining 810001, China
4) Qinghai Provincial Institute of Surveying and Mapping Science and Technology
5) Qinghai Provincial Center for Land Management and Ecological Restoration
6) Provincial Disaster Relief Materials Reserve Center, Qinghai Provincial Emergency Management Department, Xining 810008, China
7) School of Geophysics and Geomatics, China University of Geosciences (Wuhan), Wuhan 430074, China 3.4 Statistical Analysis of Model Residuals

**Corresponding author's e-mail:* zhangyongming@qhbsmi.cn

ARTICLE INFO

Article history:
Received 16 Dcember 2025
Accepted 10 April 2026
Available online 21 April 2026

Keywords:
ICESat-2 ATL08
GF-7; DEM
Geodetector
Spatial validation
Machine learning

ABSTRACT

Accurate evaluation of ICESat-2 ATL08 terrain elevations in rugged mountainous areas is important for reliable geodetic applications. In this study, ATL08 Version 006 100 m terrain segments in the upper Yellow River Basin were evaluated against a 2 m GF-7 stereo DEM after converting ATL08 ellipsoidal heights to orthometric heights to ensure vertical-datum consistency. After quality screening, 15,904 segments were retained, and the mean absolute elevation difference was 14.08 m. Geodetector analysis identified terrain_slope and h_te_std as the dominant explanatory factors, whereas snr, n_ca_photons, h_te_uncertainty, and h_te_interp_minus_median played secondary but still significant roles. To reduce spatial leakage, the retained segments were partitioned into spatially disjoint blocks of approximately 10 km and systematically assigned to five folds. Six machine-learning models were then compared for predicting absolute elevation error under spatially blocked cross-validation. HistGradientBoosting achieved the best overall performance, with a mean test R^2 of 0.452 ± 0.016, RMSE of 10.65 ± 0.64 m, and MAE of 7.24 ± 0.40 m across the five spatial folds. The results indicate moderate but stable predictive skill in geographically unseen areas and support terrain-aware screening of ATL08 segments for DEM validation and related geodetic applications.

1. INTRODUCTION

Accurate elevation information is essential for geodesy, tectonic deformation studies, geomorphological analysis, and resource investigation (Schumann et al., 2008; Liu et al., 2022; Rexer and Hirt, 2014). NASA launched the Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) in September 2018 to provide global photon-counting laser altimetry observations of the Earth's surface (Cavanaugh, 2019; Kurtz et al., 2019). The ATL08 product, derived from geolocated photon data, provides terrain and vegetation height information and has been widely used in geodetic applications, topographic mapping, and forest height monitoring (Dandabathula et al., 2020; Tian and Shan, 2021). Compared with ICESat/GLAS, ICESat-2 offers much higher spatial resolution, with a nominal footprint diameter of about 17 m and an along-track spacing of about 0.7 m (Carabajal and Boy, 2020; Neumann et al., 2019). Under optimal conditions over flat and bare surfaces, its vertical accuracy can reach about 0.1 m (Brunt et al., 2019; Carabajal and Boy, 2020).

Previous studies have evaluated the accuracy of ICESat-2 ATL08 elevation data over different surface types (Brunt et al., 2019; Urbazaev et al., 2022). Although sub-meter or near-meter accuracy has been reported over flat and sparsely vegetated terrain, accuracy degrades substantially in mountainous areas with steep topography, heterogeneous surface conditions, and strong local roughness (Liu et al., 2022). Signal-to-noise ratio (SNR), photon density, slope, and land-surface characteristics all influence ATL08 elevation quality (Li et al., 2022; Magruder et al., 2021; Pang et al., 2022). Large elevation errors, especially in mountainous and forested environments, are often associated with signal attenuation, terrain-induced complexity, and insufficient sampling of the lidar returns (Popescu et al., 2018). Consistent and reliable elevation accuracy is therefore critical for applications such as the selection of ground control points (GCPs) for DEM production and validation (Kandil et al., 2024; Zhang et al., 2024).

Validation of satellite-derived elevation products requires an independent and reliable reference dataset (Paul et al., 2017; Purinton and Bookhagen, 2021).

Cite this article as: Zhang S, Zhang Y, Zhang Y, Hao G, Li H: Improving ICESat-2 ATL08 land elevation accuracy in complex hilly terrain using Machine Learning Models and spatial analysis. Acta Geodyn. Geomater., 23, No. 2 (222), 235–248, 2026. DOI: 10.13168/AGG.2026.0015

The GaoFen-7 (GF-7) satellite, China's first sub-meter stereo mapping satellite launched in November 2019, provides high-quality stereo imagery and supports high-precision DEM generation (Haiyi et al., 2021; Li et al., 2024). Equipped with a dual-line-array stereo mapping camera and a laser altimeter for elevation control, GF-7 can produce DEMs with high vertical accuracy when supported by suitable ground control and photogrammetric processing (Liu et al., 2021; Liu et al., 2024; Zhu et al., 2023). This makes GF-7 DEMs a valuable reference for evaluating ICESat-2 elevation products, especially in regions where dense ground control is unavailable. However, relatively few studies have used GF-7 DEMs to validate spaceborne altimetry data.

Although several studies have compared ICESat-2 heights with airborne LiDAR or existing DEM products, fewer have systematically examined the mechanisms that control ATL08 elevation errors in complex terrain (Pronk et al., 2023; Zhang et al., 2024). Most previous studies have focused mainly on summary accuracy statistics or simple relationships between elevation error and individual explanatory factors (Xu et al., 2024). Geodetector, proposed by Wang and Xu, is an effective method for quantifying spatial stratified heterogeneity and detecting the explanatory power of potential driving factors (Wang and Xu, 2017). It has been widely used in geography and environmental studies, but has rarely been applied to satellite altimetry error analysis (Guo et al., 2024; Liang and Xu, 2023; Wang et al., 2021). At the same time, machine-learning approaches have shown promise for DEM error modelling and elevation-related prediction tasks, but comparative evaluation of such methods for ATL08 terrain-elevation error prediction remains limited (Maxwell et al., 2018). Thus, an integrated framework that combines factor explanation and predictive modelling for ATL08 errors is still lacking.

In this study, we aim to: (1) quantify the vertical accuracy of ICESat-2 ATL08 terrain heights in a complex hilly region using a high-resolution GF-7 DEM as reference; (2) identify and quantify the main factors controlling ATL08 elevation errors using Geodetector; and (3) develop and compare machine-learning models for predicting absolute elevation errors. The novelty of this work lies in the combined use of Geodetector and machine learning to link explanatory factor analysis with practical error prediction. The study area is located in the Upper Yellow River Basin of Qinghai Province, China, where complex terrain poses substantial challenges for both spaceborne altimetry and stereo mapping. The proposed framework is expected to provide useful guidance for terrain-aware use of ICESat-2 ATL08 data in geodetic and geomorphological applications.

2. DATA AND METHODS

2.1. STUDY AREA

Our study area is located in the Upper Yellow River Basin, Qinghai Province, northwestern China (Fig. 1). It extends approximately from 100°E to 103°E and lies within a broad elevational and geomorphic transition zone along the northeastern margin of the Qinghai–Tibet Plateau. The study area covers about 22,398 km^2 and includes the upper main course of the Yellow River together with the surrounding mountainous and hilly terrain.

The region is characterized by complex topography, high relief, diverse slope and aspect conditions, and heterogeneous landform assemblages ranging from river valleys and piedmont zones to more rugged uplands. Land cover is dominated mainly by plateau grassland, with patches of shrubs and bare ground, whereas relatively developed surfaces are concentrated primarily in river valleys. This combination of complex terrain and heterogeneous surface cover provides an ideal setting for evaluating the performance of ICESat-2 ATL08 terrain elevations and for investigating how terrain complexity and signal-quality metrics influence elevation errors.

To minimize seasonal variation and potential confounding effects of snow/ice cover and vegetation phenology, all ATL08 segments and GF-7 imagery used in this study were acquired in June 2023 under summer, largely snow-free conditions. This temporal consistency supports a direct comparison between ATL08 terrain elevations and the GF-7 DEM and allows the observed elevation differences to be interpreted mainly in relation to topographic and surface-condition effects. Therefore, the present analysis should be regarded as a summer-condition assessment rather than a universal year-round calibration.

2.2. ICESAT-2 ATL08 DATA PRODUCT

NASA's Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2), launched in September 2018, carries the Advanced Topographic Laser Altimeter System (ATLAS), a photon-counting lidar operating at 532 nm. ATLAS uses six beams arranged in three strong/weak pairs to provide dense along-track sampling and improved spatial coverage of land-surface observations (Kurtz et al., 2019).

