Article | Received 12 April 2025; Accepted 2 July 2025; Published 15 July 2025
https://doi.org/10.55092/sc20250017

Exploring the potential of backpack SLAM LiDAR for metro tunnel inspection: explainable modeling and optimization of point cloud quality

Wenbo Qin[1,2], Shangbin Gao[1,2] and Cheng Zhou[1,2,*]

[1] National Center of Technology Innovation for Digital Construction, Huazhong University of Science and Technology, Wuhan, Hubei Province, China

[2] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei Province, China

* Correspondence author; E-mail: chengzhou@hust.edu.cn.

Highlights:

- Developed an explainable framework for modeling SLAM LiDAR point cloud quality.
- Performed 12 field experiments analyzing five key metro tunnel inspection factors.
- Speed and scan density were dominant factors across multiple quality metrics.

Abstract: Digital modeling of tunnels plays a critical role in the refined management of urban metro infrastructure. Backpack LiDAR systems, due to their portability and operational efficiency, show great potential for tunnel applications. However, the quality of point cloud data collected by such systems is often compromised by various factors during acquisition, leading to issues such as instability, uneven density, and structural distortion. To address these challenges, this study proposes a point cloud quality modeling framework that integrates statistical modeling with interpretable machine learning, and validates it through metro tunnel field experiments. Five key inspection factors were identified based on operational characteristics and tunnel environment. A total of 12 experimental groups were conducted, resulting in 360 point cloud samples. Four quality metrics were extracted and modeled using the Skew-Normal distribution to capture their statistical characteristics. A CatBoost regression model was then constructed to predict the distribution parameters, and the SHAP method was employed for global and local interpretability, revealing the causal pathways between inspection conditions and quality responses. The results indicate significant differences in how quality dimensions respond to inspection variables. Speed and scan density emerged as dominant factors across multiple metrics, while interaction terms had particularly strong effects on structure-related indicators. The proposed framework provides a quantifiable foundation for understanding point cloud quality, optimizing data acquisition strategies, which demonstrates strong engineering applicability and scalability.

Keywords: tunnel inspection; backpack; LiDAR; point cloud quality; interpretable machine learning; SHAP

1. Introduction

With the rapid expansion of urban metro transit networks, ensuring the structural safety of tunnels has become an increasingly critical challenge [1–3]. Tunnel structures are often subjected to complex environmental conditions, which lead to a variety of risks including structural deformation, water leakage, and lining cracks [4–6]. These issues demand efficient and fine-grained inspection technologies. As a key enabler of modern metro operation and maintenance, 3D digital inspection provides essential data support for structural recognition, and risk warning [7,8]. Compared with traditional manual or image-based inspection methods, 3D point cloud inspection offers higher precision, better spatial completeness, and richer geometric information, making it particularly suitable for the complex and enclosed environment of metro tunnels [9–11]. However, the success of such tasks fundamentally relies on the availability of high-quality point cloud data. Therefore, developing robust methods for point cloud quality assessment and enhancement has become a crucial prerequisite for the widespread application of 3D digital inspection in tunnel scenarios.

Current mainstream technologies for tunnel spatial information acquisition primarily include terrestrial laser scanning (TLS) and vehicle-mounted LiDAR systems. TLS systems have been widely used in tunnel surveying and deformation monitoring due to their ability to generate high-precision, high-density point clouds [12,13]. Chen *et al.* utilized TLS to achieve high-precision and automated estimation of circumferential seam misalignment, longitudinal seam misalignment, and ring ovality [14]. However, TLS typically requires manual deployment at discrete stations along the tunnel, followed by labor-intensive registration of multiple scans. Such operational requirements reduce inspection efficiency and hinder its applicability to large-scale or time-sensitive scenarios. Vehicle-mounted LiDAR systems enable efficient large-scale tunnel scanning by integrating LiDAR sensors onto mobile platforms such as inspection vehicles [15,16]. These vehicle systems have shown promise in high-speed tunnel modeling and automation scenarios, significantly improving data acquisition efficiency [17,18]. Cui *et al.* developed a mobile laser scanning system for subway shield tunnels, capable of rapidly acquiring 3D point cloud data to detect tunnel deformations, with millimeter-level accuracy under operational time constraints [19]. Kim and Choi proposed a 3D location estimation and tunnel mapping system for autonomous robots operating in underground space, utilizing ICP-based point cloud registration and multi-sensor fusion to achieve accurate trajectory tracking and mapping on inclined rampways [20]. Liao *et al.* developed a mobile measurement system combining spinning mid-range LiDAR and multi-cameras to reconstruct shield tunnel point clouds and generate unfolded panoramic images [21]. Nevertheless, their operation must be scheduled within narrow off-service maintenance windows. During these periods, the tunnel workspace is exclusively occupied by the scanning system, which restricts concurrent inspection or repair activities and introduces additional logistical complexity. Moreover, the use of such vehicle systems incurs high hardware and operational costs. These limitations highlight the need for more flexible, efficient, and minimally invasive scanning approaches for metro tunnel environments, especially under operational constraints.

Backpack SLAM LiDAR systems offer a practically promising alternative for tunnel inspection, particularly under the operational constraints of metro environments [22,23]. Compared to manual visual inspection, they provide substantially higher spatial accuracy and completeness, without requiring dedicated measurement windows or specialized vehicles. Compared with vehicle-based LiDAR platforms,

backpack SLAM LiDAR systems demonstrate greater adaptability in metro tunnel environments. Their compact design makes them especially effective for scanning areas that are difficult to access with larger equipment, such as cross-passages, ventilation shafts [24,25]. In practice, they can also operate within tight maintenance windows without occupying the entire tunnel section, which is particularly valuable in settings where inspection access is limited and multiple tasks must be coordinated [26]. Additionally, backpack systems offer lower operational costs and can be easily integrated into routine patrol workflows, making them a practical choice for frequent, fine-grained inspections in metro systems. These systems integrate LiDAR with inertial sensors and employ real-time SLAM algorithms to achieve trajectory reconstruction and point cloud generation in GNSS-denied environments, offering high efficiency and strong adaptability [27,28]. Compared to traditional trolley-based platforms, backpack systems demonstrate superior flexibility in narrow, curved, or densely equipped tunnel environments, thereby expanding the spatial coverage of inspection tasks. However, in practical deployments, point cloud quality generated by backpack systems is often degraded due to combined influences from walking speed, body movement, tunnel geometry, and scan configuration [29,30]. This quality problems manifest as sparsity, structural blurring, or geometric distortion, which significantly affect the reliability of downstream analysis and detection tasks [31]. Existing research has primarily focused on trajectory accuracy and mapping errors, with limited attention to quality evaluation metrics and quantitative models that capture the influence of inspection parameters on point cloud quality. This gap hinders theoretical guidance for task-specific parameter optimization and quality control. Furthermore, current approaches rarely provide interpretable causal modeling that connects inspection settings to point cloud quality and ultimately to application performance. Thus, their value in practical deployment optimization is limited.

To address these issues, this study proposes the following three main contributions for tackling point cloud quality challenges in SLAM-based tunnel scanning: (1) A four-dimensional quality evaluation framework based on the Skew-Normal distribution, covering roughness, density, planarity, and sphericity; (2) An interpretable regression modeling approach that combines CatBoost and SHAP analysis to quantitatively reveal how scanning conditions affect point cloud quality; (3) A series of field inspection experiments to systematically evaluate the effects of key factors on point cloud quality. The results of this study are expected to provide theoretical support and practical guidance for the deployment optimization and task adaptation of backpack SLAM LiDAR systems in complex tunnel environments.

2. Methodology

2.1. Overview of the point cloud quality modeling framework

This study proposes a comprehensive point cloud quality modeling and analysis framework tailored for backpack SLAM LiDAR inspection tasks in metro tunnels, as illustrated in Figure 1. The framework integrates design of experiment, point cloud preprocessing, quality modeling, and interpretability analysis, aiming to systematically characterize how inspection conditions influence point cloud quality, and to provide practical guidance for optimizing data collection strategies in operational applications. An orthogonal experimental design is employed to generate inspection schemes, enabling the collection of field point cloud data under various parameter combinations. After preprocessing the data, a set of quality metrics are extracted that capture both data completeness and geometric structure. These metrics are then subjected to statistical modeling to compress and quantify their distributional characteristics. In

the modeling stage, a CatBoost regression model with embedded interpretability is employed to establish the mapping between inspection conditions and point cloud quality parameters. The model exhibits strong nonlinear modeling capabilities and incorporates SHAP-based feature attribution analysis, which enables interpretation of the direction and relative importance of each input factor in the prediction process. This approach ensures that the point cloud quality modeling maintains high predictive accuracy while also providing transparent and traceable physical insights, offering a unified analytical pathway and theoretical support for understanding the performance of backpack SLAM LiDAR systems under varying operational conditions.

