

Edge-constrained temporal superpixel segmentation and graph-structured energy optimization for PolSAR change detection

Nengcai Li ^a, Deliang Xiang ^a, Huaiyue Ding ^a, Yuzhen Xie ^a, Yi Su ^b

^a College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, China

^b College of Electronic Science and Technology, National University of Defense Technology, Changsha, 410073, China



ARTICLE INFO

Keywords:

Polarimetric Synthetic Aperture Radar (PolSAR)
Temporal Superpixel Segmentation
Change detection
Edge constrain
Graph-structured optimization

ABSTRACT

Polarimetric Synthetic Aperture Radar (PolSAR) has emerged as a vital tool for dynamic surface monitoring, owing to its ability to precisely characterize land cover scattering properties. However, conventional PolSAR change detection methods predominantly rely on pixel- or region-level direct comparisons, rendering them sensitive to speckle noise and multi-temporal radiometric inconsistencies. In addition, existing superpixel generation algorithms typically neglect temporal information and edge strength, resulting in suboptimal segmentation accuracy. To overcome these limitations, this paper introduces a novel edge-constrained temporal superpixel generation method. A new temporal polarimetric similarity metric is proposed to emphasize significant temporal variations, while an edge constraint mechanism is incorporated to prevent superpixels from crossing semantic boundaries, thereby improving segmentation fidelity. Building upon the generated superpixels, we develop a graph-structured energy optimization framework for PolSAR change detection. In this framework, superpixels serve as the fundamental processing units to construct a topological representation that integrates both temporal feature similarity and spatial adjacency. A cross-node similarity metric is further designed to enhance the detection of weak scattering changes, and a global energy function is formulated to suppress noise while preserving the structural integrity of changed regions. Extensive experiments on five PolSAR datasets validate the superior performance of the proposed approach, demonstrating significant improvements in noise suppression, temporal feature representation, and change detection accuracy over existing state-of-the-art methods. Specifically, the proposed superpixel segmentation method achieves an average improvement of 6.62% in boundary recall and 1.46% in achievable segmentation accuracy compared to the TSPol-ASLIC algorithm. For the change detection task, the proposed framework achieves a peak overall accuracy of 0.9802, an F1-score of 0.9431, and a kappa coefficient of 0.9311, significantly outperforming conventional pixel-level approaches. The code will be available at https://github.com/linengcai/Pol_ECTSP_GSEO.

1. Introduction

Synthetic Aperture Radar (SAR) employs active microwave transmission and receives backscattered signals from land surfaces, enabling high-resolution Earth observation under all-weather, day-and-night conditions. Owing to its unique capabilities, SAR has been extensively utilized in applications such as topographic mapping and disaster monitoring (Wang et al., 2024b; Xiang et al., 2022). Conventional SAR systems operate with a single polarization channel, which limits their capacity to characterize complex land cover features. In contrast, Polarimetric SAR (PolSAR) captures the full polarimetric scattering matrix of targets, offering a more comprehensive understanding of geometric structures, dielectric properties, and scattering mechanisms. This capability greatly enhances the accuracy of land cover classification

and dynamic environmental monitoring (Farooq and Manocha, 2024; Xiang et al., 2024).

Change detection, a fundamental task in PolSAR image analysis, aims to identify surface changes over the same geographical area across different time periods (Bai et al., 2023). However, PolSAR data are inherently affected by speckle noise, high-dimensional polarimetric scattering features, and multi-temporal radiometric inconsistencies, which pose significant challenges to traditional pixel-wise comparison-based approaches. Speckle noise increases false alarms and missed detections, while radiometric inconsistencies can obscure true land cover changes. In recent years, superpixel segmentation has gained traction in PolSAR change detection due to its ability to reduce redundancy while preserving important structural information (Cheng

* Corresponding author.

E-mail address: xiangdeliang@buct.edu.cn (D. Xiang).

et al., 2024). By grouping pixels into perceptually meaningful regions, superpixels simplify processing complexity and enhance regional homogeneity and edge representation, ultimately improving the quality of input data for subsequent change detection tasks.