In this study, we used the ATL08 Land and Vegetation Height product (Version 006) to evaluate terrain-related metrics. ATL08 reports along-track terrain and canopy heights relative to the WGS84 ellipsoid in the ITRF2014 reference frame using fixed 100 m segments. These data were obtained from the NSIDC DAAC.

All six ground tracks intersecting the study area (gt1l, gt1r, gt2l, gt2r, gt3l, and gt3r) were processed. A Python workflow based on the h5py library was used to extract terrain parameters and quality indicators for each 100 m segment. The candidate variables used in the factor analysis and machine-learning stage are summarized in Table 1.

After applying the quality-control criteria described in Section 2.7, a total of 15,904 valid ATL08 segments were retained; their spatial distribution is shown in Figure 2.

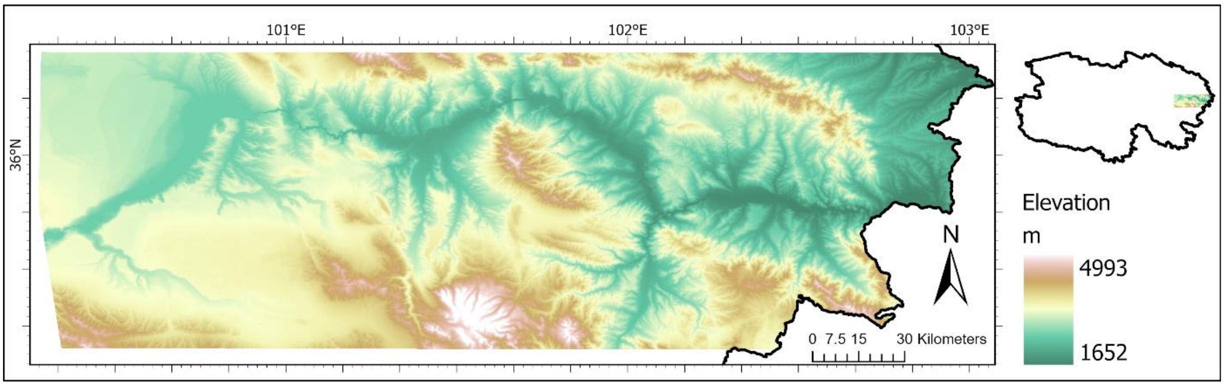

Fig. 1 Location of the study area in the Upper Yellow River Basin, Qinghai Province, China.

Table 1 ICESat-2 ATL08 V006 terrain parameters extracted for this study.

Parameter	Description	Unit
longitude	Longitude of segment center	degrees
latitude	Latitude of segment center	degrees
h_te_best_fit	Best-fit terrain elevation at segment center	m
h_te_interp	Interpolated terrain elevation from ground photons	m
h_te_median	Median terrain elevation within segment	m
h_te_mean	Mean terrain elevation within segment	m
h_te_max	Maximum terrain elevation within segment	m
h_te_min	Minimum terrain elevation within segment	m
h_te_rh25	Relative height at 25th percentile	m
h_te_std	Standard deviation of terrain elevations	m
h_te_skew	Skewness of terrain elevation distribution	-
h_te_uncertainty	Uncertainty of terrain height estimate	m
terrain_slope	Along-track terrain slope	degrees
dem_h	Reference DEM elevation used	m
h_dif_ref	Difference from reference DEM	m
cloud_flag_atm	Atmospheric cloud confidence flag	-
layer_flag	Cloud or blowing snow layer flag	-
msw_flag	Multiple scattering warning flag	-
night_flag	Day/night acquisition flag	-
Snr	Signal-to-noise ratio	-
terrain_flg	Terrain quality flag	-
urban_flag	Urban area indicator (DLR GUF)	-
dem_removal_flag	DEM-based outlier removal flag	-
photon_rate_te	Terrain photon rate	photons/shot
n_te_photons	Number of terrain photons	count
n_ca_photons	Number of canopy photons	count
n_toc_photons	Number of top-of-canopy photons	count
segment_landcover	Copernicus land cover classification	-

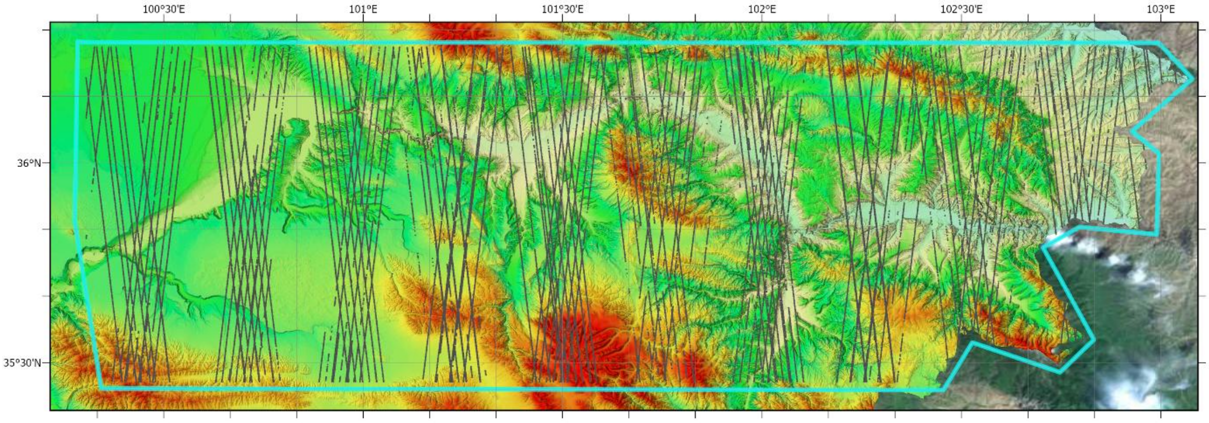

Fig. 2 Spatial distribution of retained ICESat-2 ATL08 100 m segments within the study area after quality screening.

2.3. GF-7 REFERENCE DEM

A 2 m GF-7 stereo DEM was used as the reference dataset for evaluating ATL08 terrain elevations in the study area. ATL08 terrain heights are provided relative to the WGS84 ellipsoid in the ITRF2014 reference frame. To ensure vertical-datum consistency with the GF-7 DEM, the ATL08 ellipsoidal heights were converted to normal heights using the Qinghai provincial quasigeoid model before elevation differencing. After datum harmonization, the GF-7 reference elevation corresponding to each ATL08 100 m segment location was extracted by bilinear interpolation and used for subsequent error calculation. ATL08 V006 reports terrain and canopy heights for fixed 100 m segments relative to the WGS84 ellipsoid, which makes this vertical-datum conversion necessary before comparison with a regional height-reference surface.

2.4. ELEVATION ERROR DEFINITION

The signed elevation difference between the ICESat-2 ATL08 terrain height and the GF-7 DEM was defined as:

$$\Delta H = H_{\mathrm{ATL08}}^{\mathrm{norm}} - H_{\mathrm{GF\text{-}7}} \quad (1)$$

where $H_{\mathrm{ATL08}}^{\mathrm{norm}}$ denotes the ATL08 terrain height after conversion from ellipsoidal height to normal height using the Qinghai provincial quasigeoid model, and $\mathrm{H}_{\mathrm{GF-7}}$ denotes the reference elevation extracted from the GF-7 DEM at the corresponding ATL08 100 m segment location by bilinear interpolation.

For the subsequent Geodetector analysis and machine-learning modelling, the dependent variable was defined as the absolute elevation error,

$$\mathrm{abs\_h\_dif\_ref} = |\Delta H| \quad (2)$$

so that the analysis focused on error magnitude irrespective of sign.

2.5. GEODETECTOR FACTOR ANALYSIS

2.5.1. ALGORITHM PRINCIPLE

Geodetector is a spatial statistical method for detecting spatial stratified heterogeneity and identifying driving factors. Based on variance decomposition, it estimates the explanatory power of independent variable on the differentiation of dependent variable through the comparison of within- group variance to total variance (Wang and Xu, 2017). In the analysis of elevation errors, we used the factor detector, interaction detector and the ecological detector to quantify the effects of terrain and signal-quality variables.

2.5.2. MATHEMATICAL FORMULATION

The factor detector is measured by the q-statistic:

$$q = 1 - (1/N\sigma^2)\sum_{h=1}^{L} N_h\sigma_h^2 \quad (3)$$

$$q = 1 - SSW/SST \quad (4)$$

where q ranges from 0 to 1, with larger values indicating stronger explanatory power; N is the total number of samples; L is the number of strata; and Nh and σh² are the sample size and the variance of the dependent variable in stratum h, respectively. Equation (2) is a compact form of Equation (1), where SSW denotes the within-stratum variance term and SST denotes the total variance term. Because Geodetector requires categorical inputs, all continuous variables were discretized into five quantile-based classes before analysis.