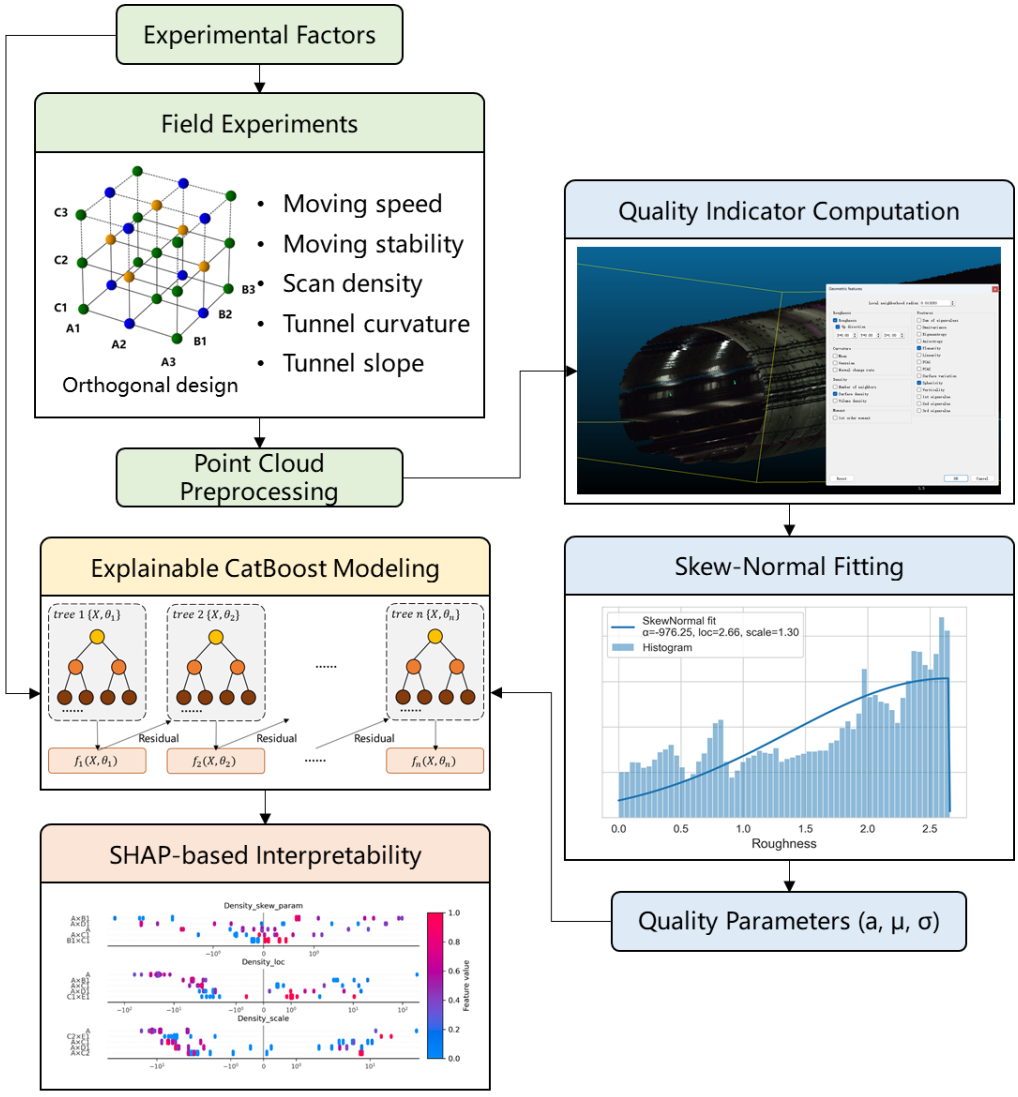

Figure 1. Overview of the proposed framework for point cloud quality explainable modeling.

2.2. Experimental factors and quality feature definition

With the increasing availability of commercial backpack SLAM LiDAR systems, such devices are now can be widely used for tasks like tunnel inspection. However, most of these systems rely on closed-source SLAM algorithms, where internal settings cannot be accessed or modified by users. In practice, only external factors can be adjusted during data collection. Therefore, this study focuses on modeling the impact of such controllable conditions on point cloud quality, offering practical guidance for real-world

deployment. To systematically analyze the variations in point cloud quality of backpack SLAM LiDAR systems under different inspection conditions, five key experimental factors were identified based on device operating characteristics and the practical tunnel environment. These factors encompass two major categories: The operation-related factors include moving speed, moving stability, and scan density, all of which are controllable and adjustable parameters during data collection. The tunnel-related factors consist of tunnel curvature and tunnel slope, representing typical geometric variations encountered in tunnel environments. To ensure both representativeness and statistical independence of the experimental combinations, a two-level design was adopted, and sample collection was conducted through on-site data acquisition. The selected factors and their corresponding level settings are summarized in Table 1.

Table 1. Experimental factors and levels.

No.	Experimental Factors	Level 1	Level 2
A	Moving speed	Normal walk	Brisk walk
B	Moving stability	Flat track	Between rails
C	Scan density	1 Scan	2 Scans
D	Tunnel curvature	Straight section	Curved section
E	Tunnel slope	Gentle slope	Steep Slope

The five experimental factors may influence point cloud quality through different mechanisms during the data acquisition process. Excessive moving speed can lead to reduced overlap between successive point cloud frames, resulting in increased motion estimation errors in the SLAM system. The accumulation of errors may further induce trajectory drift and registration deviations, ultimately affecting point cloud density distribution, surface smoothness, and structural reconstruction accuracy. Moving stability is influenced by gait and ground conditions. In particular, uneven terrains such as walking between railway tracks can induce slight body shaking, which introduces fluctuations in LiDAR pose estimation and reduces registration precision. Regarding scan density, single-pass acquisitions are often subject to occlusions or limited rearward field of view, potentially leading to blind spots. Increasing the number of scans enhances area overlap, thereby improving point cloud completeness and the fidelity of structural details. Tunnel curvature introduces challenges related to viewpoint shifts and loop closure in low-texture regions, which may result in accumulated trajectory drift and local reconstruction errors. Tunnel slope affects device pose stability; during uphill or downhill movement, inaccuracies in gravity direction estimation and rapid changes in scanning angle may disrupt spatial consistency in the point cloud. These mechanisms collectively constitute the primary pathways through which the experimental factors influence point cloud quality.

To evaluate point cloud quality, four representative metrics were selected, corresponding to three key dimensions: geometric stability, data completeness, and structural reconstruction capability [32]. The definitions and physical interpretations of these metrics are presented in Table 2. These indicators are capable of characterizing local surface morphology and distribution properties of the point cloud, offering strong engineering interpretability and general applicability. Each point cloud quality metric is computed based on geometric analysis within a fixed local neighborhood. Roughness is quantified as

the mean distance from points to a best-fit plane, capturing surface fluctuations. Density is defined as the number of points within a unit volume, representing the spatial distribution of data. Planarity and sphericity are derived from the eigenvalues of the covariance matrix computed via principal component analysis (PCA), and expressed as ratio-based indicators reflecting planar and volumetric structural characteristics [33,34].

Table 2. Definition of point cloud quality dimensions and descriptors.

Point Cloud Feature	Quality Dimension	Description
Roughness	Geometric Stability: Reflects how smooth or irregular a surface is based on local geometry	Mean distance from a point to its fitted plane within the local neighborhood $R = \frac{1}{N}\sum_{i=1}^{N} d_i$ Where N is number of neighboring points; d_i is orthogonal distance from point i to the fitted local plane.
Density	Data Integrity: Indicates how completely the area has been scanned by measuring the concentration of points	Number of points per unit volume or within a local neighborhood $D = N/V$ Where V is volume of neighborhood.
Planarity	Structural Recoverability (Surface): Measures how well the local point distribution fits a flat surface	Ratio of variance in the principal direction to total variance $P = \frac{\lambda_2 - \lambda_3}{\lambda_1}$ Where $\lambda_1 \geq \lambda_2 \geq \lambda_3$, are eigenvalues from PCA of local neighborhood.
Sphericity	Structural Recoverability (Volumetric): Assesses whether the local point distribution is balanced in all directions, resembling a sphere	Degree to which the variance along the three principal axes approximates a sphere $S = \frac{\lambda_3}{\lambda_1}$

Roughness measures the smoothness of the point cloud surface. Elevated roughness values typically indicate device jitter or trajectory instability, which may introduce mapping errors in the SLAM process and result in local surface irregularities or discontinuities. Density reflects the number of points within a unit volume. Lower density often arises from occlusions, blind spots, or loop closure difficulties during acquisition, potentially leading to uneven point distribution and missing geometric information. Planarity is used to assess the accuracy of planar structure reconstruction, such as tunnel linings. A low planarity score usually indicates the presence of SLAM drift or registration errors that compromise the correct reconstruction of planar features. Sphericity captures directional consistency from the perspective of volumetric structure. Significantly low sphericity may suggest that the point cloud is flattened or overly directional, usually caused by pose estimation errors or limited scanning trajectories. Conversely, excessively high sphericity may reflect interference from environmental noise or abnormal distributions of non-structural points.

2.3. Point cloud quality indicators computation

By jointly modeling these four features, the quality of point clouds acquired by backpack SLAM LiDAR systems under varying inspection conditions can be comprehensively characterized. In practice, point cloud quality metrics such as roughness, planarity, and density often display non-Gaussian behaviors,

including asymmetry, long tails, and local discontinuities. These problems due to factors like SLAM drift, occlusion, surface irregularities, or abrupt changes in motion, which make standard normal distributions insufficient for accurate modeling. To address this, the Skew-Normal distribution is adopted, a flexible extension of the normal distribution that introduces a shape parameter to explicitly capture skewness in the data. This allows the model to reflect the direction and degree of asymmetry, while preserving interpretability through its location and scale parameters. Compared to other nonparametric or heavy-tailed models, the Skew-Normal distribution provides a parsimonious yet expressive framework for describing distributional patterns across diverse tunnel scenarios. the parameterization facilitates regression modeling and quality interpretation.

The probability density function (PDF) of the Skew-Normal distribution is defined as follows:

$$f(x;\mu,\sigma,a) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(a\cdot\frac{x-\mu}{\sigma}\right) \quad (1)$$

Where μ is the location parameter, σ is the scale parameter, a is the shape (skewness) parameter, $\phi(\cdot)$ denotes the standard normal probability density function, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

Each point cloud quality metric was individually fitted using the Skew-Normal distribution to obtain the corresponding parameter triplet (a, μ, σ), which serves as a compact representation of the metric's overall behavior under specific experimental conditions. All distribution parameters, including a, were estimated from the empirical data using maximum likelihood estimation (MLE), without preset values. The Skew-Normal distribution compresses the complex distributional shape of the original metric, and retains essential statistical characteristics such as central tendency, dispersion, and skewness direction, which provides a unified parametric form for subsequent regression modeling and interpretability analysis. Figure 2 presents representative examples of the histograms and fitted curves for the four selected metrics. It can be observed that roughness and sphericity demonstrate asymmetric long tails, indicating the presence of localized noise or structural irregularities. Planarity displays a high mean and right-skewed shape, suggesting occasional highly planar regions amid curved surfaces. Density appears broadly symmetric, reflecting generally uniform sampling. These distributional patterns reflect the spatial heterogeneity and statistical skewness of point cloud quality under complex tunnel environments and dynamic inspection conditions. The Skew-Normal distribution is shown to be effective in capturing these characteristics.

Table 3 further summarizes the meanings and interpretations of the Skew-Normal distribution parameters for each of the four quality metrics. In the case of roughness, a positive shape parameter $a > 0$ indicates the presence of localized regions with noise or registration errors within an otherwise stable surface. An increase in the scale parameter σ reflects greater surface fluctuation, which may result from motion instability or occlusions. The location parameter μ represents the baseline level of overall surface roughness. Similarly, variations in the parameters of density, planarity, and sphericity can be mapped to spatial characteristics of point cloud quality and the stability of SLAM-based mapping. the interpretations enable a more comprehensive understanding of how inspection conditions influence geometric integrity and reconstruction fidelity.

