Deep Learning Framework and Performance Evaluation for Point-Cloud Time-Series Change Detection in Long-Term Monitoring of Ancient Bridges

Jian Li[1#], Xiyin Ma[2#]
[1]College of Urban Construction and Ecological Technology, Shanghai Institute of Technology, Shanghai 201418, China
[2]School of Design and Art, Shanghai Technical Institute of Electronics & Information, Shanghai 201411, China

Corresponding author: Xiyin Ma (e-mail: maxiyin1010@126.com).
#These authors contributed equally to this work and should be considered co-first authors.

ABSTRACT
Ancient bridges exhibit progressive deformation and localized spalling due to long-term material weathering, humidity–temperature cycles, foundation settlement, and traffic loading, making accurate structural health monitoring essential for heritage preservation. In advanced electromagnetic and intelligent sensing systems, high-precision geometric monitoring complements radar- and wave-based inspection technologies for comprehensive infrastructure assessment. Conventional geometric-differencing methods for point-cloud change detection are susceptible to accumulated registration errors, noise occlusion, and non-uniform point density during long-term monitoring, limiting reliable identification of millimeter-level deformations. This study establishes a deep learning framework for multi-temporal point-cloud change modeling to satisfy the requirements of life-cycle structural monitoring. The temporal characteristics of structural changes and the propagation mechanisms of registration errors are analyzed, and mathematical formulations for cloud-to-cloud (C2C) and cloud-to-plane (C2P) change metrics are presented. The proposed framework provides a theoretical foundation for deep learning-based change detection and performance evaluation, while offering technical support for multimodal structural sensing and intelligent monitoring applications in engineering infrastructures.

INDEX TERMS
long-term monitoring of ancient bridges, time-series point clouds, change detection, deep learning, electromagnetic sensing

I. INTRODUCTION
STRUCTURES from the past provide value relating to history and use. These structures undergo service across time and show effects from changes in conditions, movement in the basis, loads from use, and changes in material [1]. This process produces changes including movement in the basis, changes in form, breaking of material, and development in separation. Changes in the initial stages occur at levels of small measure, and this establishes a requirement for methods that allow measurement across time and provide assessment of possible issues [2,3]. High-accuracy data from three-dimensional scanning reveals the intricate geometric form across woven structures, offering an essential foundation for examining the technical construction and material degradation of ancient artifacts. Data from multiple points in time show effects from variation, limitations in measurement, differences in the process of measurement, and differences that develop from combining data. In methods that examine differences in form, changes in the structure and differences from variation show relationships that appear in the data. Changes include multiple types relating to measure and form, and this establishes limitations in approaches that use set values [4,5]. The approaches show limitations in finding changes at small levels and in providing control of incorrect findings. Developments in recent work present models that use data showing form and include approaches that examine pairs and examine changes across time. These models provide approaches for identifying structural variations in fibers and yarn density under changing environmental conditions. Monitoring the longitudinal patterns of these fabric deformations facilitates more precise measurements of material performance and

wear across time.

II. THEORY AND PRINCIPLES OF POINT-CLOUD TIME-SERIES CHANGE DETECTION FOR ANCIENT BRIDGES

A. TYPES OF STRUCTURAL CHANGES AND TEMPORAL CHARACTERISTICS

Ancient bridges that remain in service over time show different types of change occurring across multiple measures of scale. These changes occur through the process of material breaking down over time, and loads and conditions in the surrounding context also produce these effects [6]. In general, changes that occur can be separated into two main categories: overall changes in form and particular defects that appear in specific areas. Overall changes in form include situations such as supports and foundations that settle in ways that differ across locations, changes in the degree of curve in structural elements, and movement or tilting that affects the structure as a whole [7-9]. These changes appear as processes that develop in ways that are continuous and that occur at rates that are relatively slow, and the changes show patterns that build up over time and that follow trends that are clear. Particular defects that appear in specific areas include surfaces of stone material that break down through exposure and that lose portions of material, defects that increase in extent, separation that develops at points where elements meet, and cracks that develop and that increase. These changes often appear as breaks in the continuity of form that are limited to particular areas and that differ across locations, and the process through which these develop may show periods in which change occurs at rates that increase or events that occur in ways that are sudden. For work that involves tracking structures over extended periods, changes in structures also show interference that comes from multiple sources and that is significant. For example, plants that appear in certain seasons and that block observation, differences in moisture on surfaces that affect the properties relating to how light reflects, and blocking that occurs for limited periods and variation that changes over time that results from individuals walking and from vehicles may all produce differences in form in data collected as points across time that appear similar to changes in structures that are actual. The process of identifying change across time for ancient bridges must identify areas where breakdown that is genuine occurs, and this process must also separate changes that occur in ways that are gradual from changes that occur in ways that are sudden, and areas that remain stable from areas that do not remain stable, using patterns that appear across sequences that extend over time [10-11]. This approach provides results from the process of identification that show greater reliability and that offer interpretation that is more useful for work in the field.