1.1. Related work in PolSAR superpixel segmentation

Superpixel segmentation clusters spatially adjacent pixels with similar features into homogeneous regions, serving as an efficient intermediate representation in image processing. Its primary advantages include: (1) reducing data redundancy and computational complexity, (2) preserving spatial continuity to mitigate noise, and (3) providing flexible region-based units suitable for advanced models such as deep learning and graph neural networks. In PolSAR image analysis, superpixel techniques have been successfully applied to semantic segmentation (Li et al., 2024a), target detection (Deng et al., 2024; Wu et al., 2024), and land cover classification (Shi et al., 2023; Wang et al., 2024a), solidifying their role as effective preprocessing tools.

Among the various superpixel algorithms, Simple Linear Iterative Clustering (SLIC) is widely used in optical image processing for its computational efficiency and straightforward parameter tuning (Barcelos et al., 2024). However, when directly applied to PolSAR data (e.g., Pauli pseudo-color images) (Zhang et al., 2021; Guo et al., 2024), SLIC's reliance on spatial-color features makes it vulnerable to speckle noise and poorly adapted to polarimetric statistical properties. To overcome these limitations, improved SLIC variants have been proposed that integrate polarimetric distance measures — such as Wishart distance (Feng et al., 2014), Revised Wishart (RW) distance (Qin et al., 2014), and Bartlett distance (Song et al., 2015) — into the clustering framework, thereby enhancing segmentation consistency in homogeneous regions. (Yin et al., 2021) further conducted a systematic comparison of Hotelling–Lawley Trace (HLT) distance and Symmetric Revised Wishart (SRW) distance, demonstrating the superiority of SRW distance in preserving boundaries. Li et al. (2023b) introduced the Determinant Ratio Test (DRT) distance and demonstrated its advantages in computational efficiency and detail preservation in heterogeneous regions through experimental validation. Nevertheless, these methods are built on the Wishart distribution assumption, which may not hold in heterogeneous areas. To address this, Xiang et al. (2017) proposed a similarity metric based on the Spherically Invariant Random Vector (SIRV) model, while (Li et al., 2025) utilized polarimetric target decomposition and Riemannian geometry to design improved similarity metrics.

Despite their effectiveness in single-temporal segmentation, these approaches often fail to capture the joint statistical characteristics of multi-temporal data, resulting in misaligned superpixel boundaries. To address this, some studies have proposed post-segmentation fusion methods. (Gong et al., 2017) independently applied SLIC to bi-temporal images and aligned their superpixels via boundary integration strategy. (Lei et al., 2019) proposed a superpixel refinement strategy, merging superpixels from two temporal instances to achieve a multi-scale segmentation. However, these methods rely on manually designed fusion rules and do not fully exploit joint temporal features. To overcome these limitations, some studies employ joint temporal segmentation strategies to generate temporal superpixels in a single step. (Xie et al., 2015) and Bao et al. (2017) constructed a joint covariance matrix by stacking multi-temporal covariance matrices and defined a temporal similarity metric based on the Wishart distribution assumption. However, Gao et al. (2021) pointed out that such temporal similarity metrics essentially perform linear weighting of single-temporal similarities, failing to reflect meaningful temporal evolution. To address this, they proposed a root mean square-based temporal polarimetric similarity metric, which highlights the time sequence of significant change.

1.2. Related work in PolSAR change detection

Compared to optical imagery, PolSAR change detection is more challenging due to speckle noise and the complexity of scattering mechanisms, resulting in relatively limited research. Traditional PolSAR change detection methods can be broadly categorized into pixel-level and region-level approaches. The pixel-level methods include statistical information-based and hypothesis testing-based techniques. The former utilize information-theoretic measures — such as entropy (Nascimento et al., 2018), mutual information (Gueguen et al., 2011), and Kullback–Leibler divergence (Erten et al., 2011) — to quantify differences between bi-temporal pixels. The latter test the equality of covariance matrices to detect changes. (Conradsen et al., 2003) pioneered the likelihood ratio test (LRT) with an asymptotic distribution for hypothesis testing, followed by developments such as the complex HLT (Akbari et al., 2013) and the DRT (Bouhlel et al., 2020). However, the presence of speckle noise and temporal radiometric inconsistencies, along with the lack of spatial contextual information, significantly compromises the robustness of these methods.