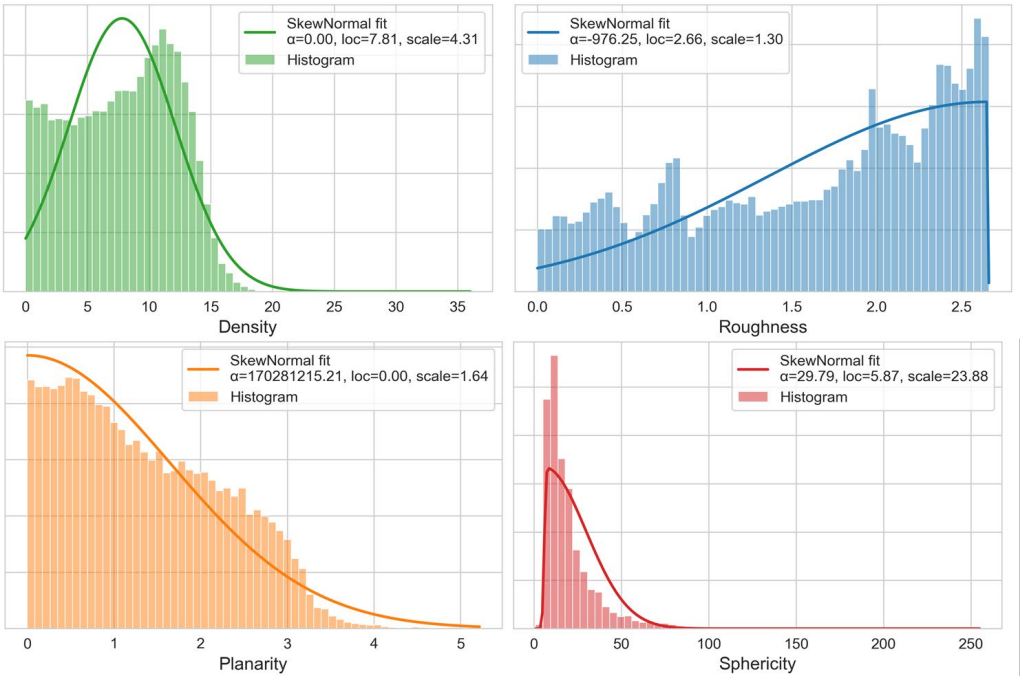

Figure 2. Point cloud quality indicators fitting results.

Table 3. Interpretation of Skew-Normal distribution parameters.

Indicator	Parameter	Interpretation
Roughness	Shape (a)	Positive values indicate a tail toward high roughness, suggesting local noise or registration errors.
	Location (μ)	Larger values reflect greater surface variation, indicating unstable motion or reconstruction inconsistency.
	Scale (σ)	High skewness implies outliers with excessive roughness, often due to vibration or occlusion.
Density	Shape (a)	Negative values suggest sparse regions exist despite overall high density, indicating potential occlusion or dropouts.
	Location (μ)	Higher values indicate uneven point distribution, reflecting inconsistent coverage or scanning paths.
	Scale (σ)	High skewness highlights localized sparsity, potentially caused by fast motion or poor registration.
Planarity	Shape (a)	Positive values indicate a few highly planar areas, such as floors or track surfaces.
	Location (μ)	Larger values reflect high variability in planar structures, possibly due to drift or curved surfaces.
	Scale (σ)	High skewness indicates structural degradation or poorly reconstructed regions.
Sphericity	Shape (a)	Positive values suggest scattered or isotropic point patterns, possibly from noise or SLAM failure.
	Location (μ)	Larger values imply structural inconsistency, such as compression, stretching, or dispersion.
	Scale (σ)	High skewness denotes the presence of noisy, non-structural points.

2.4. Explainability CatBoost regression model

To model the nonlinear relationship between inspection conditions and point cloud quality parameters, the CatBoost regression model was employed. As an enhanced implementation of gradient boosting decision trees, CatBoost offers strong fitting capability, robust generalization performance, and intrinsic model interpretability [35]. It is particularly well-suited for structured modeling tasks characterized by limited feature dimensionality and complex variable dependencies. CatBoost efficiently handles this mixed data type structure (numerical and categorical features) without requiring extensive preprocessing. Its ordered target encoding for categorical variables helps reduce overfitting, which is critical given the relatively small size and experimental nature of our dataset. Furthermore, the relationship between inspection conditions and quality parameters involves nonlinear interactions, which CatBoost models more effectively than linear regression or shallow tree-based methods. Moreover, CatBoost integrates seamlessly with the SHAP interpretability framework. This allows to quantify the marginal effect of each input variable on predicted quality parameters in a consistent and theoretically grounded manner, supporting both global attribution and local diagnosis. This capability is especially valuable in practical settings, where understanding the influence of specific inspection choices is essential for field deployment optimization. Figure 3 illustrates the architecture of the interpretable CatBoost regression model developed in this study, which integrates input feature processing, residual-based decision tree modeling, and SHAP-based interpretability mechanisms. The model takes five inspection control factors as inputs. Among them, walking speed and scan density are numerical features and can be directly used as model inputs. In contrast, movement stability, tunnel curvature, and tunnel slope are categorical variables, which are processed using CatBoost's specialized ordered target statistics encoding method. The encoding strategy leverages historical statistics of the target variable to construct feature values. The general form of ordered target encoding is given as follows:

$$V_E = \frac{\sum y + \alpha \cdot p}{n + \alpha} \tag{2}$$

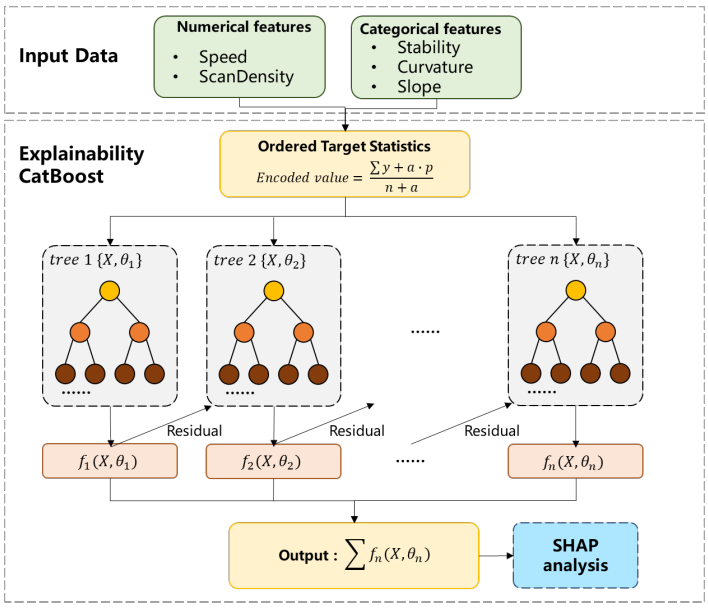

Figure 3. Framework of explainability CatBoost regression model.

Where V_E is the encoded value, $\sum y$ represents the cumulative target value of previously observed samples within the same category, p denotes the global prior expectation of the target variable, n is the number of observed samples in the current category, and α is a smoothing parameter.

Compared to traditional one-hot encoding, this method significantly reduces the risk of dimensionality explosion and overfitting, while maintaining robustness under conditions of sparse or imbalanced categorical variables.

The central part of Figure 3 illustrates the core boosting structure of the CatBoost model. In each iteration, a regression tree is constructed to fit the residuals of the previous ensemble, resulting in a sequence of weak predictors $f_m(x)$. The final prediction is obtained by aggregating the weighted outputs of all trees. At the output stage, a SHAP-based interpretability module is integrated to perform factor attribution analysis for each prediction. At the mathematical level, the basic prediction function of CatBoost can be expressed as follows:

$$\hat{y}(x) = \sum_{m=1}^{M} \gamma_m \cdot f_m(x) \tag{3}$$

Where $f_m(x)$ represents the m-th regression tree, γ_m is its learning rate (or scaling factor), and M is the total number of trees. The model iteratively fits the residuals of the loss function in each boosting round to progressively optimize the prediction accuracy. The residuals at iteration m are computed as follows:

$$r_m^{(i)} = -\left[\frac{\partial L\left(y_i, \hat{y}_{m-1}(x_i)\right)}{\partial \hat{y}_{m-1}(x_i)}\right] \tag{4}$$

Where $L(\cdot)$ denotes the loss function, and $r_m^{(i)}$ represents the residual of sample i at iteration m.

To enable unified modeling of the Skew-Normal distribution parameters, a multi-output regression approach was adopted, in which the three target variables (a, μ, σ) are jointly represented as a single output vector:

$$\hat{Y}(x) = \begin{bmatrix} \hat{a}(x) \\ \hat{\mu}(x) \\ \hat{\sigma}(x) \end{bmatrix} = \sum_{m=1}^{M} \gamma_m \cdot \begin{bmatrix} f_m^{(a)}(x) \\ f_m^{(\mu)}(x) \\ f_m^{(\sigma)}(x) \end{bmatrix} \tag{5}$$

To enhance model transparency and interpretability, CatBoost integrates the SHAP (SHapley Additive Explanations) mechanism [36], which is a game-theoretic approach that quantifies the marginal contribution of each input feature to the model's prediction. SHAP formulates the model output as a linear summation of individual feature contributions:

$$\hat{y}(x) = \phi_0 + \sum_{i=1}^{n} \phi_i(x_i) \tag{6}$$

Where ϕ_0 denotes the global baseline value, and $\phi_i(x_i)$ represents the marginal contribution of the i-th input feature for the current sample, satisfying the following property:

$$\sum_{i=1}^{n} \phi_i(x_i) = \hat{y}(x) - \phi_0 \tag{7}$$

The theoretical definition of each SHAP value is given as:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(n - |S| - 1)!}{n!} \left[f_{S \cup \{i\}}(x) - f_S(x) \right] \tag{8}$$

Where N denotes the full set of input features, S represents all possible subsets that do not contain feature i, and $f_S(x)$ is the model output when trained using only the features in subset S.

Therefore, the interpretable CatBoost regression model enables joint modeling and accurate prediction of multidimensional point cloud quality parameters, and quantifies the marginal contributions of individual inspection factors through the integrated SHAP mechanism. Through this process, the model provides a quantitative understanding of how different inspection conditions influence point cloud quality, offering a reliable data-driven foundation for subsequent mechanism identification and operational strategy optimization through causal inference.

3. Field tests

3.1. Description of test tunnel and equipment

This study conducted a case study based on the Jiangji Tunnel section of Metro Line 2 in Wuhan, China. The Jiangji Tunnel is located between Jianghan Road Station and Jiyuqiao Station, with a total length of approximately 3.322 km. The tunnel was constructed using the shield tunneling method and traverses the Yangtze River. To perform the point cloud quality modeling experiments, four test sections were chosen along the tunnel, as shown in Figure 4. These sections were selected based on orthogonal combinations of distinct geometric characteristics, with detailed parameters summarized in Table 4. The selected sections included representative variations in tunnel morphology, such as straight and curved sections, as well as gentle and steep slopes, in order to evaluate the inspection system's modeling capability and the stability of point cloud quality under different spatial scenarios.