2.6. MACHINE LEARNING MODELS

Six regression models were compared to predict ATL08 absolute elevation error: HistGradientBoosting, LightGBM, CatBoost, ExtraTrees, XGBoost, and Linear Regression. The final predictor set consisted of six variables: terrain_slope, h_te_std, snr, n_ca_photons, h_te_uncertainty, and h_te_interp_minus_median.

The machine-learning stage was designed as a predictive screening exercise rather than as a signed-error correction procedure. Model comparison therefore focused on independent test performance and generalization under spatially separated training and testing data.

2.7. DATA PREPROCESSING AND MODEL TRAINING

For the 15,904 retained ATL08 segments, the six selected predictors were scaled to the interval [0, 1] using MinMax normalization. To reduce spatial leakage, spatially blocked cross-validation was implemented using the blockCV framework (Valavi et al., 2019). Preliminary tests with block sizes of approximately 5 km, 10 km, and 15 km were conducted to evaluate the trade-off between spatial separation and fold-wise sample distribution. Among these candidate configurations, the 10 km scheme provided the best balance between reducing spatial dependence and maintaining reasonably even sample sizes across folds; it was therefore adopted in the final analysis. As shown in Figure 3, the retained samples were partitioned into five spatial folds using approximately 10 km blocks and systematic fold assignment. For each of the five iterations, one fold was exported as the test set and the remaining four folds were combined as the training set, generating five train–test pairs in total. Thus, each fold served once as the held-out test set and four times as part of the training data. Final model performance was summarized across all five spatial folds.

3. RESULTS

3.1. BASIC CHARACTERISTICS OF ATL08 ELEVATION ERRORS

General statistics of ATL08 absolute elevation errors with respect to the GF-7 DEM (via all-metrics) can be found in Table 2, where |ΔH| have a mean and std of 14.08 m and 14.44 m with a range of 0.002 to 113.62 m. The distribution is slightly right skewed, indicating that most ATL08 100 m segments have relatively small errors, with a smaller number of larger errors.

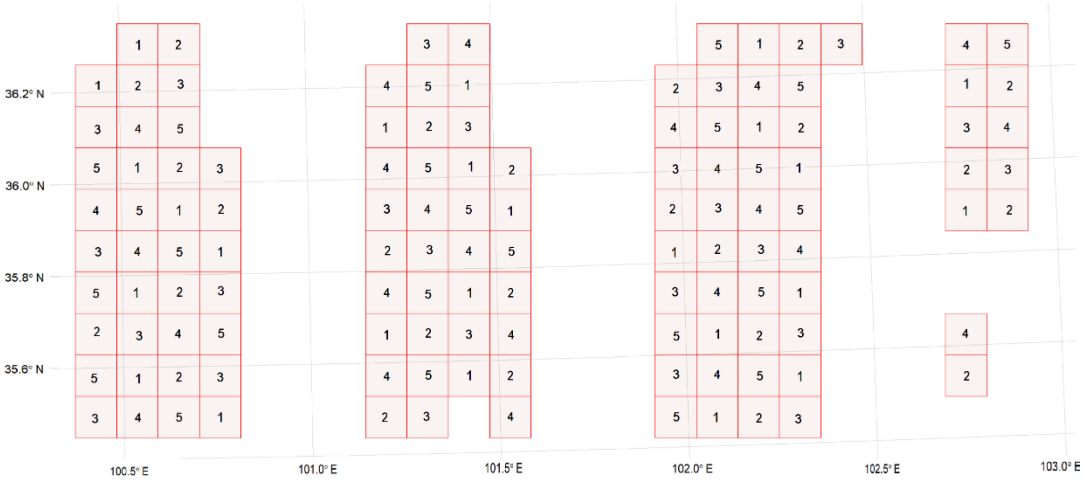

Fig. 3 Spatially blocked 5-fold cross-validation scheme based on ~10 km blocks.

Table 2 Summary statistics of ATL08 absolute elevation errors (|ΔH|) (abs_h_dif_ref).

Statistic	Value	Statistic	Value
Mean	14.08 m	Min	0.002 m
Std. Dev.	14.44 m	Max	113.62 m
Samples	15,904	Factors	28

Note: 28 candidate factors were initially evaluated in the Geodetector analysis.

Table 3 Significant factors and q-values for ATL08 absolute elevation errors ($p < 0.001$).

Factor	**Description**	**q-value**	**p-value**
terrain_slope	ATL08 along-track terrain slope (degrees)	0.351	<0.001
h_te_std	Standard deviation of terrain elevation estimates (m)	0.348	<0.001
Snr	Signal-to-noise ratio	0.210	<0.001
n_ca_photons	Number of canopy photons	0.209	<0.001
h_te_uncertainty	Uncertainty of terrain height estimate (m)	0.181	<0.001
h_te_interp_minus_median	Difference between interpolated and median terrain height (m)	0.153	<0.001

3.2. GEODETECTOR FACTOR ANALYSIS RESULTS

Geodetector factor detection was used to evaluate the influence of the final six terrain- and signal-related variables on ATL08 absolute elevation error. All six variables were statistically significant ($p< 0.001$; Table 3).

Among the six predictors, terrain_slope ($q = 0.351$) and h_te_std ($q = 0.348$) showed the strongest explanatory power, confirming the dominant role of terrain complexity in controlling ATL08 elevation-error heterogeneity. By comparison, the remaining variables contributed additional but weaker explanatory power, including snr ($q = 0.210$), n_ca_photons ($q = 0.209$), h_te_uncertainty ($q = 0.181$), and h_te_interp_minus_median ($q = 0.153$).

The interaction detector was then applied to the six selected variables (Table 4; Fig. 4). All pairwise combinations showed enhancement effects. Among the 15 tested pairs, 14 were classified as Two Factor Enhance, whereas the interaction between snr and n_ca_photons showed the strongest Nonlinear Enhance effect ($q = 0.783$), substantially exceeding the explanatory power of either variable alone.

The strongest two-factor enhancement effects involved terrain_slope or h_te_std interacting with signal-quality or uncertainty-related variables, including terrain_slope × snr ($q = 0.435$), terrain_slope × n_ca_photons ($q = 0.408$), h_te_std × snr ($q = 0.422$), and h_te_std × n_ca_photons ($q = 0.379$). The interaction between terrain_slope and h_te_std was also strong ($q = 0.402$). These results indicate that terrain complexity and photon-return conditions reinforce each other in controlling ATL08 error magnitude.

The ecological detector was used to test whether the explanatory power of different variables differed significantly. The effects of terrain_slope and h_te_std were not significantly different ($p = 0.731$), and snr and n_ca_photons also showed no significant difference ($p = 0.591$), indicating two internally coherent factor groups: a terrain-complexity group and a signal-quality group.

Table 4 Selected interaction detection results showing combined effects of factor pairs.

Factor 1	Factor 2	q_1	q_2	q_{12}	Interaction Type
snr	n_ca_photons	0.212	0.209	0.790	Nonlinear Enhance
n_te_photons	n_ca_photons	0.189	0.209	0.749	Nonlinear Enhance
snr	n_ca_photons	0.210	0.209	0.783	Nonlinear Enhance
terrain_slope	snr	0.351	0.210	0.435	Two Factor Enhance
h_te_std	snr	0.348	0.210	0.422	Two Factor Enhance
terrain_slope	n_ca_photons	0.351	0.209	0.408	Two Factor Enhance
terrain_slope	h_te_std	0.351	0.348	0.402	Two Factor Enhance
h_te_std	n_ca_photons	0.348	0.209	0.379	Two Factor Enhance

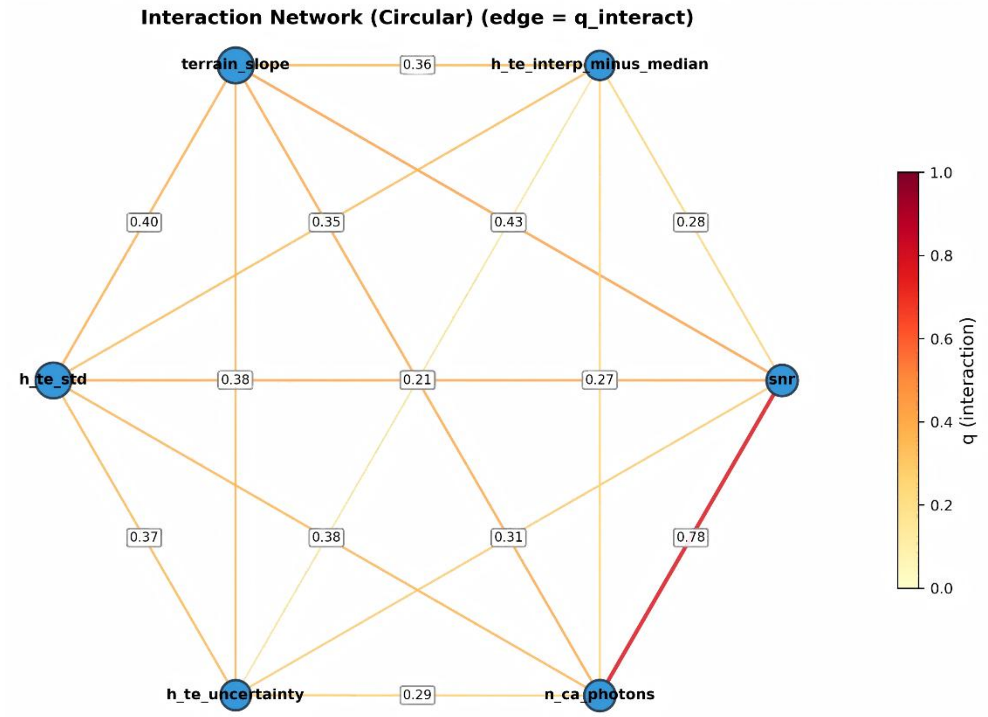

Fig. 4 Circular interaction network showing pairwise q-values among the six selected variables.