B. POINT-CLOUD TIME-SERIES DATA MODEL AND PROBLEM DEFINITION

Let the point cloud acquired from the t-th monitoring campaign of an ancient bridge be denoted as

$$P_t = \{x_i^t \in \mathbb{R}^3 \mid i = 1, 2, \ldots, N_t\},$$

where $x_i^t = (x_i^t, y_i^t, z_i^t)$ represents the three-dimensional coordinates and N_t is the number of points. Long-term monitoring produces a point-cloud sequence $\{P_t\}_{t=1}^{T}$, whose data characteristics typically include inconsistent sampling density, significant occlusions and missing regions, and temporally varying noise distributions [12-14]. Point-cloud time-series change detection can be formalized as a problem of “inter-temporal difference identification and quantification.” After performing inter-epoch registration, the model is required to make point-wise predictions, including the change probability $\hat{y}_i^t \in [0, 1]$ (change segmentation) and the change magnitude $\hat{\delta}_i^t \in \mathbb{R}$ (change regression). The change magnitude may correspond to displacement amplitude, defect depth, or the projected differences along the normal direction [15]. For long-term monitoring, change detection should further incorporate temporal consistency constraints, i.e., non-changing regions are expected to maintain stable responses across multiple epochs, whereas true change regions should exhibit continuous or stage-wise trends. Therefore, this task is not merely a static differencing problem between two point clouds, but rather a dynamic inference problem that integrates multi-temporal information. The model must be capable of handling the complex coupling among registration bias, noise occlusion, and structural changes, thereby enabling reliable identification and quantitative assessment of subtle changes.

C. POINT-CLOUD REGISTRATION AND ANALYSIS OF ERROR SOURCES

Establishing spatial relationships between point clouds from different time periods requires aligning data within a single coordinate system. This process typically follows two stages. The initial stage uses feature matching or control points to establish starting positions. The following stage applies methods that reduce distances between point sets through repeated adjustments. In long-term monitoring of ancient bridges, errors from this alignment process show significant accumulation over time. Multiple factors contribute to these errors. Instrument limitations and measurement variations introduce random shifts in recorded positions, and these shifts may change as conditions during data collection differ. Differences in scanner placement and areas where data cannot be collected reduce the shared regions between datasets, which decreases the reliability of the adjustment process and produces localized misalignments. Changes in environmental factors and surface characteristics—including water presence, biological growth, illumination differences, and reflection properties—affect data quality and increase uncertainty in alignment. When actual structural changes occur, alignment methods may incorrectly interpret changed regions as alignment errors and force these areas to match the reference model, which weakens signals indicating change or creates false indications of deformation. Consequently, detecting changes over long periods must explicitly address how alignment errors interact with genuine structural changes in geometric comparison. Stability requirements and robustness considerations should be integrated into both model development and assessment approaches to limit how error accumulation affects conclusions from long-term monitoring. It should be emphasized that the proposed deep learning framework does not assume perfectly accurate inter-epoch registration. Instead, registration is used only to provide coarse spatial alignment that enables approximate correspondence between epochs. The subsequent feature-differencing module and temporal fusion mechanism are designed to learn robust representations that are less sensitive to small residual registration errors. By modeling temporal consistency across multiple epochs, the network can suppress pseudo-changes caused by minor misalignments while reinforcing persistent structural changes, thereby mitigating the impact of registration noise on change detection.

D. PRINCIPLES AND FORMULATIONS OF POINT-CLOUD CHANGE METRICS

After registration, change metrics are used to quantify the geometric discrepancies between point clouds acquired at different epochs, forming the theoretical basis for change detection and change magnitude regression. Commonly used measures mainly include cloud-to-cloud (C2C) distance and cloud-to-plane (C2P) distance. For a reference point $x_i^a \in P_a$ and a target point cloud P_b, the C2C metric is defined as the nearest-neighbor distance:

$$d_{\mathrm{C2C}}(x_i^a, P_b) = \min_{y \in P_b} \|x_i^a - y\| \tag{1}$$

This method is simple to implement; however, under conditions of non-uniform point density and noise, it is easily affected by outliers. To improve sensitivity to continuous deformations on structural surfaces, C2P projects the discrepancy onto the normal direction of a local plane. Let y_i denote the nearest neighbor of x_i^a in P_b, and let n_i be the normal vector of the fitted local plane within the neighborhood of y_i. Then,

$$d_{\mathrm{C2P}}(x_i^a, P_b) = (x_i^a - y_i) \cdot n_i \tag{2}$$

C2P better conforms to the engineering interpretation of "true deformation along the structural surface normal," and is thus suitable for detecting micro-deformations such as settlement and deflection. In long-term monitoring, one may further define the temporal change $\Delta_t(x)$ and the cumulative change

$$C_T(x) = \sum_{t=2}^{T} \Delta_t(x) \tag{3}$$

to characterize change trends and accumulation effects, thereby providing a quantitative basis for constructing supervision signals and evaluating the performance of subsequent deep learning models.

III. DEEP LEARNING FRAMEWORK FOR POINT-CLOUD TIME-SERIES CHANGE DETECTION IN LONG-TERM MONITORING

A. OVERALL FRAMEWORK DESIGN

1) Formalization of Inputs and Outputs

Let the point-cloud sequence acquired from long-term monitoring be:

$$P = \{P_t\}_{t=1}^{T}, \quad P_t = \{x\}_i^t \in \mathbb{R}^3\}_{i=1}^{N_t} \tag{4}$$

After inter-epoch registration, the point clouds from different epochs are aligned within a unified coordinate system. The objective of the proposed framework is to predict, for each point, both its change status and the magnitude of change. The outputs include:

Change segmentation probability:

$$\hat{y}_i^t \in [0, 1] \tag{5}$$

Change magnitude regression:

$$\hat{\delta}_i^t \in \mathbb{R} \tag{6}$$

The change magnitude can be defined as the displacement projection of a point along the normal direction, the settlement amplitude, the defect depth, or an integrated change quantity (as specified by the supervised labels). To accommodate the evaluation requirements of long-term monitoring, this study formulates change detection as a point-wise multi-task learning problem:

$$F : \{P_{t-k}, \ldots, P_t\} \to \{\hat{y}_t, \hat{\delta}_t\} \tag{7}$$

2) Overall Pipeline and Module Composition

Bridge-TSCDNet consists of four core modules: Point-cloud deep feature encoder: extracts local geometric and semantic features from the point cloud of each epoch; Inter-epoch differencing and alignment module: constructs feature differences and correspondence relationships between adjacent epochs, alleviating inconsistencies caused by sampling variations and occlusions; Temporal feature fusion module: models temporal structures to capture change trends and suppress false alarms induced by random noise; Change prediction head: outputs change segmentation and change magnitude regression results, and provides interpretable quantitative representations of changes. The proposed framework is applicable to both bi-temporal change detection and can be naturally extended to multi-temporal sequences, enabling a unified formulation of "localized abrupt change detection + long-term gradual trend modeling."

B. DESIGN OF THE POINT-CLOUD DEEP FEATURE ENCODING MODULE

Point clouds of ancient bridges are characterized by non-structured layouts, large density variations, and significant noise and missing regions, making conventional convolution difficult to apply directly. Therefore, this study employs a deep point-cloud network as the encoder $\Phi(\cdot)$, which maps a point set into point-wise feature representations:

$$F_t = \Phi(P_t), F_t = \{f_i^t \cup R^d\}_{i=1}^{N_t} \tag{8}$$

where d denotes the feature dimension.

1) Local Neighborhood Modeling (Geometric Sensitivity)

For each point x_i^t, a K-nearest neighbor (KNN) set is constructed as:

$$\mathcal{N}_K(x_i^t) = \{x_j^t \mid x_j^t \in \mathrm{KNN}(x_i^t)\} \tag{9}$$

Local geometric features can be encoded using relative displacement representations, i.e.,

$$\Delta x_{ij}^t = x_j^t - x_i^t \tag{10}$$

and are jointly fed, together with the point features, into a local aggregation operator:

$$f_i^t = \mathrm{Agg}\left(\{\varphi(\Delta x_{ij}^t) \odot W f_j^t\}_{j \in \mathcal{N}_K}\right) \tag{11}$$

Where ϕ $(\cdot)$ denotes the positional encoding function (e.g., an MLP), $\odot$ represents element-wise multiplication, and Agg $(\cdot)$ can be instantiated as max/mean pooling or an attention-based aggregation. This design enhances the model's sensitivity to local geometric discrepancies such as arch-ring curvature variations, pier flatness changes, and defect boundaries.

2) Attention Mechanism (Adapting to Multi-Scale Changes)

To simultaneously attend to subtle cracks and larger spalling regions, this study introduces point-wise self-attention to enable multi-scale feature modeling:

$$f_i^t = \alpha_{ij}^t W_v f_j^t \tag{12}$$

where the attentionweights are given by:

$$\alpha_{ij}^t = \frac{\exp(q_i^t \cdot k_j^t)}{\sum_{m \in \mathcal{N}_K} \exp(q_i^t \cdot k_m^t)} \tag{13}$$

$$q_i^t = W_q f_i^t, \quad k_j^t = W_k f_j^t \tag{14}$$

This mechanism enables the model to automatically focus on change-sensitive regions without introducing manual scale thresholds, thereby providing high-quality representations for subsequent differencing and temporal fusion.

C. TEMPORAL FEATURE FUSION MODULE

A key challenge in long-term monitoring is to distinguish true structural changes from pseudo-differences caused by noise, occlusions, or registration errors. Therefore, this study designs a temporal fusion module T $(\cdot)$, which models temporal consistency of changes based on inter-epoch differencing features.

1) Inter-epoch Feature Association and Difference Construction

Given the point-cloud features of two adjacent epochs (t-1) and (t), a difference representation is constructed as:

$$d_i^t = \mathrm{Fuse}\left(f_i^t, \tilde{f}_i^{t-1}, \left|f_i^t - \tilde{f}_i^{t-1}\right|\right) \tag{15}$$

where $\tilde{f}_i^{t-1}$ denotes the aligned feature from the previous epoch, obtained via inter-epoch nearest-neighbor search and/or correspondence matching:

$$\tilde{f}_i^{t-1} = f_{\Pi(i)}^{t-1}, \quad \Pi(i) = \arg\min_j \|x_i^t - x_j^{t-1}\| \tag{16}$$

This step establishes approximate correspondences under inconsistent sampling conditions, allowing the differencing features to better reflect true geometric changes rather than sampling fluctuations.

2) Temporal Recurrent Fusion (GRU-based)

For multi-epoch sequences, the differencing features are fed into a Gated Recurrent Unit (GRU) to model long-term change trends:

$$h_i^t = \mathrm{GRU}(d_i^t, h_i^{t-1}) \tag{17}$$

where the GRU update is formulated as:

$$z_i^t = \sigma(W_z[d_i^t, h_i^{t-1}]) \tag{18}$$

$$r_i^t = \sigma(W_r[d_i^t, h_i^{t-1}]) \tag{19}$$

$$\tilde{h}_i^t = \tanh(W_h[d_i^t, r_i^t \odot h_i^{t-1}]) \tag{20}$$

$$h_i^t = (1 - z_i^t) \odot h_i^{t-1} + z_i^t \odot \tilde{h}_i^t \tag{21}$$

where σ $(\cdot)$ denotes the Sigmoid function. This structure can gradually accumulate responses when changes persist over time, while suppressing false alarms caused by occasional noise through its memory mechanism, thereby improving the long-term stability metric (TS).