Figure 4. layout of test sections in a metro tunnel.

Table 4. Geometric characteristics of the test tunnel sections.

Test Section	Slope	Curvature	Distance Range
Section 1	25.8°	Straight	DK14 + 363-DK14 + 820
Section 2	25.8°	R360	DK13 + 819-DK14 + 363
Section 3	2.0°	R312	DK13 + 629-DK13 + 819
Section 4	3.5°	Straight	DK12 + 800-DK13 + 629

The backpack SLAM LiDAR system used in this study is a commercial device (OmniSLAM R6), as shown in Figure 5. The system integrates a rotating LiDAR scanner, a 360° panoramic camera, an RTK-SLAM positioning module, and an embedded processing and power unit, making it well-suited for mobile inspection tasks in narrow and complex tunnel environments. The LiDAR module features a 32-beam structure, with a maximum range of 120–300 meters and a peak point rate of up to 640,000 points per second, enabling high-density spatial sampling (up to 10,000 points/m²). Combined with a SLAM algorithm and inertial navigation fusion (LIO), the system achieves centimeter-level trajectory estimation and point cloud registration. The absolute measurement accuracy is 3 cm, with relative accuracy better than 1 cm. The imaging unit includes two 1-inch CMOS sensors equipped with 360° panoramic lenses. Colorized point cloud reconstruction is achieved by projecting texture information from panoramic images onto the 3D surface, thereby enhancing both the semantic expressiveness and visualization quality of the point cloud. During testing, the operator carried the system through each segment, acquiring approximately 50 meters of tunnel point cloud data per section. The system automatically handled sensor synchronization, trajectory estimation, and point cloud generation throughout the inspection process.

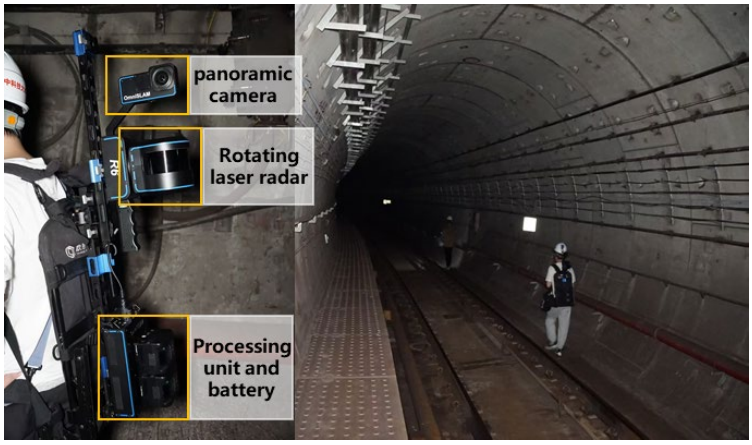

Figure 5. Illustration of the backpack SLAM LiDAR system and its deployment.

3.2. Design of experiment

A total of 12 field experiments were designed and conducted across four representative sections of the Jiangji Tunnel, in order to systematically investigate the influence mechanisms of different inspection conditions on point cloud quality. The experimental design balanced operational feasibility and variable

independence by adopting an orthogonal experimental approach in combination with an extended control group strategy. Such a strategy provided comprehensive coverage of inspection factors and strong comparability between experimental groups.

Five key inspection factors affecting point cloud quality were selected as experimental variables, each set at two levels. Specifically, moving speed was controlled by adjusting the operator's pace, with levels defined as "walk" (~ 4.2 km/h) and "brisk walk" (~ 5.4 km/h). Moving stability was modulated via walking path, categorized as "flat track" and "track between rails". Scan density was set as "single scan" or "double scan" depending on whether the tunnel was passed once or twice. Tunnel geometric parameters were based on actual section structures, including curvature (straight or R312/R360 curved segments) and longitudinal slope (ranging from 2.0° to 25.8°).

On this basis, an additional four extended experiments (No. 9–12) were designed to isolate and evaluate the effect of moving speed as a single variable under fixed spatial conditions. These tests were all conducted within Section 4 (straight + gentle slope), with moving speed varied continuously from 2 km/h to 8 km/h, covering stroll, walk, brisk walk, and jogging. A comparative group was thereby constructed for analyzing speed-dependent variation in point cloud quality.

The complete configuration of experimental groups and factor levels is summarized in Table 5. Actual moving speeds were recorded using timing and trajectory evaluation. To ensure experimental consistency, the same inspection system and operator were used throughout, with fixed LiDAR and IMU configurations.

To minimize external interference, all experiments were conducted during non-operational nighttime hours under stable environmental conditions, with no ongoing construction or equipment activity inside the tunnel. Point clouds were segmented and indexed by complete trajectory sections, resulting in a structured and well-controlled experimental dataset for subsequent quality feature extraction and modeling analysis.

Table 5. Design of field experiments and test conditions.

No.	Test Section	Speed (Measured speed)	Stability	Scan Density	Curvature	Slope
1	Section 4	Walk: 4.2 km/h	flat	1 scan	straight	gentle
2	Section 2	Walk: 4.2 km/h	track	2 scans	curved	steep
3	Section 3	Brisk walk: 5.4 km/h	flat	2 scans	curved	gentle
4	Section 1	Brisk walk: 6 km/h	track	1 scan	straight	steep
5	Section 1	Walk: 4.5 km/h	flat	2 scans	straight	steep
6	Section 3	Walk: 4 km/h	track	1 scan	curved	gentle
7	Section 2	Brisk walk: 5.6 km/h	flat	1 scan	curved	steep
8	Section 4	Brisk walk: 5.2 km /h	track	2 scans	straight	gentle
9	Section 4	Brisk walk: 6 km/h	flat	1 scan	straight	gentle
10	Section 4	Walk: 3.6 km/h	flat	1 scan	straight	gentle
11	Section 4	Jog: 8 km/h	flat	1 scan	straight	gentle
12	Section 4	Stroll: 2 km/h	flat	1 scan	straight	gentle

3.3. Point cloud preprocessing

The standardized and preprocessed point cloud data were used to extract quality metrics and construct modeling features, aimed at characterizing geometric quality variations of tunnel point clouds under different inspection conditions. The overall workflow involved three main stages: segmentation of point cloud samples, computation of quality metrics, and fitting of distribution parameters. Based on uniformly filtered and cropped point clouds, each main experimental dataset was further segmented into approximately 30 structurally complete and spatially consistent segment ring samples, with each sample covering a full tunnel segment ring. Overlapping or damaged regions at the segment boundaries were excluded, resulting in a total of 360 high-quality point cloud samples for subsequent feature analysis. Four quality metrics were computed for each sample using CloudCompare software. As shown in Figure 6, each metric was calculated at the point level to form a scalar field, and its spatial distribution characteristics were visualized accordingly. These four indicators capture different aspects of point cloud quality. To avoid altering the intrinsic spatial characteristics of data, no resampling or smoothing process was performed. Only obvious outliers and non-tunnel artifacts were manually removed. When computing local descriptors in CloudCompare, the neighborhood radius was set using the software's default estimation method, which adapts to point density and ensures stable, unbiased results across all samples.

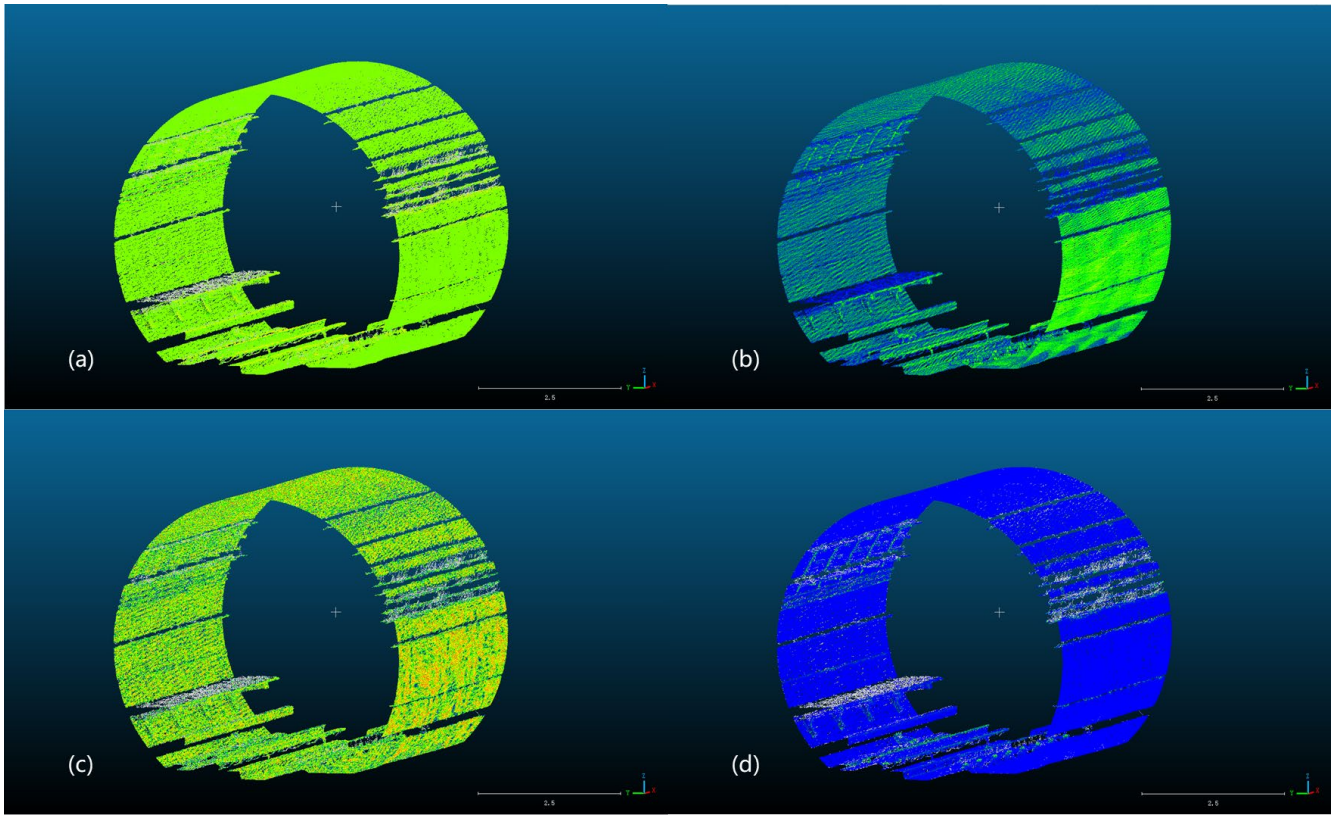

Figure 6. Visualization of point cloud quality metrics using scalar fields: **(a)** Roughness; **(b)** Planarity; **(c)** Sphericity; **(d)** Point density.