By contrast, most comparisons between terrain-related and signal-related variables were statistically significant. For example, terrain_slope differed significantly from snr, n_ca_photons, h_te_uncertainty, and h_te_interp_minus_median, and h_te_std showed the same pattern. These results suggest that terrain and photon-return variables explain ATL08 error through related but distinct mechanisms, supporting the inclusion of both groups in the subsequent predictive modelling stage.

To support the Geodetector analysis, the continuous predictors were discretized into strata before factor detection. As shown in Figure 5, the four continuous variables included in this step — terrain_slope, h_te_std, h_te_uncertainty, and h_te_interp_minus_median — were divided into eight groups (G0–G7). The resulting distributions show distinct characteristics: terrain_slope is centered near zero with moderate asymmetry, whereas h_te_std, h_te_uncertainty, and h_te_interp_minus_median all exhibit clear right-skewed distributions with long upper tails. This pattern indicates that most ATL08 segments are concentrated in relatively low-to-moderate ranges, while a smaller proportion of segments occupy high-value strata associated with more complex terrain or larger internal uncertainty.

The discretization results also help explain the final factor-detection pattern. The two dominant terrain-related variables, terrain_slope and h_te_std, each showed strong explanatory power with q-values of 0.3512 and 0.3477, respectively, whereas h_te_uncertainty and h_te_interp_minus_median contributed weaker but still significant explanatory power (0.1812 and 0.1527). These results indicate that the adopted discretization scheme was sufficient to preserve the main structure of the continuous predictors while yielding an interpretable and compact variable set for the subsequent Geodetector and machine-learning analyses.

The Geodetector results indicate that ATL08 elevation-error assessment in complex terrain should prioritize terrain_slope and h_te_std as the primary descriptors of local terrain complexity, while retaining snr, n_ca_photons, h_te_uncertainty, and

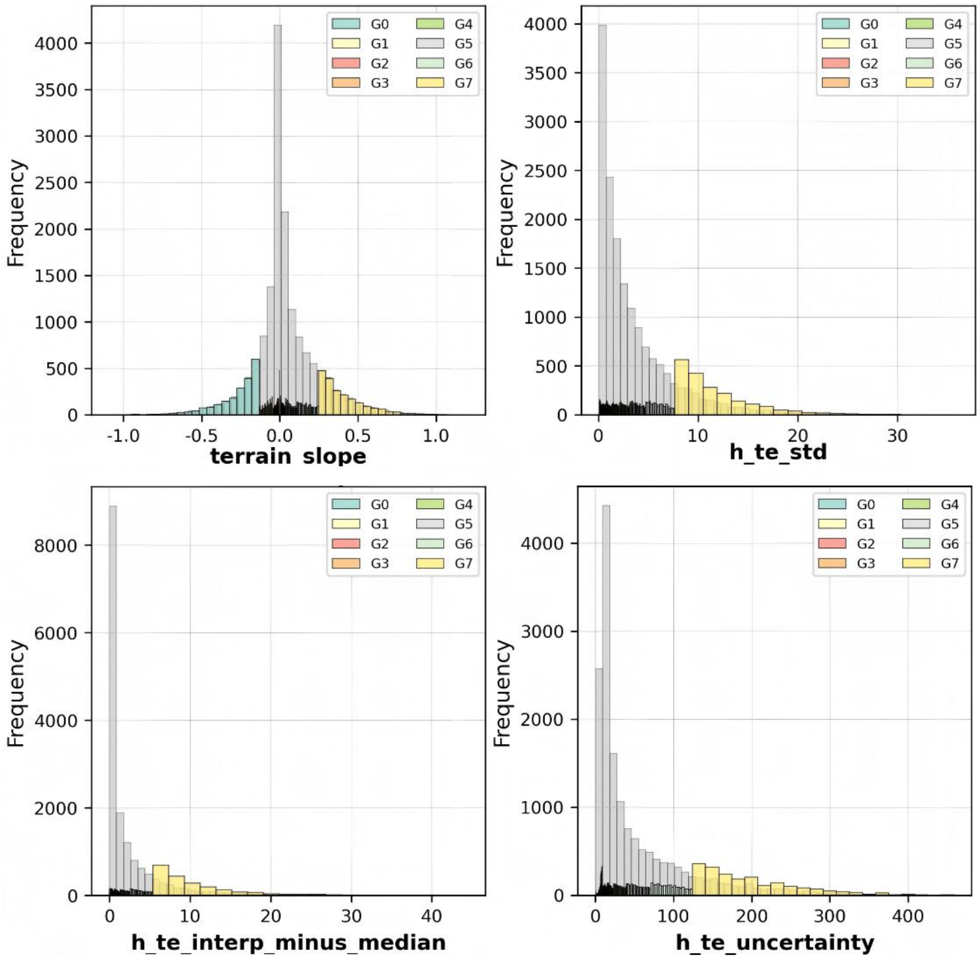

Fig. 5 Discretization results of the continuous predictors.

Table 5 Mean performance of the candidate machine-learning models under spatially blocked 5-fold cross-validation.

Model	Mean Train R²	Mean Test R²	Mean RMSE (m)	Mean MAE (m)	Mean Overfitting Index
Hist Gradient Boosting	0.581 ± 0.019	0.452 ± 0.016	10.65 ± 0.64	7.24 ± 0.40	0.129 ± 0.034
CatBoost	0.688 ± 0.004	0.442 ± 0.016	10.74 ± 0.60	7.30 ± 0.38	0.246 ± 0.018
LightGBM	0.630 ± 0.005	0.441 ± 0.013	10.75 ± 0.64	7.29 ± 0.40	0.189 ± 0.018
ExtraTrees	1.000 ± 0.000	0.438 ± 0.020	10.78 ± 0.59	7.36 ± 0.40	0.562 ± 0.020
Linear Regression	0.399 ± 0.005	0.390 ± 0.020	11.23 ± 0.64	7.90 ± 0.38	0.008 ± 0.025
XGBoost	0.824 ± 0.002	0.386 ± 0.014	11.27 ± 0.67	7.61 ± 0.40	0.437 ± 0.015

h_te_interp_minus_median as secondary but meaningful predictors. These insights directly motivated the final six-variable predictor set used in Section 3.3.

3.3. MACHINE LEARNING MODEL COMPARISON

3.3.1. OVERALL PERFORMANCE UNDER SPATIALLY BLOCKED VALIDATION

To evaluate predictive performance under geographically independent conditions, six regression models were compared using the spatially blocked 5- fold cross-validation framework described in Section 2.7. The candidate models were HistGradientBoosting, CatBoost, LightGBM, ExtraTrees, Linear Regression, and XGBoost. For each iteration, one spatial fold was used as the test set and the remaining four folds were combined as the training set. Before model fitting, the six selected predictors were normalized to the range [0, 1] using MinMax scaling. The fold-averaged model performance metrics are summarized in Table 5

As shown in Table 5 HistGradientBoosting achieved the best overall performance among the six models. Its mean training R² was 0.581 ± 0.019, and