To address these drawbacks, region-level methods have emerged as a promising alternative. By incorporating spatial context, these methods suppress noise more effectively. (Xie et al., 2015) proposed a superpixel-based PolSAR change detection framework, where a modified SLIC algorithm was used for joint segmentation of bi-temporal images. They also introduced a color difference map to assist in analyzing polarimetric statistical differences, combined with a majority voting strategy and a difference map fusion approach to enhance detection accuracy. Experimental results demonstrated that superpixel-level processing significantly reduced the speckle noise observed in pixel-level methods. To address the complexity of heterogeneous land cover scattering characteristics, Yang et al. (2016) employed a Wishart mixture model to characterize the statistical distribution within superpixel regions. They quantified inter-region differences using Cauchy–Schwarz divergence and incorporated Markov random field (MRF) optimization to refine threshold segmentation results. This method exhibited greater robustness than traditional pixel-level approaches, particularly in highly heterogeneous regions such as forest–urban boundaries. More recently, Zhang et al. (2022) systematically compared pixel- and region-level methods, highlighting the latter's advantages in noise suppression and detection reliability through the use of averaged coherence matrices.

In recent years, deep learning-based methods have significantly advanced the field of change detection in remote sensing. Supervised approaches have been widely adopted for optical image change detection, benefiting from the availability of large-scale annotated datasets. However, their application to hyperspectral, SAR, and particularly PolSAR image change detection remains limited due to the scarcity of reliable pixel-level labels in these domains. To mitigate this limitation, Lv et al. (2023) proposed an iterative training sample augmentation (ITSA) strategy that expands a small set of manually labeled samples by automatically generating new ones, demonstrating its effectiveness when combined with deep neural networks. Other studies employ traditional unsupervised techniques, such as clustering, to produce pseudo-labels that are subsequently used to supervise the training of deep models (Xie et al., 2024). To further reduce annotation dependency, semi-supervised methods have gained increasing attention. These approaches utilize a small number of labeled samples in conjunction with large volumes of unlabeled data to improve generalization. Lin and Lin (2023) combines graph neural networks with convex deep learning to achieve accurate change detection using only a small amount of labeled data, demonstrating strong performance under semi-supervised settings. Li et al. (2024c) achieves robust and accurate semi-supervised SAR image change detection by leveraging structure-optimized complex-valued graph contrastive learning under limited labels and heavy speckle noise. Han et al. (2024) enhances change detection accuracy under limited annotations by employing

a coarse-to-fine framework with consistency regularization in high-resolution remote sensing images. In parallel, self-supervised learning has emerged as a promising paradigm for learning transferable representations without relying on any labeled data. Li et al. (2024b) introduces a causal graph contrastive learning framework to effectively capture causal differences between changed and unchanged regions, enabling efficient and generalizable self-supervised SAR image change detection. Zhao et al. (2024) constructs self-supervised tasks guided by image-level supervision, effectively avoiding costly pixel-level annotations and significantly reducing labeling costs for change detection. Zhou et al. (2025) leverages depth prompting to guide change detection, enabling self-supervised remote sensing change detection without the need for pixel-level annotations. Despite these advances, most deep learning-based change detection methods remain dependent on large labeled datasets and lack interpretability. Moreover, they are seldom designed specifically for PolSAR imagery, which poses unique challenges such as speckle noise, radiometric inconsistency, and complex polarimetric scattering mechanisms. These limitations highlight the need for unsupervised and physically interpretable methods that are tailored to the characteristics of PolSAR data.