Table 6 summarizes the mean values of the Skew-Normal distribution parameters fitted to each quality metric across the 12 experimental groups. The results reveal substantial differences in distribution characteristics under varying inspection conditions.

Table 6. Computed Skew-Normal parameters for each test sample.

Group	Roughness	Density	Planarity	Sphericity
test01	(2.00e+8, 1.01e−3, 1.48e+0)	(3.06e+8, 1.07e−3, 8.50e+0)	(−3.29e+1, 3.05e+0, 1.68e+0)	(2.91e+1, 3.71e+0, 2.08e+1)
test02	(−5.75e+2, 2.46e+0, 1.20e+0)	(−5.89e−1, 9.56e+0, 4.79e+0)	(1.37e+8, 1.17e−3, 4.26e+0)	(3.09e+1, 5.90e+0, 2.39e+1)
test03	(−7.95e+2, 2.41e+0, 1.18e+0)	(1.12e+3, 2.49e−1, 9.55e+1)	(2.53e+8, 1.00e−3, 1.57e+0)	(2.36e+1, 4.68e+0, 1.90e+1)
test04	(2.63e+0, 1.65e+0, 3.90e+0)	(8.70e−3, 3.04e+1, 1.67e+1)	(−2.57e−1, 2.59e+0, 1.40e+0)	(2.89e+1, 4.67e+0, 2.27e+1)
test05	(9.11e+0, 5.26e−1, 4.45e+0)	(9.09e−1, 1.62e+1, 1.80e+1)	(2.32e−1, 1.39e+0, 9.46e−1)	(1.41e+0, 7.19e+1, 3.38e+1)
test06	(1.94e+8, 1.00e−3, 7.06e-1)	(1.41e−1, 2.56e+1, 1.51e+1)	(1.05e+8, 1.04e-3, 1.61e+0)	(1.61e+1, 4.60e+1, 7.23e+1)
test07	(−1.11e+3, 3.05e+0, 1.53e+0)	(1.57e+8, 1.57e−2, 4.44e+1)	(1.85e+8, 1.00e−3, 1.64e+0)	(2.64e+1, 4.67e+1, 1.14e+2)
test08	(4.60e−1, 2.98e+1, 2.55e+1)	(5.40e−1, 1.40e+2, 1.34e+2)	(2.35e+0, 5.26e−1, 1.27e+0)	(1.35e+1, 4.29e+1, 6.29e+1)
test09	(5.29e+6, 1.72e−2, 2.87e+0)	(4.51e−3, 2.63e+1, 1.70e+1)	(−1.77e−1, 1.91e+0, 1.06e+0)	(1.95e+1, 5.35e+0, 1.97e+1)
test10	(7.76e+8, 1.04e−3, 2.48e+1)	(1.24e+9, 1.20e−3, 8.78e+1)	(5.44e+7, 1.04e−3, 1.77e+0)	(3.15e+1, 4.17e+0, 2.42e+1)
test11	(−4.17e−1, 2.01e+1, 1.10e+1)	(−1.35e−1, 2.42e+1, 1.44e+1)	(−2.00e+0, 2.14e+0, 1.13e+0)	(2.34e+1, 5.51e+0, 2.70e+1)
test12	(−2.04e+2, 5.56e+1, 2.61e+1)	(−1.73e+2, 4.01e+2, 1.80e+2)	(2.39e+8, 1.00e−3, 2.53e+0)	(1.86e+1, 7.29e+0, 1.91e+1)

Among the first eight orthogonal experimental groups, test01, characterized by low walking speed and stable operation, exhibits relatively concentrated distributions in all metrics. Although the shape parameters for roughness and density are large due to numerical instability, the overall distributions remain unimodal with moderate values of μ and σ, indicating stable mapping performance. In contrast, test02 and test03, both conducted in sections with pronounced curvature or slope and unstable walking conditions, show clear left-skewness in roughness and significantly increased dispersion in density and planarity, reflecting potential risks of registration error and structural drift. Test04 also demonstrates amplified fluctuations in roughness and sphericity, confirming that excessive speed under sloped conditions suppresses SLAM mapping accuracy.

Under the combined condition of moderate walking speed and high overlap, test05 achieved favorable distribution patterns in density and structural metrics, with increased μ and reduced σ in planarity, validating the effectiveness of repeated scanning in enhancing geometric detail reconstruction. Although test06 was conducted at a low walking speed, the unstable path (track between rails) caused pose disturbances, which manifested as extreme shape values and decreased μ in roughness and sphericity, highlighting the critical role of movement stability in quality control. Test07 and test08 represent contrasting combinations: “fast & stable” and “double scan & unstable,” respectively. While test07 retained acceptable structural consistency, the density metric showed significantly low mean and high dispersion, indicating data sparsity in some regions. Test08 exhibited substantial skewness and dispersion across multiple metrics, revealing that high overlap alone cannot compensate for instability during inspection, despite the elevated mean density.

To isolate the effect of walking speed on point cloud quality, four additional speed-controlled experiments (test09–12) were conducted within Section 4, with all other variables held constant. The results indicate a nonlinear relationship between speed and quality. Within the 4.0–5.0 km/h range (test09, test10), quality metrics such as roughness and planarity remained well-centered, with stable μ and σ, suggesting this range as an optimal operational interval. In contrast, both very slow (2 km/h, test12) and very fast (8 km/h, test11) movements led to degraded quality: the former caused localized stacking and density anomalies, while the latter resulted in registration drift and blurred structural reconstruction, as evidenced by expanded and skewed distributions in density and sphericity.

In summary, the parameter features derived from Skew-Normal fitting effectively capture distributional differences in point cloud quality under varying inspection conditions. These features further support interpretable modeling and quantitative analysis of individual factor effects, providing a robust foundation for mechanism understanding and inspection strategy optimization.

3.4. Point cloud quality modeling

After completing the Skew-Normal fitting of all 360 point cloud samples, the resulting parameter sets were used as modeling outputs. Specifically, the distribution parameter triplets were used to construct a regression task in which the inspection control factors serve as inputs and the distribution parameters as outputs. Each sample included five input variables and twelve output variables, constituting a multi-output regression problem. During training, CatBoost first automatically encodes categorical variables using ordered target statistics and ranks numerical variables for optimal splitting. It then iteratively constructs a series of symmetrically structured regression trees. In each boosting iteration, a new tree is fitted to the residuals from the previous round, specifically targeting the residuals of the multi-output parameters. These residuals are progressively accumulated to form the final prediction.

The CatBoost regression model was implemented using the official open-source Python library. To balance model complexity and generalization, the following hyperparameters were applied: depth = 8, learning rate = 0.05, and number of iterations = 1000. Early stopping was used with a patience of 50 rounds to prevent overfitting during training. The hyperparameters were selected based on preliminary tests and prior experience, without extensive tuning, as the focus of this study lies in analyzing quality factor relationships rather than maximizing predictive accuracy.

Figure 7 illustrates the structure of a typical regression tree within the CatBoost model. Each non-leaf node represents a split condition that directs the sample down different branches based on the values of input variables. In the illustrated example, walking speed serves as a primary decision variable, in combination with conditions such as “whether the slope is gentle” or “whether the section is curved”, to hierarchically classify sample characteristics. Each path encodes a specific combination of inspection conditions and leads to a corresponding leaf node, which provides a numerical output—representing the predicted increment for the target parameter in that tree. In CatBoost’s gradient boosting framework, the overall prediction is computed as the weighted sum of outputs from all trees’ leaf nodes. Although each individual tree contributes only a local increment, its internal structure clearly reflects the model’s decision logic and interactions among variables, thereby facilitating interpretation of the modeling process.

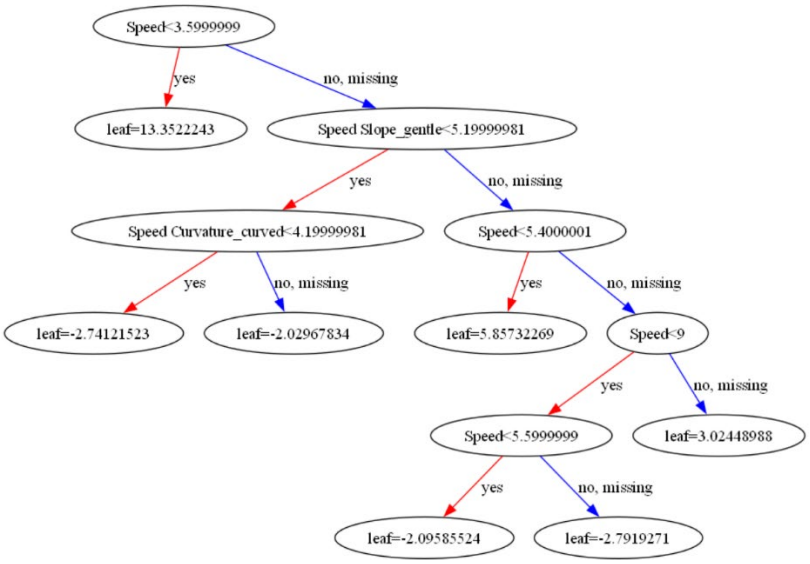

Figure 7. CatBoost decision tree for point cloud quality metrics.

To assess the predictive capability of the proposed model in practical tasks, a five-fold cross-validation strategy was adopted for training and testing. For each output dimension, the average coefficient of determination (R^2) and Relative Mean Absolute Error (RMAE) were recorded. The results are summarized in Table 7. The CatBoost model demonstrated consistently high predictive accuracy and robustness across all output dimensions. Among the quality metrics, sphericity and density achieved the best fitting performance, with both location and scale parameters exhibiting R^2 values exceeding 0.998 and RMAE values below 0.01. These results indicate that these metrics are highly responsive to changes in inspection conditions and are well-suited for regression modeling. In comparison, the prediction errors for the shape parameter (α) were slightly higher, reflecting its greater sensitivity to distribution tails. However, this also provides enhanced contrast in subsequent interpretability analysis by highlighting structural variations in quality distributions. Overall, the constructed CatBoost-based multi-output regression framework effectively maps inspection control variables to distributional characteristics of point cloud quality. The model exhibits strong regression performance and inherent structural interpretability, providing a solid foundation for variable attribution and inspection strategy optimization.

Table 7. Performance metrics of the explainable CatBoost regression model.