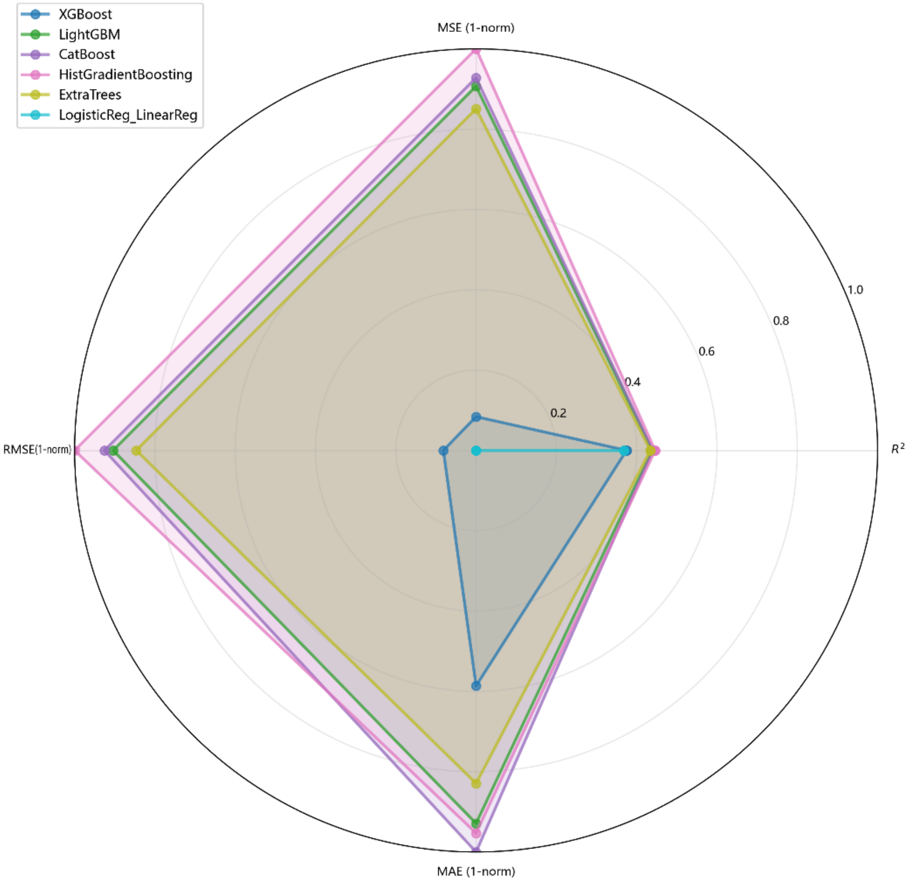

Fig. 6 Radar-chart comparison of the six machine-learning models under spatially blocked 5-fold cross-validation.

its mean test R² reached 0.452 ± 0.016, with a mean RMSE of 10.65 ± 0.64 m and a mean MAE of 7.24 ± 0.40 m. CatBoost and LightGBM formed a competitive second tier, with mean test R² values of 0.442 ± 0.016 and 0.441 ± 0.013, respectively, and only slightly larger prediction errors than HistGradientBoosting. By contrast, Linear Regression showed the weakest capacity to capture nonlinear structure, yielding a mean test R² of 0.390 ± 0.020, whereas XGBoost had the lowest mean test R² among the nonlinear ensemble models (0.386 ± 0.014).

ExtraTrees produced relatively competitive test scores, with a mean test R² of 0.438 ± 0.020, but its training R² remained 1.000 in all five folds, indicating severe overfitting. This behavior is consistent with the very large mean overfitting index (0.562 ± 0.020), which was much higher than those of the other models. Overall, the fold-averaged comparison in Table 5 indicates that HistGradientBoosting provides the most favorable balance between predictive skill and spatial generalization under geographically unseen conditions.

To facilitate an integrated comparison of the six models, the fold-averaged performance metrics are further visualized in Figure 6. The radar chart confirms the ranking summarized in Table 5, showing that HistGradientBoosting achieved the most balanced overall performance, with the highest test R² and the most favorable normalized error metrics. CatBoost and LightGBM formed a competitive second tier with very similar performance profiles, whereas ExtraTrees showed relatively strong test accuracy but substantially weaker generalization because of severe overfitting. Linear Regression produced the most conservative fit but failed to capture the nonlinear structure of ATL08 absolute elevation error, while XGBoost showed the weakest overall performance under spatially blocked validation. Overall, Figure 6 supports the selection of HistGradientBoosting as the final model.

3.3.2. SELECTION OF THE FINAL MODEL

HistGradientBoosting was selected as the final model because, as summarized in Table 5. Figure it achieved the best fold-averaged test performance among all six models. Second, its mean overfitting index (0.129 ± 0.034) was substantially smaller than those of CatBoost (0.246 ± 0.018), LightGBM (0.189 ± 0.018), ExtraTrees (0.562 ± 0.020), and XGBoost (0.437 ± 0.015), indicating a more favorable compromise between model flexibility and generalization. Third, the model ranked first or near first in all spatial folds, demonstrating stable predictive behavior across geographically distinct subsets.

These results indicate that ATL08 absolute elevation error in the study area is only moderately predictable under spatially independent validation, but

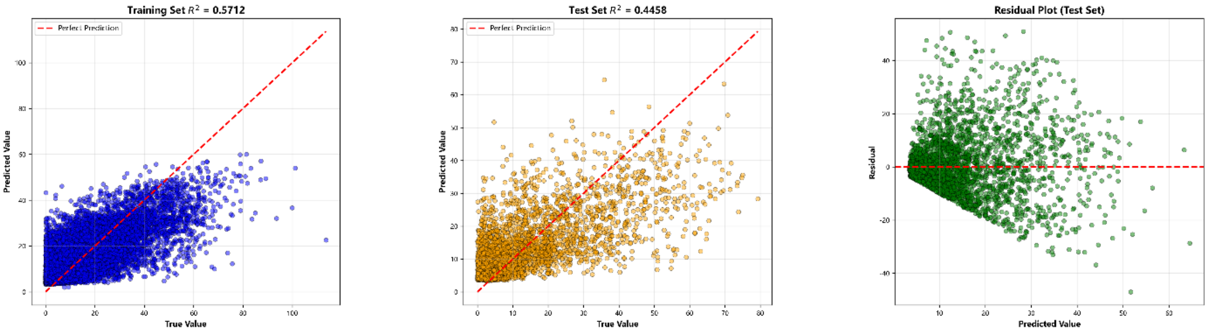

Fig. 7 Diagnostic plots of the final HistGradientBoosting model.

that the error field still contains stable nonlinear structure that can be captured effectively by histogram-based gradient boosting. Compared with the originally used random split, the spatially blocked validation yields more conservative but more realistic estimates of predictive performance, and therefore provides a more reliable basis for evaluating model transferability to geographically unseen areas.

Left: predicted versus observed values for the training set. Middle: predicted versus observed values for the test set. Right: residuals versus predicted values for the test set. The red dashed line in the scatterplots indicates the 1:1 relationship, and the horizontal dashed line in the residual plot indicates zero residual.

3.3.3. SHAP-BASED INTERPRETATION OF THE FINAL MODEL

To interpret the final HistGradientBoosting model, SHapley Additive exPlanations (SHAP) were used to quantify the contribution of each predictor to the predicted absolute elevation error. Tree SHAP is specifically designed for tree-based models and tree ensembles and provides fast, exact explanations of feature contributions, making it appropriate for HistGradientBoosting. As shown in Figure 7, the fold-wise SHAP results were highly consistent and revealed a stable feature hierarchy across the five spatial folds.

Based on the mean absolute SHAP values, h_te_std was the most influential predictor, accounting for approximately 42.7 % of the total relative contribution, followed by terrain_slope (24.8 %) and h_te_uncertainty (13.7 %). The remaining predictors played secondary but still interpretable roles, including n_ca_photons (7.2 %), snr (7.2 %), and h_te_interp_minus_median (4.4 %). This ranking is broadly consistent with the Geodetector results presented in Section 3.2, which also identified terrain_slope and h_te_std as the dominant explanatory variables. Taken together, the Geodetector and SHAP analyses indicate that ATL08 absolute elevation error in the study area is controlled primarily by terrain complexity, whereas signal- and uncertainty-related variables provide secondary but meaningful explanatory power.

The SHAP beeswarm plot in Figure 8 further shows that larger values of h_te_std are generally associated with positive SHAP values, indicating increased predicted error under stronger local terrain variability. A similar positive contribution pattern is observed for higher h_te_uncertainty, whereas the effects of snr and n_ca_photons are weaker and more concentrated near zero. Overall, the SHAP results confirm that the final HistGradientBoosting model is dominated by terrain-related predictors, with signal-return and uncertainty variables acting as additional modifiers of the predicted error field.

The internal feature attribution structure of the final model is further examined using SHAP in Figure 8.

Although HistGradientBoosting was selected as the final model, the alternative models remain informative. CatBoost and LightGBM produced performances very close to HistGradientBoosting and may therefore be regarded as strong alternative choices. Linear Regression showed the smallest train–test discrepancy, with a mean overfitting index of 0.008 ± 0.025, but its lower test R^2 and larger prediction errors indicate that a purely linear representation is insufficient for the nonlinear structure of ATL08 absolute elevation error. ExtraTrees and XGBoost showed substantially stronger in-sample fitting but weaker spatial generalization, suggesting greater sensitivity to local training patterns.