3) Optional: Temporal Transformer (Stronger Long-Range Dependency)

Alternatively, a Transformer can be adopted to capture stronger long-range temporal dependencies. Specifically, the differencing features within a temporal window of length L can be stacked as:

$$D_i = [d_i^{t-L+1}, \cdots, d_i^t] \tag{22}$$

The self-attention fusion is formulated as:

$$H_i = \mathrm{softmax}\left(\frac{Q_i K_i^\top}{\sqrt{d}}\right) \tag{23}$$

thereby enhancing the capability of capturing stage-wise abrupt changes (e.g., sudden expansion of spalling). In practice, either GRU or Transformer can be selected in the paper depending on the experimental setting, and both alternatives are well aligned with the proposed evaluation metric system.

D. DESIGN OF THE CHANGE PREDICTION HEAD

1) Multi-task Outputs (Classification + Regression)

The point-wise hidden representation h_i^t after temporal fusion is fed into a prediction head $G(\cdot)$, which outputs the change segmentation probability and the change magnitude:

$$\hat{y}_i^t = \sigma(W_c h_i^t + b_c) \tag{24}$$

$$\hat{\delta}_i^t = W_r h_i^t + b_r \tag{25}$$

where $\hat{y}_i^t$ is used for evaluating segmentation performance metrics such as IoU and F1-score, and $\hat{\delta}_i^t$ corresponds to change-magnitude accuracy metrics such as MAE.

2) Physical-Consistency Constraint on Change Magnitude

Considering that micro-deformations of ancient bridges are more appropriately expressed along the surface normal direction, a normal-projection consistency term can be introduced (either as supervision or as a regularization constraint):

$$\delta_i^t = (x_i^t - x_{\pi(i)}^{t-1}) \cdot n_{\pi(i)}^{t-1} \tag{26}$$

where n denotes the normal vector of the local fitted plane. This definition is consistent with the C2P principle, endowing the regression target with a clear engineering meaning.

3) Loss Functions (Consistent with Long-Term Monitoring Objectives)

To address the class imbalance where changed points are far fewer than unchanged points, the classification loss adopts Focal Loss:

$$L_{\mathrm{cls}} = -\frac{1}{N}\sum_i \alpha_i (1-\hat{y}_i^t)^\gamma y_i^t \log(\hat{y}_i^t) \tag{27}$$

The regression loss for change magnitude can be formulated using MAE or Huber loss:

$$L_{\mathrm{reg}} = \frac{1}{N}\sum_i |\hat{\delta}_i^t - \delta_i^t| \tag{28}$$

A key requirement for long-term monitoring is to suppress random fluctuations in non-change regions. Therefore, this study introduces a temporal consistency constraint:

$$L_{\mathrm{temp}} = \frac{1}{N}\sum_i \sum_{t=2}^{T} |\hat{\delta}_i^t - \hat{\delta}_i^{t-1}| \tag{29}$$

Moreover, to prevent false alarms from accumulating over time in stable regions, an additional stable-region suppression term can be incorporated:

$$L_{\mathrm{stb}} = \frac{1}{N}\sum_i (1-y_i^t)\cdot \hat{y}_i^t \tag{30}$$

Finally, the overall loss is defined as:

$$L = \lambda_1 L_{\mathrm{cls}} + \lambda_2 L_{\mathrm{reg}} + \lambda_3 L_{\mathrm{temp}} + \lambda_4 L_{\mathrm{stb}} \tag{31}$$

IV. SIMULATION-BASED EXPERIMENTAL VALIDATION AND CONTROLLED DATASET CONSTRUCTION

A. OBJECTIVES AND OVERALL SIMULATION SCHEME

To validate the proposed Bridge-TSCDNet under controllable conditions in the context of long-term monitoring of ancient bridges, this study constructs a simulation-based experimental system for multi-temporal point clouds. The simulation objectives consist of three aspects. First, a baseline point cloud with typical geometric characteristics of ancient bridge structures is generated, and engineering-realistic change patterns are injected into multi-epoch sequences to simulate representative deterioration processes, including pier settlement, arch-ring deflection, surface spalling defects, and crack propagation. Second, acquisition disturbance mechanisms such as noise, occlusions, and sampling variations are incorporated to mimic common sources of errors encountered in long-term field measurements and to analyze their impacts on change detection. Third, the simulated point-cloud sequences provide point-wise change labels and ground-truth change magnitudes, enabling the model to perform both change segmentation and change magnitude regression, and to be systematically evaluated using metrics such as IoU, F1-score, MAE, Chamfer Distance, and long-term stability (TS). The overall workflow follows a closed-loop design of "baseline modeling → point-cloud sampling → disturbance injection → change injection → label generation → training/testing and evaluation," ensuring reproducibility of experiments and verifiability of conclusions. Although the simulated dataset allows controlled evaluation of the proposed framework under different noise and occlusion levels, it should be noted that the deterioration patterns are synthetically injected according to engineering-informed assumptions. Such simulation enables reproducible benchmarking and provides precise ground-truth labels for both change segmentation and magnitude regression. Nevertheless, real-world structural deterioration processes may exhibit more complex behaviors than the simulated scenarios. In practice, deformation patterns can be influenced by heterogeneous material properties, environmental interactions, and measurement uncertainties that cannot be perfectly reproduced through simulation. Therefore, the current study focuses on validating the methodological feasibility of the proposed Bridge-TSCDNet under controlled conditions. Future work will incorporate long-term monitoring datasets acquired from terrestrial laser scanning campaigns on real historical bridges, which will enable further evaluation of the model under authentic structural deformation processes.