1.3. Motivation and contribution of this paper

Change detection in PolSAR images plays a crucial role in dynamic land cover monitoring. However, existing methods still face significant challenges. Pixel-level techniques such as HLT (Akbari et al., 2013) and DRT (Bouhlel et al., 2020) are capable of detecting fine-grained changes but are highly sensitive to speckle noise and radiometric inconsistencies, leading to noise interference and fragmentation in the detection results. Although superpixel-based region-level methods mitigate noise interference by maintaining spatial consistency, their performance remains constrained by three key bottlenecks:

- 1. Insufficient fusion of temporal information:** Existing superpixel generation techniques struggle to capture joint polarimetric features across time, resulting in inconsistent segmentation boundaries in regions with subtle or gradual changes.
- 2. Lack of dynamic edge constraints:** Most superpixel methods use edge information only during initialization and do not enforce edge adherence during clustering, often resulting in superpixels crossing land cover boundaries.
- 3. Inadequate exploitation of topological structure:** Many current region-level approaches rely solely on direct feature differences without modeling the underlying spatial or cross-temporal relationships between image regions. As a result, they are often sensitive to temporal radiometric inconsistencies and struggle to detect weak scattering changes.

To address these challenges, this paper proposes a robust PolSAR change detection framework, with the following contributions:

- 1. Edge-Constrained Temporal Superpixel Generation:** We propose a temporal similarity metric based on Jensen-Bregman LogDet (JBLD) divergence using a maximum selection strategy to highlight significant changes while suppressing minor fluctuations. A dynamic edge constraint mechanism is introduced by embedding a normalized edge strength map into the SLIC clustering metric, which penalizes pixel assignments that cross strong edges, thereby ensuring that superpixels strictly adhere to actual land cover boundaries.
- 2. Dual-Temporal Topological Representation and Association Analysis:** We model superpixels as nodes to construct a graph representation that simultaneously captures temporal feature similarity and spatial adjacency. Beyond conventional intra-node comparisons, we introduce a novel cross-node feature similarity metric that models relationships between temporally and spatially adjacent superpixels. This enriched topological structure

enables the graph to encode both short-range and long-range dependencies across time, providing a more comprehensive context for semantic reasoning. Such modeling significantly enhances the detection of weak scattering changes while improving robustness to temporal radiometric inconsistencies.

- 3. Graph-Structured Energy Optimization Framework:** To leverage the constructed topological graph, we formulate change detection as a global energy minimization problem that jointly considers node costs (imposing a change sparsity prior) and edge costs (enforcing topological consistency across temporal and spatial domains). The edge cost functions are explicitly designed to incorporate intra-node, inter-node, and cross-node similarity constraints, allowing the model to better capture structural coherence and suppress noise-induced label fragmentation. The optimization is efficiently solved using the Quadratic Pseudo-Boolean Optimization (QPBO) algorithm (Kolmogorov and Rother, 2007), which guarantees global optimality and ensures label smoothness along complex topological boundaries.

The remainder of this paper is organized as follows: Section 2 details the proposed methodology, including the edge-constrained temporal superpixel generation (Section 2.1) and the graph-structured energy optimization-based change detection framework (Section 2.2). Section 3 presents experimental results and comparative evaluations on five real PolSAR datasets. Section 4 concludes the paper.

2. Methodology

The overall workflow of the proposed method is shown in Fig. 1, comprising two main components: temporal superpixel segmentation and graph-structured energy optimization-based change detection. In the first stage, temporal JBLD divergence is used to measure similarity between multi-temporal covariance matrices, while a dynamic edge constraint mechanism prevents superpixels from crossing actual land cover boundaries, ensuring accurate segmentation. In the second stage, superpixels serve as graph nodes from which polarimetric features are extracted. A graph is constructed by integrating spatial adjacency and feature similarity, and the corresponding node and edge cost functions are formulated. The QPBO algorithm is then employed to solve the global energy minimization problem, producing a change map with reduced noise and well-preserved boundaries.

2.1. Edge-constrained temporal superpixel generation for PolSAR images

Building upon our previously proposed single-temporal SLIC-based method (Li et al., 2025), this paper introduces a superpixel generation approach for temporal PolSAR image. The proposed method incorporates two key improvements. First, a temporal JBLD divergence-based similarity measure is defined based on a maximum-value criterion. This measure is utilized both in the generation of the temporal edge strength map and during the SLIC clustering process. Second, a dynamic edge constraint mechanism is introduced, which significantly enhances the alignment of generated superpixels with the actual boundaries of land cover targets. The following subsections describe these components in detail.