Overall, the six-model comparison suggests that the ATL08 error surface is nonlinear, moderately structured, and more amenable to gradient boosting than to either a purely linear model or a strongly overfitted ensemble model. This result supports the adoption of HistGradientBoosting as the final model for terrain-aware prediction of ATL08 absolute elevation error.

4. DISCUSSION

4.1. DOMINANT FACTORS AFFECTING ATL08 ACCURACY

Both Geodetector and SHAP indicate that ATL08 absolute elevation error in the upper Yellow River Basin is controlled primarily by terrain complexity, whereas signal-return and uncertainty-related variables play secondary but still meaningful roles. In the Geodetector analysis, terrain_slope and h_te_std showed the strongest explanatory power,

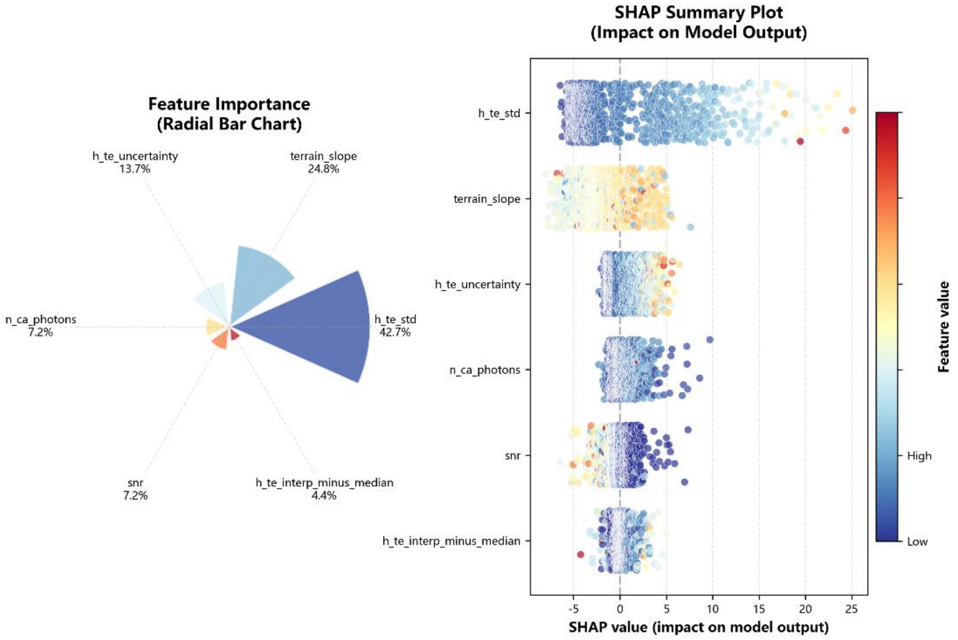

Fig. 8 SHAP summary plot and relative feature importance for the final HistGradientBoosting model.

while the SHAP results of the final HistGradientBoosting model identified h_te_std as the most influential predictor, followed by terrain_slope and h_te_uncertainty. Together, these results show that local relief variability and along-track terrain steepness are the dominant controls on ATL08 error magnitude, while photon-return conditions and terrain-height uncertainty act as additional modifiers rather than primary drivers.

This pattern is physically consistent with the photon-counting retrieval process. In steep and topographically variable terrain, the local ground surface becomes more difficult to represent using segment-based fitting, and the identification of ground photons becomes more uncertain because terrain-induced variability broadens the local elevation distribution. Under such conditions, even relatively acceptable signal conditions cannot fully compensate for the complexity introduced by rugged terrain. Conversely, in relatively gentle terrain with lower h_te_std and terrain_slope, ATL08 tends to perform better even when snr or photon counts are not optimal. This explains why terrain-related variables dominate both the statistical factor-detection results and the model-based attribution patterns.

An additional implication is that ATL08 error in complex mountainous terrain should not be interpreted simply as a function of “low SNR” or “few photons”. Instead, the present results suggest that terrain complexity sets the primary limit on achievable accuracy, while signal-quality variables modulate the magnitude of the error around that terrain-controlled baseline. This conclusion is important for practical screening of ATL08 segments: reducing error risk in rugged terrain requires attention first to terrain heterogeneity, and only secondarily to photon-return conditions.

4.2. MODEL PERFORMANCE AND IMPLICATIONS

From a methodological perspective, the six-model comparison indicates that the ATL08 error surface is clearly nonlinear, but still sufficiently structured to be captured by a moderately flexible gradient-boosting model. Linear Regression provided the most conservative generalization behavior, with the smallest train–test discrepancy, but its lower test R^2 and larger errors indicate that a purely linear representation is insufficient for describing ATL08 absolute elevation error in complex terrain. At the other extreme, ExtraTrees and XGBoost achieved strong in-sample fitting but exhibited substantially larger train–test gaps, implying greater sensitivity to local training patterns and weaker spatial transferability. This behavior is especially evident for ExtraTrees, whose training R^2 remained 1.000 in all folds, consistent with the known tendency of fully grown, unpruned tree ensembles to overfit when model complexity is not explicitly controlled.

HistGradientBoosting achieved the best overall balance between predictive skill and spatial generalization. It produced the highest fold-averaged test R^2 and the lowest mean RMSE and MAE among the six models, while maintaining a markedly smaller overfitting index than CatBoost, LightGBM, ExtraTrees, and XGBoost. This result suggests that the histogram-based gradient boosting framework is well suited to the present dataset size and to the nonlinear but moderately structured nature of the ATL08 error field. Compared with the original random split, the spatially blocked validation yields more conservative but more realistic performance estimates, implying that the current results provide a more credible assessment of model transferability to geographically unseen areas.

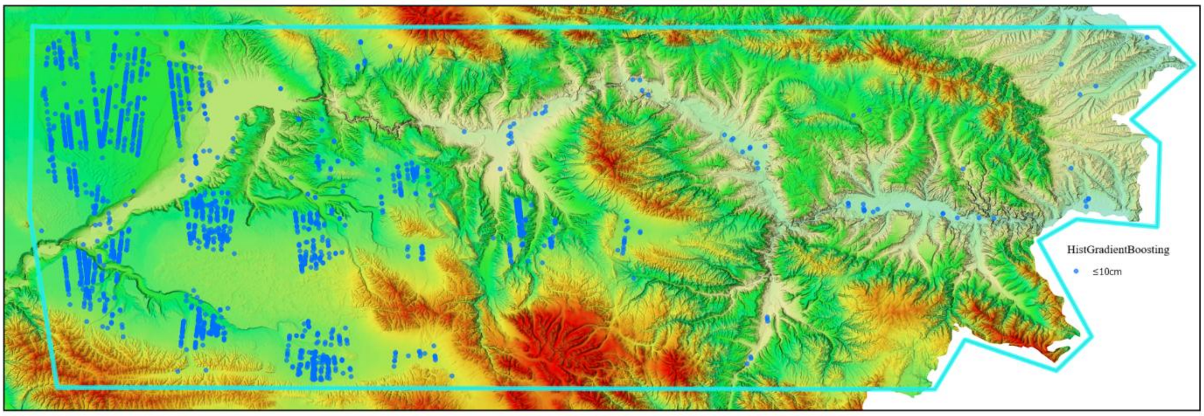

Fig. 9 Spatial distribution of ATL08 segments predicted by HistGradientBoosting to have very low absolute elevation error (|ΔH| ≤ 0.10 m).

The practical implication is that HistGradientBoosting should be viewed as a terrain-aware error prediction tool rather than an exact point-level correction model. Its performance indicates meaningful but moderate predictive skill: the model can distinguish relatively higher-risk and lower-risk ATL08 segments, but it does not fully recover very small absolute errors at the individual-segment level. In this sense, the model is more suitable for reliability screening and spatial prioritization than for exact deterministic correction of each ATL08 segment. The remaining errors are most likely concentrated in the steepest and most heterogeneous terrain, where local relief complexity still exceeds what can be represented by the current six-variable predictor set and 100 m segment-scale modelling framework.

4.3. *IMPLICATIONS FOR GCP SELECTION*

The present results have direct implications for the practical use of ATL08 in DEM validation and preliminary control-point screening. Because the final HistGradientBoosting model predicts the expected magnitude of ATL08 absolute elevation error, it can be used to identify relatively lower-risk and higher-risk ATL08 segments before they are considered in downstream geodetic applications. In this sense, the model is more appropriately viewed as a terrain-aware reliability screening tool than as a deterministic point-level correction model.

As shown in Figure 9, ATL08 segments predicted to have very low absolute elevation error (|ΔH| ≤ 0.10 m) are concentrated mainly in gentle river valleys, broad interfluves, and low-relief hillslopes, whereas steep mountain fronts, sharp ridges, and deeply incised gullies contain relatively few such candidates. This spatial pattern is consistent with the Geodetector and SHAP results, which both indicate that terrain-related variables dominate the error structure. Therefore, ATL08 segments located in terrain with lower predicted error may be prioritized as candidate supporting points for DEM assessment or regional elevation-quality screening, whereas segments located in terrain with higher predicted error should either be excluded or assigned lower confidence.