B. CONSTRUCTION OF BASELINE ANCIENT-BRIDGE POINT CLOUD AND SAMPLING STRATEGY

The baseline bridge model used for simulation represents the typical geometric characteristics of historical masonry arch bridges. The structure includes key components such as arch rings, piers, deck surfaces, and parapets, which together provide representative elements for analyzing both global structural deformation and localized damage patterns. These components allow the simulation of common deterioration processes in ancient bridges, including pier settlement, arch deflection, surface spalling, and crack propagation. The initial bridge geometry is converted into a point-cloud representation through spatial sampling. To approximate the characteristics of terrestrial laser scanning (TLS) data, a non-uniform sampling strategy is adopted. Regions with relatively simple geometry are sampled with lower density, while structurally significant regions—such as arch crowns, pier edges, and deck boundaries—are sampled with higher density. This approach produces point distributions that resemble those observed in practical bridge monitoring datasets. To further reflect realistic acquisition conditions, the sampling process introduces variability in point density across different structural regions. Such variability reflects the influence of scanning distance, incident angle, and surface reflectivity, which commonly lead to heterogeneous point distributions in field measurements. As a result, the generated baseline dataset captures both the geometric complexity of historical bridge structures and the irregular sampling characteristics typical of real scanning environments. The resulting baseline point cloud serves as the reference geometry for constructing multi-temporal monitoring sequences. Based on this reference model, subsequent experimental stages introduce controlled disturbances and simulated structural changes to generate labeled time-series datasets for change segmentation and change magnitude estimation.

C. NOISE AND OCCLUSION MODELING

To represent errors from measurement and conditions in the scanning process, the study provides a model using perturbations that follow zero-mean patterns. These perturbations affect the three-dimensional positions, and they produce deviations in positions at the millimeter level. The study examines multiple levels of this type to examine stability in detection and the degree to which errors can be reduced under different conditions. Missing regions that occur from occlusion also appear in monitoring that continues over time for the bridges. Occlusions from vegetation, occlusions from pedestrians and vehicles that change over time, and limitations from the positions used for scanning can produce loss of data in particular regions. The study represents missing data using two approaches. The first approach removes points following a pattern that is not based on position. The second approach uses regions that represent occlusion at particular locations such as the deck, the edge region of the arch ring, or the side of piers that is opposite from wind. This approach removes points within regions that are continuous in space to represent the features that occur in occlusions. The study also considers that occlusions change across seasons in monitoring over time. Different levels of occlusion are used for different time points to produce samples that are more difficult to examine across time. This requires the model to use patterns that are consistent across time rather than differences between two time points in the data to provide detection that remains stable.

D. TEMPORAL CHANGE INJECTION MECHANISMS

To develop sequences of point clouds that show changes with clear meaning for structures, four main types of change are applied to the initial point cloud to produce multiple sequences that simulate monitoring over time. First, settlement of the pier is represented as displacement moving downward that increases in a linear manner, representing overall deformation that develops from uneven conditions in the foundation. Second, deflection in the arch ring is applied using a pattern that moves downward at the middle span, such that the arch shows deformation that continues over time where deflection becomes larger in later measures. Third, defects from spalling are created by removing points in specific areas and expanding the edges of these areas, representing the process of material deterioration and defect expansion on surfaces of the structure. Fourth, propagation of cracks is generated by creating narrow grooves on surfaces of components that become deeper or wider in later sequences, showing how cracks develop from initial formation to later expansion. To provide the data required for learning from examples with supervision, labels indicating changes are generated for each point at the same time: labels for segmentation of changes indicate which areas show change and which areas remain the same, and labels for magnitude of changes provide the actual values such as displacement or depth of defects for each point that shows change. The sequences across multiple times include changes that develop gradually following trends and changes that occur suddenly in specific locations, allowing the tests to confirm that the model can identify both types of changes that develop over time and changes that appear abruptly, and providing support for assessment of how well the module for combining information across time performs.