2.1.1. Similarity measure between temporal polarimetric covariance matrices

For reciprocal targets, a fully polarimetric SAR system measures the backscattered signal from the target, resulting in a complex scattering vector k :

$$k = [S_{\text{HH}}, \sqrt{2}S_{\text{HV}}, S_{\text{VV}}]^T \quad (1)$$

Here, S_{HH} , S_{HV} , and S_{VV} represent the complex scattering coefficients, where the subscript H denotes horizontal polarization, and the subscript V denotes vertical polarization, and the superscript T indicates

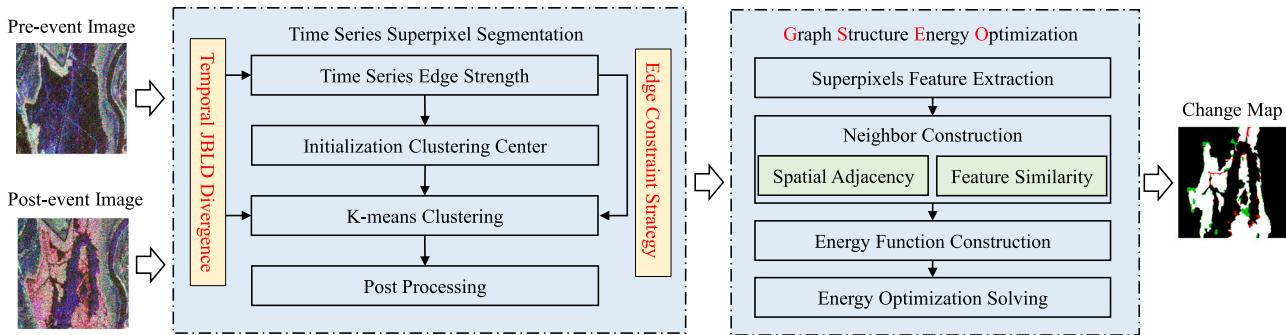


Fig. 1. Overall workflow of the proposed algorithm.

the transpose operation. For multi-look PolSAR data, in order to describe the statistical properties of the target, the second-order statistics of the complex scattering vector k are typically used. This is commonly represented by the polarimetric coherence matrix or the covariance matrix. The polarimetric covariance matrix is defined as:

$$\mathbf{C} = \langle \mathbf{k}\mathbf{k}^H \rangle \quad (2)$$

where $\langle \cdot \rangle$ denotes the sample ensemble average, and the superscript H represents the Hermitian operator. It can be readily proven that the polarimetric covariance matrix is a Hermitian symmetric positive definite (SPD) matrix (Gao et al., 2021).

The similarity measure plays a crucial role in SLIC-based superpixel generation. On the one hand, it is used in the computation of the edge strength map. On the other hand, it constitutes an essential part of the SLIC clustering similarity metric. For PolSAR data, traditional similarity measures include the Wishart distance, RW distance, and SRW distance, etc (Qin et al., 2022). Among them, the SRW distance is a robust and commonly used similarity measure, defined as:

$$D_{SRW}(\mathbf{C}_1, \mathbf{C}_2) = \frac{1}{2} (\mathbf{C}_1^{-1} \mathbf{C}_2 + \mathbf{C}_2^{-1} \mathbf{C}_1 - 2\mathbf{I}) \quad (3)$$

where \mathbf{C}_1 and \mathbf{C}_2 denote two polarimetric covariance matrices, and \mathbf{I} is the identity matrix of the same dimension. It is evident that the computation of Eq. (3) involves matrix inversion, which leads to relatively high computational complexity. To address this issue for SPD matrices, the JBLD divergence was introduced as an alternative similarity measure (Cherian et al., 2012). This metric significantly improves computational efficiency while maintaining comparable accuracy and robustness to the SRW distance. The JBLD divergence is defined as:

$$D_{JBLD}(\mathbf{C}_1, \mathbf{C}_2) = \log \left| \frac{\mathbf{C}_1 + \mathbf{C}_2}{2} \right| - \frac{1}{2} \log |\mathbf{C}_1 \mathbf{C}_2| \quad (4)$$

where $|\cdot|$ denotes the determinant of a matrix, and $\log(\cdot)$ represents the natural logarithm. Similar to the D_{SRW} , the D_{JBLD} satisfies all axioms of a metric except the triangle inequality. Given that the determinant of an SPD matrix in Eq. (4) can be efficiently computed using Cholesky decomposition, while Equation (3) requires matrix inversion, the JBLD divergence offers significantly higher computational efficiency compared to the SRW distance.