Parameter		RMAE	R^2
Roughness	Shape (a)	0.9411	0.1737
	Location (μ)	0.9852	0.0563
	Scale (σ)	0.9907	0.0298
Density	Shape (a)	0.9511	0.1426
	Location (μ)	0.9997	0.0114
	Scale (σ)	0.9984	0.0086
Planarity	Shape (a)	0.9602	0.0928
	Location (μ)	0.9680	0.0885
	Scale (σ)	0.9873	0.0181
Sphericity	Shape (a)	0.9990	0.0054
	Location (μ)	0.9999	0.0020
	Scale (σ)	0.9999	0.0042

4. Results and discussion

4.1. Global contribution patterns of experimental factors

To comprehensively evaluate the overall influence of inspection control factors across different point cloud quality dimensions, SHAP values were extracted from the trained CatBoost model for each output parameter. The absolute mean SHAP value across all samples was used as the indicator of feature importance. This process quantifies the average marginal impact of each input variable on the model's predictions, thereby enabling a global ranking of variable contributions. For each quality metric, SHAP means were computed separately for the three Skew-Normal distribution parameters (a, μ, σ), then normalized and visualized in the form of radar plots to facilitate intuitive comparison across both quality dimensions and inspection variables. Figure 8 presents the visualization results of these SHAP statistics. The radar chart axes correspond to five inspection control factors: Speed (A), Stability (B), Scan Density (C), Tunnel Curvature (D), and Slope (E). Since categorical variables are included in the model, their SHAP contributions are displayed by level (e.g., B1: "flat track", B2: "track between rails"), reflecting the actual marginal effect of a specific inspection condition. It is important to note that for categorical variables, only one level is active per sample, and thus SHAP values for different levels are independent and not cumulative.

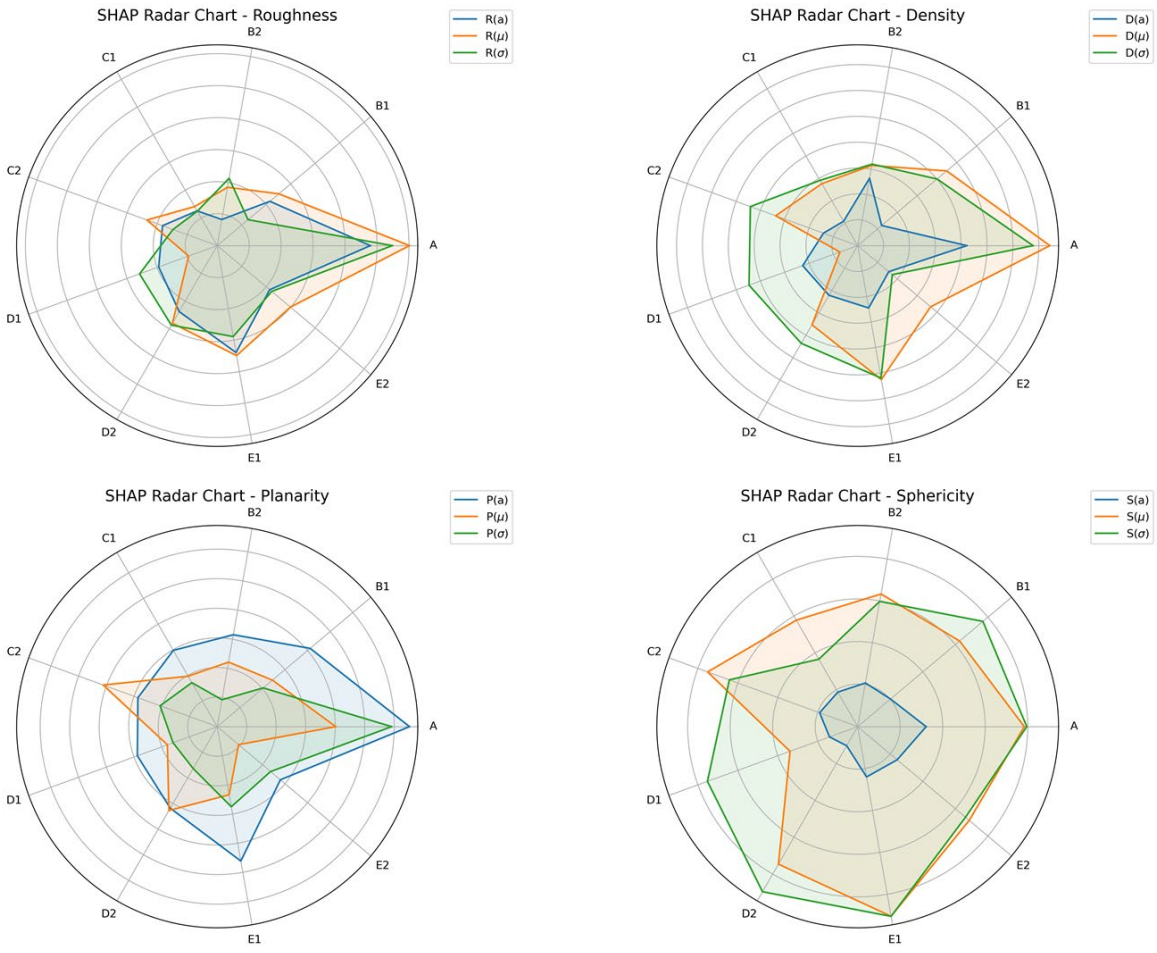

Figure 8. Radar plots of variable influence on each quality dimension.

The radar plots reveal substantial variation in dominant influencing factors across quality dimensions. Density is predominantly influenced by speed (A) and slope (E1/E2). The location parameter and scale parameter both exhibit strong sensitivity, suggesting that variations in elevation gradients and mobility significantly affect the spatial distribution of point. Roughness exhibits high dependency on speed (A), especially in the shape and scale parameters, suggesting that fast movement tends to induce trajectory instability, thereby affecting surface geometric continuity. The influence of other parameters is comparatively moderate. Planarity is sensitive to multiple input variables, scan density (C2), curvature (D2), and slope (E1) show prominent SHAP contributions across all distribution parameters, with the shape parameter and location parameter being most responsive. These patterns suggest that surface planarity is influenced by both sampling resolution and trajectory geometry. Sphericity shows broad sensitivity across multiple factors. While speed (A) and stability (B1/B2) is a key influence, curvature (D2) and slope (E1) also significantly contribute to both the location and scale parameters. This implies that inconsistencies in device posture, trajectory bending, and elevation jointly affect the geometric uniformity of point cloud data.

4.2. Directional effects of experimental factors on quality parameters

A directional interpretation of how each inspection variable affects the predicted quality parameters was conducted. To this end, SHAP summary plots were generated for each output parameter, in which variable names (including interaction terms) are shown on the vertical axis, and their corresponding marginal SHAP contributions are plotted on the horizontal axis. Color encoding indicates the feature value across different samples. This plot format simultaneously conveys feature importance, influence direction, and value dependence, offering a comprehensive perspective for uncovering the predictive mechanisms embedded in the model. To enable analysis of not only individual factor effects but also potential interactions between inspection conditions, the experiment adopted a full factorial design across five factors. This ensured that the influence of pairwise combinations—such as speed × slope or scan density × curvature—could be quantitatively assessed through interpretable machine learning.

Figure 9 to Figure 12 are SHAP dependence plots of top features and feature interactions for three density distribution parameters. Each point represents a sample's SHAP value for the given feature. The x-axis indicates the contribution magnitude and direction to the model output; the color encodes the actual feature value (blue = low, red = high). Feature combinations (e.g., A × C1) reflect interaction effects. The plots reveal how both individual factors and their pairwise interactions affect predicted density characteristics.

As shown in Figure 9, the SHAP dependence plots provide actionable insights into how individual variables and their interactions affect the three predicted density distribution parameters. For shape parameter, interaction terms like A × B1 and A × C1 show clear opposite effects: skewness decreases under low speed and low stability, but increases significantly when both are high. This implies that unstable motion combined with fast scanning tends to distort point distribution symmetry. In location parameter, speed is the dominant factor—high-speed scenarios consistently reduce the predicted mean density. Interaction terms such as A × B1 and C1 × E1 further reinforce this reduction, suggesting that fast motion and slope or dense sampling together suppress average point accumulation. For scale parameter, speed again reduces the scale, but some interactions (e.g., A × C2) show the opposite effect under specific settings, where fast movement combined with dense scanning can lead to wider point

dispersion. Overall, the analysis confirms that high-speed operation generally reduces both the average and spread of point cloud density, while certain variable combinations can either suppress or exaggerate local variation.

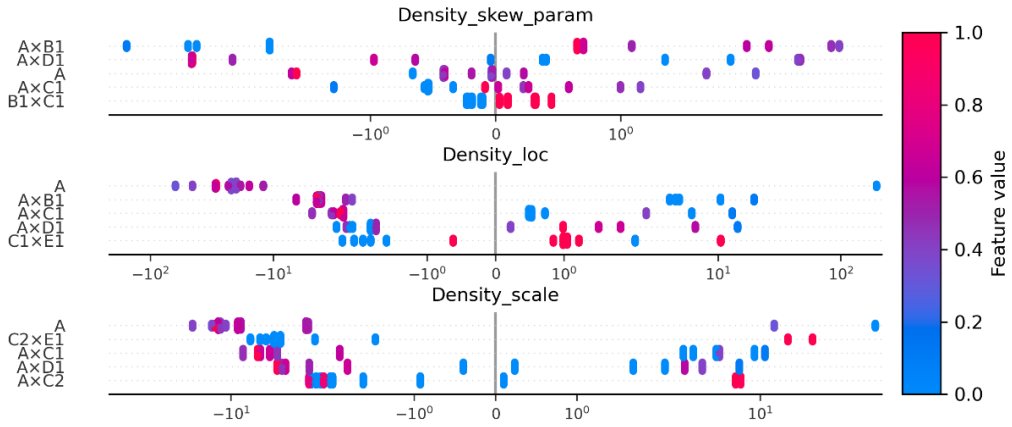

Figure 9. SHAP summary plots grouped by density dimensions.

Figure 10 illustrates the directional SHAP effects for the roughness metric. For the shape parameter, the pattern is more complex. While many feature combinations (e.g., A × C1, A × E1, C1 × E1) show that low-speed and low-slope samples reduce skewness, certain cases like A × D2 indicate that skewness can increase under specific geometric configurations, such as scanning fast in curved sections. This result highlights that while speed reduces overall roughness, it may also introduce asymmetric distortions depending on terrain shape and scanning conditions. For the location and scale parameters, high-speed samples are clearly associated with lower SHAP values, indicating that faster motion tends to reduce both the average surface roughness and its variability. This effect is further strengthened when combined with other factors, all of which show similar suppressive patterns under high-speed conditions. These results suggest that fast scanning generally produces smoother and more uniform surfaces, but may also suppress local surface detail.