E. BASELINE METHODS AND PARAMETER SETTINGS

To examine performance of the approach using deep learning methods, the study compares results with methods that provide baseline findings and with models that use deep point-cloud analysis. The methods that provide baseline results include an approach using threshold measures with distance from cloud to plane and a method that examines change in point-cloud data across multiple scales. These methods allow the study to show advantages of the approach using deep learning when data include noise and occlusions and when density differs across the sample. This study proposes uses a model that provides encoding of deep point-cloud features to obtain characteristics for individual points, and this combines with a component that allows fusion across time to support inference of change across multiple observation periods. The study follows principles for training that support reproduction of findings. The approach applies normalization of scale that remains consistent for inputs of point-cloud data. The study uses a fixed value for neighborhood size K, and this allows modeling of local geometric features that remains consistent. The study sets length of the temporal window according to the number of observation periods, and this means the approach uses information from differences between adjacent periods and information from trends across multiple periods. The study applies optimization that combines multiple tasks for losses in classification and regression, and this includes constraints for consistency across time that function to reduce accumulation of false detections. The approach uses a strategy with fixed division for training and testing samples, and the study conducts repeated tests under different levels of noise and different proportions of occlusion to provide stability in results for evaluation that uses statistical measures. We analyze multiple measures of performance that include performance for segmentation of change, which uses IoU and F1-score, and that include accuracy for estimation of change magnitude, which uses MAE. The measures also include consistency in geometric features, which uses Chamfer Distance, and stability for monitoring over extended time, which uses TS. These measures allow comparison across three aspects that include accuracy in detection, capability for providing quantitative results, and reliability over extended periods. In this study, the baseline comparison primarily focuses on classical geometric change detection methods (C2P-threshold and M3C2), which are widely used in engineering deformation monitoring. These methods provide a representative benchmark for evaluating the advantages of learning-based approaches under noise and occlusion conditions. While modern deep-learning point-cloud segmentation networks such as KPConv and RandLA-Net have demonstrated strong performance in semantic segmentation tasks, the primary objective of this work is to investigate temporal change modeling rather than purely spatial segmentation. Therefore, the comparison emphasizes the improvement over traditional geometric approaches commonly used in structural monitoring workflows. Incorporating additional deep-learning baselines will be considered in future studies.

V. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION METRIC SYSTEM

We provide validation for Bridge-TSCDNet in monitoring over time for bridges showing use in the past. This involves examining the method for finding changes, for measuring change levels, and for maintaining patterns across time. The analysis uses data that include point clouds from multiple periods that the study creates in Chapter 4. The comparison includes methods that differ in approach. The first method uses differences in form with a level for deciding changes and is called C2P-threshold. The second method uses scales for finding changes and is called M3C2. The third method uses the approach that applies learning from data and is called Bridge-TSCDNet or Ours. Results from the analysis present findings from four areas. These areas include how well the method separates changes, how well it measures change levels, how much the form remains consistent, and how patterns remain across time. The study measures these factors by comparing data from periods that follow in sequence. Examples include comparing P2 with P1 and comparing P5 with P4. The analysis also finds the average and the variation across all periods in the sequence. This provides evidence that conclusions from the study apply to monitoring across time.

A. EVALUATION METRICS FOR CHANGE SEGMENTATION PERFORMANCE

Change segmentation performance measures the model's ability to distinguish changed regions from unchanged regions, which constitutes a fundamental requirement for long-term monitoring. In this study, IoU and F1-score are adopted to quantitatively evaluate different methods, and the per-epoch comparison results as well as the averaged performance over the sequence are reported. The data in Figure 1 show that change measures increase over time as monitoring continues across multiple periods. The IoU measure and F1 measure for the three approaches show increases overall. This pattern appears consistent with the design of the study, which uses settlement values and deflection values that increase across periods. However, the C2P-threshold approach that uses a fixed value produces IoU results below point six zero for all periods. This indicates that the approach shows high response to noise in data and to missing regions in the data. Also, using a threshold that remains fixed presents difficulties for different change measures across periods, and this produces both false detection of change and false absence of detection.

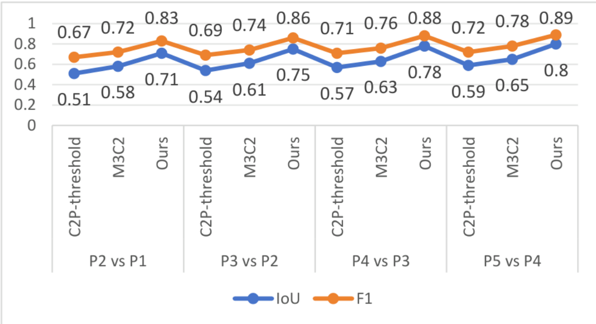

FIGURE 1. Comparison of segmentation performance at different periods

The M3C2 approach provides some reduction in effects from variations in point distribution, but registration errors and regions with missing data still affect results. This produces only limited improvement in the measure of segmentation. In contrast, the approach presented in this study shows a significant advantage across all periods. The IoU measure and F1 measure for this approach reach ranges of point seven one to point eight zero and point eight three to point eight nine. These findings indicate that using deep features from data and combining information across time periods allows the method to use structural information from different monitoring periods. This allows more consistent identification of regions showing change even when noise and missing data occur. The variation measure in Table 1 indicates that performance variation for the approach in this study remains within a limited range. This suggests that the approach provides consistent results for monitoring across long periods.

TABLE 1. Overall Segmentation Performance Statistics

Method	IoU(Mean±std)	F1(Mean±std)
C2P-threshold	0.552 ± 0.031	0.698 ± 0.022
M3C2	0.618 ± 0.028	0.750 ± 0.025
Ours	0.760 ± 0.038	0.865 ± 0.025