However, the present results do not support the unrestricted direct use of all such segments as primary GCPs. The fold-averaged performance of the final model indicates moderate predictive skill rather than exact point-level recovery of very small errors. Accordingly, the mapped low-error segments in Figure 9 should be interpreted as relatively favorable candidates for further screening and weighting, rather than as definitive control points. In practical terms, the framework is most useful for pre-screening, confidence stratification, and regional support for DEM validation rather than for replacing conventional high-precision ground control.

4.4. *LIMITATIONS AND FUTURE WORK*

Several limitations of the present study should be acknowledged, and they also point directly to priorities for future work. First, the analysis is based on ATL08 and GF-7 data acquired in June 2023 under summer, largely snow-free conditions. This improves temporal consistency between the two datasets and reduces confounding effects from seasonal snow and vegetation phenology; however, it also means that the present conclusions should be interpreted primarily as valid for summer conditions. Previous studies have shown that ATL08 terrain-height accuracy can vary with terrain slope, land-cover or canopy conditions, and season, indicating that transferability of the current model to winter or transitional seasons remains uncertain.

Second, although the study area is representative of the complex hilly terrain along the northeastern margin of the Qinghai–Tibet Plateau, it still corresponds to a single regional geomorphic and climatic setting. The relative importance of terrain-, signal-, and uncertainty-related predictors may differ in other environments, especially in densely forested, humid, or snow-dominated mountain regions. Therefore, the present model should not be interpreted as a universal calibration model for all ATL08 applications.

Third, the current framework is segment-based and relies on six selected predictors derived from terrain, photon-return, and uncertainty information. Although this compact predictor set improves interpretability and model stability, it does not explicitly represent all sources of residual variation, particularly local spatial dependence, geolocation-related effects, and fine-scale surface heterogeneity within 100 m ATL08 segments. This likely explains why the final model achieved meaningful but only moderate predictive skill under spatially blocked validation.

Fourth, although Geodetector and SHAP provide consistent evidence for the dominant role of terrain complexity, they do not eliminate all ambiguity in causal interpretation. The present framework should therefore be regarded as a physically informed predictive and screening model rather than a complete mechanistic description of ATL08 terrain-error generation.

These limitations suggest several directions for future work. Extending the analysis to multiple years and additional seasons would help determine whether the identified factor hierarchy remains stable under different snow, vegetation, and surface-moisture conditions. Transfer tests in other mountainous environments are also needed to assess the broader robustness of the final HistGradientBoosting model and to determine whether additional predictors, such as more detailed land-cover or canopy descriptors, are required in more complex surface settings. Future work could further incorporate explicit modelling of residual spatial structure, for example through geostatistical extensions or spatially varying coefficient approaches, in order to better characterize where the present segment-based model underperforms. Finally, the practical value of the framework would be strengthened by applying it to an operational DEM validation, DEM-fusion, or control-point screening workflow, so that the benefits of reliability-based ATL08 screening can be evaluated in a real mapping context.

5. CONCLUSIONS

This study developed a terrain-aware machine-learning framework for assessing and predicting ICESat-2 ATL08 absolute elevation error in complex hilly terrain. Using a 2 m GF-7 stereo DEM as reference in the Upper Yellow River Basin, China, ATL08 Version 006 terrain heights were evaluated after converting ATL08 ellipsoidal heights to orthometric heights for vertical-datum consistency. The main conclusions are as follows.

1. **ATL08 terrain elevations in the study area show clear and spatially structured errors.**

After quality screening, 15,904 ATL08 segments were retained. Relative to the GF-7 reference DEM, the mean absolute elevation difference was 14.08 m. ATL08 performance varies systematically with local terrain conditions, with relatively smaller errors in gentle terrain and larger errors in steep, heterogeneous areas.

2. **Terrain complexity is the primary control on ATL08 absolute elevation error, whereas signal-quality variables play secondary but meaningful roles.**

Geodetector analysis showed that terrain_slope and h_te_std had the strongest explanatory power, whereas snr, n_ca_photons, h_te_uncertainty, and h_te_interp_minus_median contributed additional but weaker effects. SHAP analysis of the final model produced a similar feature hierarchy, confirming that terrain-related predictors dominate the error structure.

3. **HistGradientBoosting provided the best overall predictive performance under spatially independent validation.**

Using a spatially blocked 5-fold cross-validation design based on approximately 10 km blocks, six regression models were compared. HistGradientBoosting achieved the best fold-averaged performance, with a mean test R^2 of 0.452 ± 0.016, mean RMSE of 10.65 ± 0.64 m, and mean MAE of 7.24 ± 0.40 m, indicating moderate but stable predictive skill under geographically unseen conditions.

4. **The final model is more suitable for terrain-aware screening than for exact point-level correction.**

Although HistGradientBoosting outperformed the other tested models, the achieved accuracy indicates meaningful but still moderate predictive skill. The model is therefore best interpreted as a reliability-screening tool that can distinguish relatively lower-risk and higher-risk ATL08 segments, rather than as a deterministic point-level correction model for recovering very small errors exactly.

5. **The proposed framework has practical value for DEM validation and ATL08 segment prioritization in complex terrain.**

By combining GF-7-based reference evaluation, Geodetector factor analysis, spatially blocked machine-learning prediction, and SHAP interpretation, the proposed framework provides a practical workflow for terrain-aware use of ATL08 data in mountainous environments. ATL08 segments predicted to have relatively low absolute elevation error may be prioritized for DEM assessment and regional elevation-quality screening, whereas segments predicted to have larger errors should be assigned lower confidence or excluded. With further testing across additional seasons and geomorphic settings, the framework may support broader geodetic and geomorphological applications.

ACKNOWLEDGMENTS

This work was supported by the National Basic Scale Topographic Map Surveying and Real-Scene 3D Construction Project (Grant No. 701013002). The authors thank the NASA National Snow and Ice Data Center (NSIDC) for providing ICESat-2 ATL08 data, and the relevant agency for providing GF-7 satellite imagery.

REFERENCES

Breiman, L.: 2001. Random forests. Mach. Learn., 45, 1, 5–32. DOI: 10.1023/A:1010933404324

Brunt, K.M., Neumann, T.A. and Smith, B.E.: 2019, Assessment of ICESat-2 Ice Sheet Surface Heights, based on comparisons over the interior of the Antarctic Ice Sheet. Geophys. Res. Lett., 46, 22, 13072–13078. DOI: 10.1029/2019GL084886

Carabajal, C.C. and Boy, J.-P.: 2020, I ICESAT-2 altimetry as geodetic control. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., 18, B3, 1299–1306. DOI: 10.5194/isprs-archives-XLIII-B3-2020-1299-2020

Cavanaugh, J.F.: 2019, The advanced topographic laser altimeter system (ATLAS) for nasa's ice cloud and land elevation satellite-2 (icesat-2): Design, development, and on-orbit performance. NASA Technical Reports Server (NASA). http://hdl.handle.net/2060/20190014038

Chen, K., Dai, W., Li, F., Li, S. and Wang, C.: 2025, Enhancing Large-Area DEM modeling of GF-7 stereo imagery: Integrating ICESat-2 data with Multi-characteristic constraint filtering and terrain matching correction. Int. J. Appl. Earth Obs. Geoinf., 138, 104485. DOI: 10.1016/j.jag.2025.104485

Chen, T. and Guestrin, C.: 2016, XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM. DOI: 10.1145/2939672.2939785

Dandabathula, G., Verma, M., Satyanarayana, P. and Srinivasa Rao, S.: 2020, Evaluation of ICESat-2 ATL08 Data Product: Performance assessment in inland water. Eur. J. Environ. Earth Sci., 1, 3. DOI: 10.24018/ejgeo.2020.1.3.15

Friedman, J.H.: 2001, Greedy Function Approximation: A gradient boosting machine. Ann. Stat., 29, 5, 1189–1232.