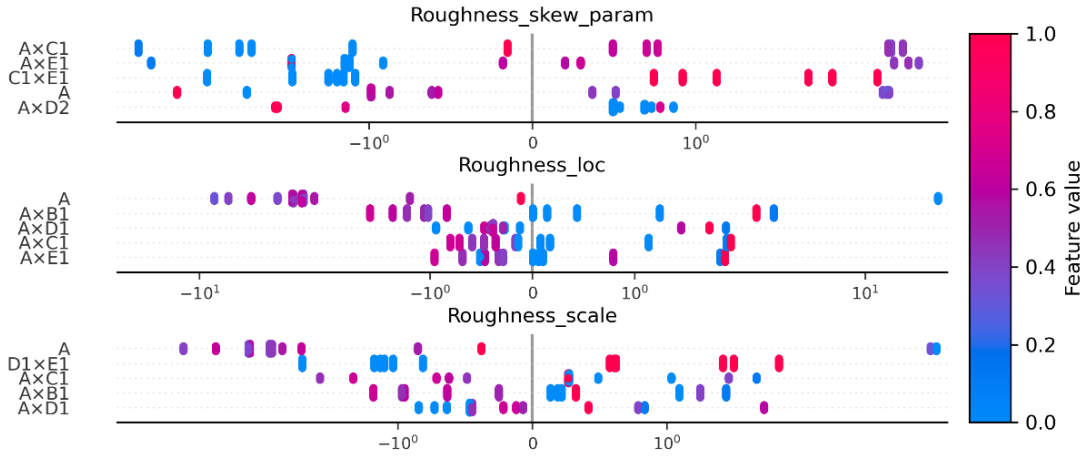

Figure 10. SHAP summary plots grouped by roughness dimensions.

Figure 11 presents the directional SHAP effect plot for the planarity metric. For the shape parameter, interaction terms such as A × D1, A × C1, and A × D2 show a clear trend where higher speed is associated with lower predicted skewness values. This result indicates that faster speed tends to reduce directional bias in the planarity distribution, making it more symmetric and balanced across the scan area. For the location parameter, combinations like A × C1, C1 × D1, and A × C2 show that higher scan density and speed lead to a lower predicted average planarity value, implying that aggressive scanning in complex geometry can reduce overall flatness. For the scale parameter, speed and its combinations with factors such as stability, curvature, and slope consistently contribute to lower predicted variation in planarity. This result suggests that high-speed operation suppresses localized deviations, resulting in a narrower range of predicted values. Overall, the results indicate that while faster scanning may produce more symmetric and concentrated planarity predictions, it also tends to reduce the average flatness, especially under unstable or inclined conditions.

Figure 12 presents the directional SHAP response for the sphericity metric. For the shape parameter, SHAP values are concentrated around zero across all features and input ranges, indicating that skewness is insensitive to variable changes. In contrast, the location parameter shows significant negative SHAP contributions across both low and high feature values. This result suggests that under conditions of low stability, steep slopes, or high curvature, the predicted mean sphericity tends to decrease, implying that the point cloud structure may become less regular or spatially coherent. The scale parameter exhibits a similar trend: variables such as C1 × D2 and B2 × E1 produce consistently strong negative SHAP values, indicating that under complex or sloped terrain, the predicted variation in sphericity is compressed. These findings suggest that motion posture and tunnel structure features primarily affect the structural regularity and compactness of the point cloud, while having limited impact on the asymmetry of its distribution.

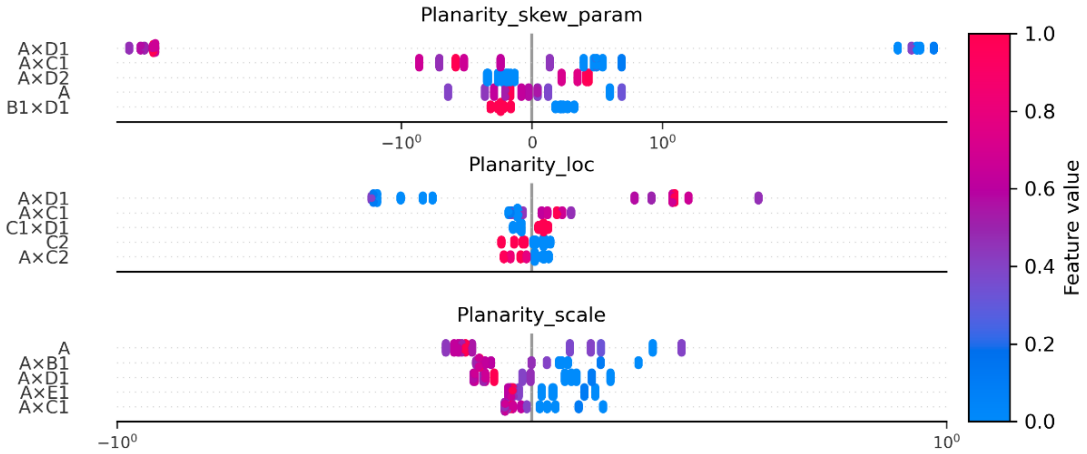

Figure 11. SHAP summary plots grouped by planarity dimensions.

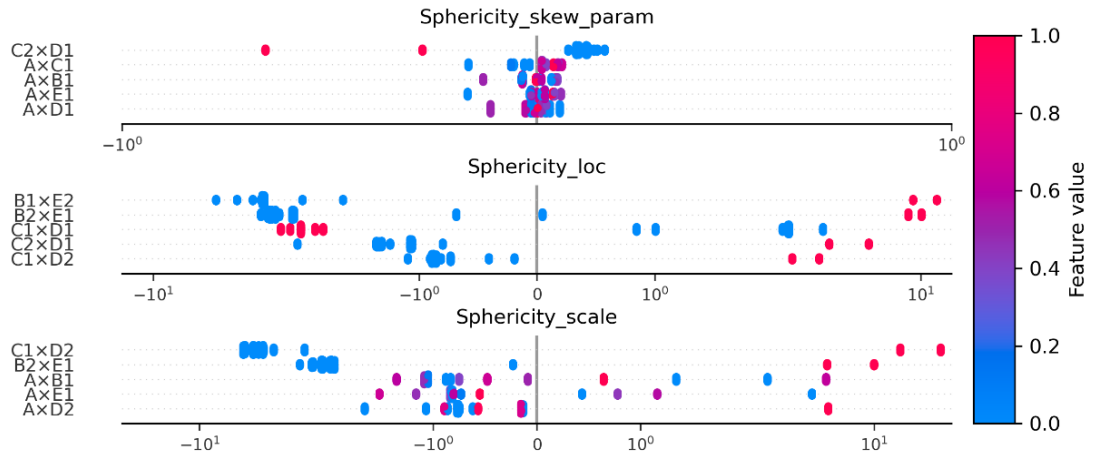

Figure 12. SHAP summary plots grouped by sphericity dimensions.

4.3. Multi-parameter comparison of experimental factor contributions

To systematically evaluate the overall influence of experimental variables on various point cloud quality parameters, a SHAP-based heatmap was constructed using the average SHAP values across all samples (Figure 13). This visualization highlights the relative importance of each input variable—including both main effects and interaction terms—on the three Skew-Normal distribution parameters of the four geometric quality metrics. The x-axis represents input variables (including interactions), and the y-axis lists the target distribution parameters. Color indicates the average magnitude of SHAP values (yellow = high importance, purple = low). This visualization highlights which inspection conditions most strongly influence each aspect of point cloud quality.

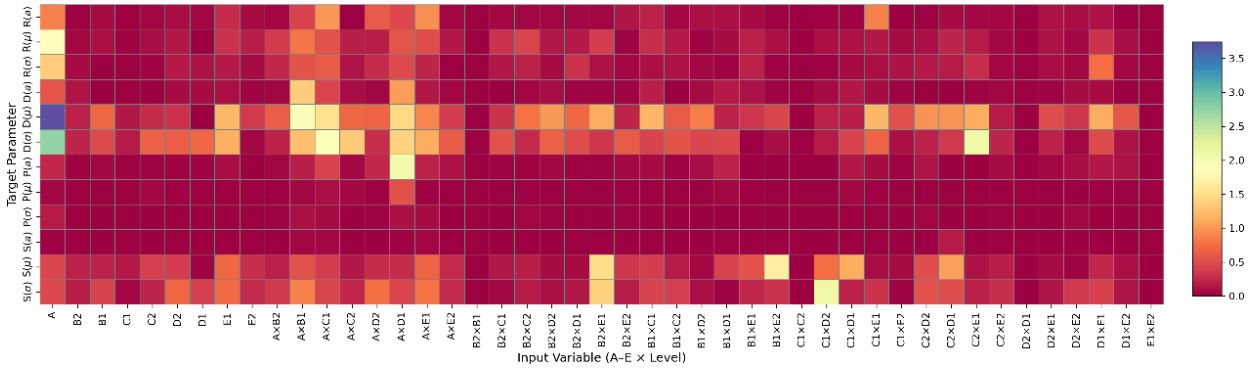

Figure 13. Heatmap of global SHAP importance across parameters.

The heatmap reveals moving speed consistently exhibits high SHAP values across almost all quality metrics and their distribution parameters. Not only does the main effect of speed dominate roughness and density, but several of its interaction terms also show strong impacts across multiple targets. These patterns suggest that inspection speed is a core driver of both density-related metrics and structural regularity. From an engineering standpoint, precise control of walking speed—particularly under specific scanning configurations and tunnel geometries—is essential to ensure stable point cloud quality

during real-world deployment. To maintain data acquisition stability, operators are advised to use speed monitoring tools or configure adaptive scanning modes.

In contrast to globally dominant factors, several input variables show strong selectivity toward specific quality metrics. For example, the interaction term B2*E1 is particularly prominent in influencing location and scale parameters of sphericity, while having negligible impact on other parameters. Similarly, C2*E1 plays a strong role in improving location and scale parameters of density, but contributes little to roughness or sphericity. These findings indicate that certain scanning conditions only affect localized aspects of point cloud structure, often in combination with geometric or operational constraints. Practically, these results support the tuning of data collection configurations for specific inspection objectives: if a task primarily focuses on surface planarity (e.g., tunnel lining inspection), optimizing factors like scan density and walking path under slope variations becomes critical. Therefore, recognizing these dimension-specific variables can guide cost-effective inspection planning, avoiding unnecessary redundancy in system settings.

The SHAP heatmap also reveals distinct response structures among the four quality dimensions. For instance, roughness-related metrics are primarily sensitive to motion-related variables like A*D1, highlighting their dependence on trajectory stability and SLAM pose consistency. In contrast, density metrics are strongly influenced by scan coverage and environmental openness, as reflected by high SHAP values from C1, C2, and E1. Notably, planarity and sphericity exhibit a broader set of contributing factors, including interaction terms involving curvature, slope, and movement stability, suggesting that these structural descriptors are shaped by more complex inspection conditions coupling. This divergence implies that different quality goals require differentiated acquisition strategies. For example, maximizing point density in narrow tunnels may call for repeated scanning, whereas achieving uniform sphericity under slope conditions may require optimized walking posture and path selection. Recognizing these sensitivities enables scenario-specific configuration planning and helps maintain robust point cloud quality across diverse tunnel conditions.