B. EVALUATION METRICS FOR CHANGE MAGNITUDE ESTIMATION ACCURACY

Long-term monitoring requires not only localizing changed regions but also quantitatively estimating change magnitudes, in order to support structural safety analysis and trend interpretation. In this study, the MAE metric is adopted to evaluate the regression accuracy of change magnitude estimation, and error comparisons across different epochs are reported for different methods. Since C2P-threshold is essentially a binary detection approach and does not provide a stable regression output, this part mainly compares M3C2 and Ours. It should be noted that the reported sub-millimeter MAE values are obtained under the controlled simulation environment, where the ground-truth deformation magnitudes are precisely defined. In practical field monitoring scenarios, achievable accuracy is influenced by scanner precision, environmental disturbances, and registration uncertainties. Therefore, the reported values should be interpreted primarily as an indicator of the relative performance improvement of the proposed method compared with baseline algorithms, rather than as the absolute achievable measurement accuracy in real-world bridge monitoring. Figure 2 shows that error measures decrease as monitoring continues. This pattern occurs in both the M3C2 approach and the approach that this study presents. The pattern appears because change becomes larger over time. Larger change makes values easier to measure. However, M3C2 shows error measures between one point three six and one point six two millimeters across the different monitoring periods. This range indicates that the method for determining surface direction and the procedure for measuring distance at multiple scales still show problems. Problems appear when data include noise and when features that block observation are present in the data.

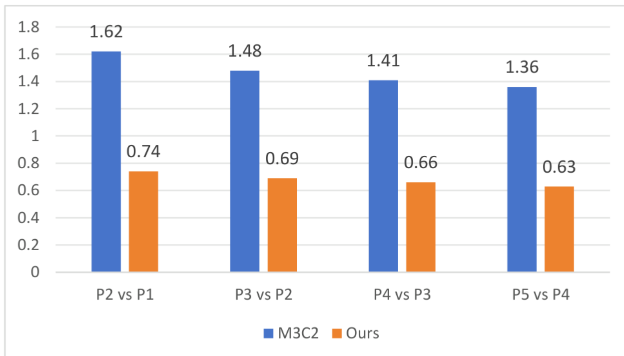

FIGURE 2. Change magnitude estimation errors (MAE, unit: mm) at different epochs

The approach that this study presents shows different results. Error measures remain between point six three and point seven four millimeters. This range shows that error decreases by more than half compared to M3C2. The results suggest that the method using learning from data captures differences in local surface features more effectively. The procedure that combines information across time also reduces variation that single measurements produce when noise affects data. Table 2 presents measures of variation that support these findings. This study proposes shows smaller variation in error measures. Smaller variation indicates that results remain more consistent between different monitoring periods. Consistency at this level provides better value for engineering use. These features meet requirements that ancient bridge monitoring establishes for measuring change at the millimeter level.

TABLE 2. Overall Statistics of Change Estimates (Mean ± Standard Deviation)

Method	MAE(Mean±std,mm)
M3C2	1.468 ± 0.114
Ours	0.680 ± 0.046

C. EVALUATION METRICS FOR GEOMETRIC CONSISTENCY

Geometric consistency measures whether the change detection results remain reasonable in terms of the overall geometric morphology. In other words, after applying the predicted changes, whether the point cloud preserves geometric consistency with the true deformed shape. In this study, Chamfer Distance (CD) is used to quantify geometric discrepancies between point clouds, where a lower value indicates a closer match between the predicted results and the ground-truth geometry. This metric is particularly suitable for evaluating structural continuity of predicted results under occlusion and missing-data conditions.

TABLE 3. Geometric consistency metrics

Method	Chamfer Distance
C2P-threshold	0.0124
M3C2	0.0098
Ours	0.0065

The method using a set value shows high response to conditions that introduce variation, and the areas of difference that it indicates appear rough and not consistent, which produces large differences in form. The approach that applies estimation at multiple levels provides improvement in local comparison, but it does not prevent the breaking of local structure that occurs from blocking and errors in alignment. The proposed method shows the lowest measure of distance, which indicates that the areas of difference that it predicts show more connection and that the edges follow the actual form of the structure. The results from prediction show higher consistency in the overall form of the structure. This finding relates to the ability of representation using multiple levels to develop patterns in local form, and it also suggests that combining data across time can provide stronger limits on the form of the structure across sequences that follow over a long period, which reduces local features that do not follow physical principles.

D. LONG-TERM MONITORING STABILITY EVALUATION

The temporal stability metric TS is defined as the proportion of points in unchanged regions that remain consistently classified as stable across all epochs. Formally, it can be expressed as

$$TS = 1 - \frac{1}{N_s T} \sum_{t=1}^{T} \sum_{i=1}^{N_s} \mathbf{1}(p_{i,t} = 1) \tag{32}$$

where N_s denotes the number of points belonging to stable regions, T is the number of monitoring epochs, and $p_{i,t}$ indicates whether point i at epoch t is falsely detected as a change point. A TS value closer to 1 therefore indicates fewer accumulated false alarms over time. A key requirement in long-term monitoring is that, in unchanged regions, detection results should remain temporally stable, preventing the accumulation of false alarms that may exaggerate structural risks. To this end, a long-term stability metric TS is introduced to evaluate the stability of different methods over multi-epoch sequences. A TS value closer to 1 indicates fewer false alarms and higher stability. As shown in Figure 3, the traditional threshold-based method yields the lowest TS, indicating that false alarms accumulate significantly over time, and noise or occlusion variations are easily misclassified as structural changes. Although M3C2 improves stability to some extent through its multi-scale strategy, it still struggles to distinguish "occasional disturbances" from "true changes" from a long-term trend perspective. In contrast, ours achieves a TS of 0.95, demonstrating a substantially lower false-alarm rate in unchanged regions over long-term sequences and an effective suppression of error propagation. This advantage mainly stems from the temporal fusion module, which introduces inter-epoch consistency learning, allowing the model to leverage long-term trend information to filter random fluctuations and thereby enhance the reliability of long-term monitoring conclusions.