Guo, Y., Cheng, L., Ding, A., Yuan, Y., Li, Z., Hou, Y., Ren, L. and Zhang, S.: 2024, Geodetector model-based quantitative analysis of vegetation change characteristics and driving forces: A case study in the Yongding River basin in China. Int. J. Appl. Earth Obs. Geoinf., 132, 104027. DOI: 10.1016/j.jag.2024.104027

Haiyi, C., Fuqiang, L., Chenguang, Z. and Jun, D.: 2021, The study of high resolution stereo mapping satellite. Nat. Remote Sens. Bull., 25, 7, 1400–1410. DOI: 10.11834/jrs.20210411

Hoerl, A.E. and Kennard, R.W.: 1970, Ridge Rregression: Biased estimation for nonorthogonal Pproblems. Technometrics, 12, 1, 55–67. DOI: 1080/00401706.1970.10488634

Kandil, W.M., Zarzoura, F.H., Salah, M. and El-Mewafi, M.: 2024, New approach for satellite DEM accuracy enhancement by combing machine learning, fuzzy majority voting, and weighted interpolation techniques. Innov. Infrastruct. Solut., 9, 4. DOI: 10.1007/s41062-024-01401-z

Kurtz, N.T., Neumann, T. and Magruder, L.A.: 2019, ICESat-2 mission overview: Measuring the height of the earth with a photon-counting laser altimeter. AGU Fall Meeting Abstracts. https://ui.adsabs.harvard.edu/abs/2019AGUFM.C31C1522K/abstract

Li, B., Zhao, T., Su, X., Fan, G., Zhang, W., Deng, Z. and Yu, Y.: 2022, Correction of terrain effects on forest canopy height estimation using ICESat-2 and high spatial resolution images. Remote Sens., 14, 18, 4453. DOI: 10.3390/rs14184453

Li, G., Guo, J., Shang, J., Zhu, S. and Chen, J.: 2024, Qinghai-Tibet Plateau Glacier-DEM product derived from the Chinese stereo mapping satellites. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., 18, 1, 311–316. DOI: 10.5194/isprs-archives-XLVIII-1-2024-311-2024

Liang, Y. and Xu, C.: 2023, Knowledge diffusion of Geodetector: A perspective of the literature review and Geotree. Heliyon, 9, 9, e19651. DOI: 10.1016/j.heliyon.2023.e19651

Liu, A., Cheng, X. and Chen, Z.: 2021, Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote Sens. Environ., 264, 112571. DOI: 10.1016/j.rse.2021.112571

Liu, K., Liu, L., Luo, Z. and Gan, Y.: 2024, Study on the Application of Water Resource Management Based on GF-7 Stereo Mapping Satellite. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., 18, 1, 395–400. DOI: 10.5194/isprs-archives-XLVIII-1-2024-395-2024

Liu, Z., Zhou, S., Yu, H., Zhang, W., Guo, F., Chen and Guo, J.: 2022, Quantitative analysis of tectonic geomorphology research based on Web of Science from 1981 to 2021. Remote Sens., 14, 20, 5227. DOI: 10.3390/rs14205227

Magruder, L., Neuenschwander, A. and Klotz, B.: 2021, Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry. Remote Sens. Environ., 264, 112621. DOI: 10.1016/j.rse.2021.112621

Maxwell, A.E., Warner, T.A. and Fang, F.: 2018, Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens., 39, 9, 2784–2817. DOI: 10.1080/01431161.2018.1433343

Neuenschwander, A., Pitts, K., Jelley, B., Robbins, J., Markel, J., Popescu, S., Nelson, R., Harding, D., Pederson, D., Klotz, B. and Sheridan, R.: 2023, ATLAS/ICESat-2 L3A land and vegetation height, version 6. NASA National Snow and Ice Data Center Distributed Active Archive Center. DOI: 10.5067/ATLAS/ATL08.006

Neumann, T.A., Martino, A.J., Markus, T., Bae, S., Bock, M.R., Brenner, A.C., Brunt, K.M., Cavanaugh, J., Fernandes, S.T., Hancock, D.W., Harbeck, K., Lee, J., Kurtz, N.T., Luers, P.J., Luthcke, S.B., Magruder, L., Pennington, T.A., Ramos-Izquierdo, L., Rebold, T., … Thomas, T. C.: 2019, The ice, cloud, and land elevation satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System. Remote Sens. Environ., 233, 111325. DOI: 10.1016/j.rse.2019.111325

Pang, S., Li, G., Jiang, X., Chen, Y., Lu, Y. and Lu, D.: 2022, Retrieval of forest canopy height in a mountainous region with ICESat-2 ATLAS. For. Ecosyst., 9, 100046. DOI: 10.1016/j.fecs.2022.100046

Paul, F., Bolch, T., Briggs, K., Kääb, A., McMillan, M., McNabb, R., Nägler, T., Nuth, C., Rastner, P., Strozzi, T. and Wuite, J.: 2017, Error sources and guidelines for quality assessment of glacier area, elevation change, and velocity products derived from satellite data in the Glaciers_cci project. Remote Sens. Environ., 203, 256–275. DOI: 10.1016/j.rse.2017.08.038

Popescu, S.C., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L. and Walsh, K.M.: 2018, Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data. Remote Sens. Environ., 208, 154–170. DOI: 10.1016/j.rse.2018.02.019

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V. and Gulin, A.: 2018, CatBoost: Unbiased boosting with categorical features. Proceedings of the 32nd International Conference on Neural Information Processing Systems, 6639–6649.

Pronk, M., Eleveld, M. and Ledoux, H.: 2023, Assessing vertical accuracy and spatial coverage of ICESat-2 and GEDI spaceborne lidar for creating global terrain models. EarthArXiv (California Digital Library). DOI: 10.31223/X5309R

Purinton, B. and Bookhagen, B.: 2021, Beyond vertical point accuracy: Assessing inter-pixel consistency in 30 m Global DEMs for the Arid Central Andes. Front. Earth Sci., 9. DOI: 10.3389/feart.2021.758606

Rexer, M. and Hirt, C.: 2014, Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2.1/v4.1) and validation against accurate heights from the Australian National Gravity Database. Aust. J. Earth Sci., 61, 2, 213–226. DOI: 10.1080/08120099.2014.884983

Rumelhart, D.E., Hinton, G.E. and Williams, R.J.: 1986, Learning representations by back-propagating errors. Nature, 323, 6088, 533–536. DOI: 10.1038/323533a0

Schumann, G., Matgen, P., Cutler, M.E.J., Black, A., Hoffmann, L. and Pfister, L.: 2008, Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. ISPRS J. Photogramm. Remote Sens., 63, 3, 283–296. DOI: 10.1016/j.isprsjprs.2007.09.004

Tang, X., Xie, J., Liu, R., Hug, G., Zhao, C., Zhen, Y., Tang, H. and Dou, X.: 2020, Overview of the GF-7 Laser Altimeter System Mission. Earth Space Sci., 7, 1, e2019EA000777. DOI: 10.1029/20 19EA000777

Tian, X. and Shan, J.: 2021, Comprehensive evaluation of the ICESat-2 ATL08 terrain product. IEEE Trans. Geosci. Remote Sens., 59, 10, 8195–8209. DOI: 10.1109/TGRS.2021.3051086

Urbazaev, M., Hess, L.L., Hancock, S., Sato, L.Y., Ometto, J.P., Thiel, C., Dubois, C., Heckel, K., Urban, M., Adam, M. and Schmullius, C.: 2022, Assessment of terrain elevation estimates from ICESat-2 and GEDI spaceborne LiDAR missions across different land cover and forest types. Sci. Remote Sens., 6, 100067. DOI: 10.1016/j.srs.2022.100067

Valavi, R., Elith, J., Lahoz-Monfort, J.J., and Guillera-Arroita, G.: 2019, blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods Ecol. Evol., 10, 2, 225–232. DOI: 10.1111/2041-210X.13107

Wang, H., Qin, F., Xu, C., Li, B., Guo, L. and Wang, Z.: 2021, Evaluating the suitability of urban development land with a Geodetector. Ecol. Indic., 123, 107339. DOI: 10.1016/j.ecolind.2021.107339

Wang, J. and Xu, C.: 2017, Geodetector: Principle and prospective. Acta Geogr. Sin., 72, 1, 116–134. DOI: 10.11821/dlxb201701010

Xu, W., Li, J., Peng, D., Yin, H., Jiang, J., Xia, H. and Wen, D.: 2024, Vertical accuracy assessment and improvement of Five High-Resolution Open-Source Digital Elevation Models using ICESat-2 data and random forest: Case study on Chongqing, China. Remote Sens., 16, 11, 1903. DOI: 10.3390/rs16111903

Zhang, X., Guo, S., Yuan, B., Mu, H., Xia, Z., Tang, P., Fang, H., Wang, Z. and Du, P.: 2024, Error-Reduced Digital Elevation Model of the Qinghai-Tibet Plateau using ICESat-2 and Fusion Model. Sci. Data, 11, 1. DOI: 10.1038/s41597-024-03428-4

Zhu, X., Ren, Z., Nie, S., Bao, G., Ha, G., Bai, M. and Liang, P.: 2023, DEM generation from GF-7 Satellite Stereo Imagery assisted by Space-Borne LiDAR and its application to active tectonics. Remote Sens., 15, 6, 1480. DOI: 10.3390/rs15061480