It should be noted that although the experimental tunnel included a range of slopes, curvatures, and scanning conditions, but all data were collected within a single metro tunnel system. The limited dataset may lead to the full generalizability of the findings. Future work should extend the framework to diverse tunnel types, geological settings, and even above-ground infrastructure environments, to evaluate its adaptability across broader application domains.

5. Conclusion

This study proposes an interpretable modeling and evaluation framework to address the issue of point cloud quality in metro tunnel inspection using a backpack SLAM LiDAR system. By designing orthogonal field experiments, extracting quality metrics, and introducing Skew-Normal distribution modeling in conjunction with interpretable machine learning, the multidimensional influence mechanisms of five inspection control factors on point cloud quality were systematically analyzed and validated through a real tunnel case study. Experimental results reveal that speed is the most dominant factor affecting point cloud quality, not only through its direct impact but also through interactions with scan density and moving stability. Roughness and density are mainly influenced by motion-related variables, while planarity and sphericity are more affected by tunnel geometry and operator posture. In particular, some variables that appear weak on average show strong effects on specific quality metrics,

such as the interaction between flat track and curved tunnel show strong effects on density. This phenomenon highlights that different inspection goals require different acquisition strategies rather than uniform inspection parameter settings.

Despite the strong interpretability and applicability of the proposed framework, certain limitations remain. First, the number of experimental samples was constrained by practical engineering conditions, leading to limited coverage of the variable space. Future work may incorporate synthetic data to extend variable gradients. Second, the current explanation framework is built on static regression structures, without modeling temporal dynamics or error propagation in the SLAM process. Moreover, the quality assessment focused primarily on geometric features, without considering multi-source data such as color or semantic information. Future research could enhance the model's dimensional adaptability and scenario generalization, and further embed this framework into practical tunnel point cloud mapping or inspection workflows.

Acknowledgements

This work is supported in part by "the National Natural Science Foundation of China" (Grant nos. 52192664). The authors thank Mr. Wan Yong, Mr. You Zhengjun and Mr. Fan Bin from Wuhan Metro Group Co., Ltd., for outstanding support in the data collection and analysis.

Authors' contribution

Conceptualization, Cheng Zhou and Wenbo Qin; methodology, Cheng Zhou and Wenbo Qin; software, Wenbo Qin; validation, Wenbo Qin, Shangbin Gao and Cheng Zhou.; formal analysis, Wenbo Qin; investigation, Wenbo Qin and Shangbin Gao; resources, Cheng Zhou; data curation, Wenbo Qin and Shangbin Gao; writing—original draft preparation, Wenbo Qin; writing—review and editing, Cheng Zhou; visualization, Wenbo Qin; supervision, Cheng Zhou; project administration, Cheng Zhou; funding acquisition, Cheng Zhou. All authors have read and agreed to the published version of the manuscript.

Conflicts of interests

The authors declare no conflict of interest.

References

[1] Liu W, Shao Y, Li C, Li C, Jiang Z. Development of a non-Gaussian copula Bayesian network for safety assessment of metro tunnel maintenance. *Reliab. Eng. Syst. Saf.* 2023, 238:109423.

[2] Qin W, Chen EJ, Wang F, Liu W, Zhou C. Data-driven models in reliability analysis for tunnel structure: a systematic review. *Tunnelling Underground Space Technol.* 2024, 152:105928.

[3] Zhou C, Gao Y, Chen EJ, Ding L, Qin W. Deep learning technologies for shield tunneling: Challenges and opportunities. *Autom. Constr.* 2023, 154:104982.

[4] Fu Y, Liang N, Guo Y, Chen X, Wu Z, *et al.* Calculation method and application of discontinuous longitudinal deformation of shield tunnels considering variable stiffness of the ring joints. *Tunnelling Underground Space Technol.* 2025, 161:106592.

[5] Gu Y, Ai Q, Xu Z, Yao L, Wang H, *et al.* Cost-effective image recognition of water leakage in metro tunnels using self-supervised learning. *Autom. Constr.* 2024, 167:105678.
[6] Yang K, Bao Y, Li J, Fan T, Tang C. Deep learning-based YOLO for crack segmentation and measurement in metro tunnels. *Autom. Constr.* 2024, 168:105818.
[7] Zhou C, Qin W, Luo H, Yu Q, Fan B, *et al.* Digital twin for smart metro service platform: evaluating long-term tunnel structural performance. *Autom. Constr.* 2024, 167:105713.
[8] Yue H, Wang Q, Zhao H, Zeng N, Tan Y. Deep learning applications for point clouds in the construction industry. *Autom. Constr.* 2024, 168:105769.
[9] Lin W, Sheil B, Zhang P, Zhou B, Wang C, *et al.* Seg2Tunnel: a hierarchical point cloud dataset and benchmarks for segmentation of segmental tunnel linings. *Tunnelling Underground Space Technol.* 2024, 147:105735.
[10] Li Y, Xiao Z, Li J, Shen T. Integrating vision and laser point cloud data for shield tunnel digital twin modeling. *Autom. Constr.* 2024, 157:105180.
[11] Ji A, Zhang L, Fan H, Xue X, Dou Y. Dual attention-based deep learning network for multi-class object semantic segmentation of tunnel point clouds. *Autom. Constr.* 2023, 156:105131.
[12] Li P, Wang Q, Li J, Pei Y, He P. Automated extraction of tunnel leakage location and area from 3D laser scanning point clouds. *Opt. Lasers Eng.* 2024, 178:108217.
[13] Han J, Guo J, Jiang Y. Monitoring tunnel deformations by means of multi-epoch dispersed 3D LiDAR point clouds: an improved approach. *Tunnelling Underground Space Technol.* 2013, 38:385–389.
[14] Chen Y, Shen S, Zhou A. Extraction of assembly quality information in shield tunnels from LiDAR data. *Measurement* 2025, 245:116541.
[15] Montero R, Victores JG, Martínez S, Jardón A, Balaguer C. Past, present and future of robotic tunnel inspection. *Autom. Constr.* 2015, 59:99–112.
[16] Jiang Y, Wang L, Zhang B, Dai X, Ye J, *et al.* Tunnel lining detection and retrofitting. *Autom. Constr.* 2023, 152:104881.
[17] Menendez E, Victores JG, Montero R, Martínez S, Balaguer C. Tunnel structural inspection and assessment using an autonomous robotic system. *Autom. Constr.* 2018, 87:117–126.
[18] Yasuda T, Yamamoto H, Enomoto M, Nitta Y. Smart tunnel inspection and assessment using mobile inspection vehicle, non-contact radar and AI. In *Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC 2020)*, Kitakyushu, Japan, October 27–29, 2020, pp. 1373–1379.
[19] Cui H, Ren X, Mao Q, Hu Q, Wang W. Shield subway tunnel deformation detection based on mobile laser scanning. *Autom. Constr.* 2019, 106:102889.
[20] Kim H, Choi Y. 3D location estimation and tunnel mapping of autonomous driving robots through 3D point cloud registration on underground mine rampways. *Underground Space* 2025, 22:1–20.
[21] Jian L, Qiu W, Cheng Y. Centerline-based registration for shield tunnel 3D reconstruction using spinning mid-range LiDAR point cloud and multi-cameras. *Autom. Constr.* 2025, 171:105950.
[22] Bao S, Shi W, Chen P, Xiang H, Yu Y. A systematic mapping framework for backpack mobile mapping system in common monotonous environments. *Measurement* 2022, 197:111243.
[23] Chow JK, Liu K, Tan PS, Su Z, Wu J, *et al.* Automated defect inspection of concrete structures. *Autom. Constr.* 2021, 132:103959.

[24] Shi W, Chen P, Wang M, Bao S, Xiang H, *et al.* PolyU-BPCoMa: a dataset and benchmark towards mobile colorized mapping using a backpack multisensorial system. *Int. J. Appl. Earth Obs. Geoinf.* 2022, 112:102962.

[25] Wang Y, Song W, Zhang Y, Huang F, Tu Z, *et al.* Four years of multimodal odometry and mapping on the rail vehicles. *J. Field Rob.* 2024, 41:227–257.

[26] Ivina D, Palmqvist CW. Railway maintenance windows: discrepancies between planning and practice in Sweden. *Transp. Res. Interdiscip. Perspect.* 2023, 22:100927.

[27] Eugster H, Huber F, Nebiker S, Gisi A. Integrated georeferencing of stereo image sequences captured with a stereovision mobile mapping system—approaches and practical results. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 2012, 39:309–314.

[28] Zhao S, Fang Z, Li H, Scherer S. A robust laser-inertial odometry and mapping method for large-scale highway environments. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China, November 3–8, 2019, pp. 1285–1292.

[29] Wajs J, Kasza D, Zagożdżon PP, Zagożdżon KD. 3D modeling of underground objects with the use of SLAM technology on the example of historical mine in Ciechanowice (Ołowiane Range, The Sudetes). 2018. Available: https://doi.org/10.1051/e3sconf/20182900024 (accessed on 3 April 2025).

[30] Sammartano G, Spanò A. Point clouds by SLAM-based mobile mapping systems: accuracy and geometric content validation in multisensor survey and stand-alone acquisition. *Appl. Geomatics* 2018, 10:317–339.

[31] Štroner M, Urban R, Křemen T, Braun J, Michal O, *et al.* Scanning the underground: comparison of the accuracies of SLAM and static laser scanners in a mine tunnel. *Measurement* 2025, 242:115875.

[32] Letard M, Lague D, Le Guennec A, Lefèvre S, Feldmann B, *et al.* 3DMASC: accessible, explainable 3D point clouds classification. Application to bi-spectral topo-bathymetric lidar data, *ISPRS J. Photogramm. Remote Sens.* 2024, 207:175–197.

[33] Daghigh H, Tannant DD, Daghigh V, Lichti DD, Lindenbergh R. A critical review of discontinuity plane extraction from 3D point cloud data of rock mass surfaces. *Comput. Geosci.* 2022, 169:105241.

[34] Günen MA, Öztürk KF, Aliyazıcıoğlu Ş. Supervised classification-based framework for rock mass discontinuity identification using point cloud data. *Eng. Geol.* 2025, 350:107987.

[35] Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review, *J. Big Data* 2020, 7:94.

[36] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, USA, December 4–9, 2017, pp. 4768–4777.