FIGURE 3. Long-term stability evaluation

VI. CONCLUSION

This study proposes a deep learning framework, Bridge-TSCDNet, for detecting structural changes in multi-temporal point clouds acquired during long-term monitoring of ancient bridges. The framework integrates point-cloud feature encoding, inter-epoch differencing, and temporal feature fusion to jointly perform change segmentation and change magnitude estimation. Experimental results on controlled multi-temporal datasets demonstrate that the proposed approach improves the accuracy of change detection compared with conventional geometric methods such as C2P-threshold and M3C2. In particular, the temporal fusion mechanism effectively suppresses noise-induced fluctuations and improves long-term stability in unchanged regions. The results indicate that Bridge-TSCDNet provides a promising computational framework for monitoring structural deformation processes such as pier settlement, arch deflection, and surface deterioration in historical bridges. Future work will focus on validating the method using real-world long-term monitoring datasets and integrating structural constraints to further enhance physical interpretability.

AUTHOR CONTRIBUTIONS

Conceptualization –Jian Li and Xiyin Ma; methodology – Jian Li and Xiyin Ma; investigation – Jian Li and Xiyin Ma; writing-original draft preparation – Jian Li and Xiyin Ma. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] H. Zhang, M. Yuan, Q. Jiang, et al., “Deep learning model compression for real-time vibration monitoring of wood-boring activities,” Journal of Beijing Forestry University, vol. 43, no. 6, pp. 9, 2021, doi: 10.12171/j.1000-1522.20200100.
[2] Z. Li, C. Quan, H. Chen, et al., “Building change detection method based on point clouds generated by dense image matching,” Geospatial Information, vol. (3), pp. 022, 2024, doi: 10.3969/j.issn.1672-4623.2024.03.003.
[3] C. Meng, Y. Zhao, W. Han, et al., “Urban building change detection using airborne LiDAR point clouds based on RandLA-Net,” Remote Sensing for Natural Resources, pp. 34(4), 2022, doi: 10.6046/zrzyyg.2021402.
[4] J. Chai, Y. Zhang, Z. Yang, et al., “Multi-Level building change detection using DSM and DOM generated from UAV imagery,” Remote Sensing for Land & Resources, vol. (2), pp. 9, 2024, doi: 10.6046/zrzyyg.2023001.
[5] I. De Gélis, S. Lefèvre, and T. Corpetti, “3D urban change detection with point cloud Siamese networks,” Copernicus GmbH, 2021, doi: 10.5194/ISPRS-ARCHIVES-XLIII-B3-2021-879-2021.
[6] H. D. Seo and E. M. Kim, “Object Classification and Change Detection in Point Clouds Using Deep Learning,” LX Spatial Information Research Institute, 2020, doi: 10.22640/LXSIRI.2020.50.2.37.
[7] G. Kaur and Y. Afaq, “Developments in deep learning for change detection in remote sensing: A review,” Transactions in GIS, pp. 28(2), 2024, doi: 10.1111/tgis.13133.
[8] Léo Renaut, H. Frei, and NüchteAr, “Deep learning on 3D point clouds for fast pose estimation during satellite rendezvous,” Acta Astronautica, vol. 232, pp. 231-243, 2025, doi: 10.1016/j.actaastro.2025.03.009.
[9] J. Li, P. Fearnhead, P. Fryzlewicz, et al., “Automatic change-point detection in time series via deep learning,” Journal of the Royal Statistical Society: Series B (Statistical Methodology), pp. 86(2), 2024, doi: 10.1093/jrsssb/qkae004.
[10] T. D. Ryck, M. D. Vos, and A. Bertrand, “Change Point Detection in Time Series Data Using Autoencoders with a Time-Invariant Representation,” IEEE Transactions on Signal Processing, 2021, doi: 10.1109/TSP.2021.3087031.
[11] H. He, J. Yan, D. Liang, et al., “Time-Series land cover change detection using deep learning-based temporal semantic segmentation,” Remote Sensing of Environment, vol. 305, Art. no. 114101, 2024, doi: 10.1016/j.rse.2024.114101.
[12] T. Han and G. A. Sánchez-Azofeifa, “A Deep Learning Time Series Approach for Leaf and Wood Classification from Terrestrial LiDAR Point Clouds,” Remote Sensing, vol. 14, no. 13, pp. 3157, 2022, doi: 10.3390/rs14133157.
[13] S. Shirowzhan, M. Samad, et al., “Comparative analysis of machine learning and point-based algorithms for detecting 3D changes in buildings over time using bi-temporal lidar data - ScienceDirect,” Automation in Construction, vol. 105, Art. no. 102841, 2019, doi: 10.1016/j.autcon.2019.102841.
[14] A. Floris, L. Frittoli, D. Carrera, et al., “Composite Convolution: A Flexible Operator for Deep Learning on 3D Point Clouds,” Pattern Recognition, Art. no. 110557, 2024, doi: 10.1016/j.patcog.2024.110557.
[15] T. Chaton, N. Chaulet, S. Horache, et al., “Torch-Points3D: A Modular Multi-Task Frameworkfor Reproducible Deep Learning on 3D Point Clouds,” 2020 International Conference on 3D Vision (3DV), 2020, doi: 10.1109/3DV50981.2020.00029.