Based on the statistical independence among multi-temporal PolSAR data, a joint covariance matrix can be constructed by stacking multi-temporal PolSAR data (Gao et al., 2021), thereby effectively characterizing the scattering properties across multiple temporal acquisitions:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{C}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{C}_N \end{bmatrix}_{3N \times 3N} \quad (5)$$

where \mathbf{C}_i denotes the covariance matrix of the i th temporal acquisition, $\mathbf{0}$ denotes a 3×3 zero matrix, and N is the number of temporal acquisitions. It can be easily proven that the joint covariance matrix \mathbf{Z}

is also a SPD matrix. Thus, the distance between two joint covariance matrices can be computed using Eq. (4):

$$D_{JBLD}(\mathbf{Z}_1, \mathbf{Z}_2) = \log \left| \frac{\mathbf{Z}_1 + \mathbf{Z}_2}{2} \right| - \frac{1}{2} \log |\mathbf{Z}_1 \mathbf{Z}_2| \quad (6)$$

where \mathbf{Z}_1 and \mathbf{Z}_2 denote the joint covariance matrix of the two regions, respectively. Due to the joint covariance matrix is a block matrix, Eq. (6) can be simplified as follows.

$$D_{JBLD}(\mathbf{Z}_1, \mathbf{Z}_2) = \sum_{i=1}^N \left(\log \left| \frac{\mathbf{C}_{i,1} + \mathbf{C}_{i,2}}{2} \right| - \frac{1}{2} \log |\mathbf{C}_{i,1} \mathbf{C}_{i,2}| \right) \quad (7)$$

where $\mathbf{C}_{i,j}$ represents the covariance matrix of the j th region in the i th time sequence. In order to avoid the influence brought by the number of sequences, Eq. (7) can be normalized as:

$$\begin{aligned} D_{JBLD}^{mean}(\mathbf{Z}_1, \mathbf{Z}_2) &= \frac{1}{N} \sum_{i=1}^N \left(\log \left| \frac{\mathbf{C}_{i,1} + \mathbf{C}_{i,2}}{2} \right| - \frac{1}{2} \log |\mathbf{C}_{i,1} \mathbf{C}_{i,2}| \right) \\ &= \frac{1}{N} \sum_{i=1}^N D_{JBLD}(\mathbf{C}_{i,1}, \mathbf{C}_{i,2}) \end{aligned} \quad (8)$$

It is evident that Eq. (8) essentially performs a simple equal-weighted summation of the distance metrics for each temporal acquisition. This approach poses a clear limitation: it assigns equal weight to the contributions of both significantly fluctuating and weakly fluctuating temporal acquisitions, thereby failing to accurately reflect the temporal variation characteristics. To address this issue, this study proposes a temporal distance metric based on the maximum criterion:

$$D_{JBLD}^{max}(\mathbf{Z}_1, \mathbf{Z}_2) = \max_{i=1, \dots, N} D_{JBLD}(\mathbf{C}_{i,1}, \mathbf{C}_{i,2}) \quad (9)$$

Compared to Eq. (8), which computes the average divergence across all temporal acquisitions, Eq. (9) focuses on extracting the most significant difference by selecting the maximum divergence across time. This design addresses a common limitation in existing superpixel generation techniques, which typically rely on mean-based or linearly weighted temporal similarity metrics. Such approaches often fail to capture joint polarimetric characteristics, especially in regions undergoing subtle or gradual changes, leading to boundary inconsistencies and segmentation artifacts. In contrast, the proposed maximum-based similarity measure enhances sensitivity to temporally localized changes while suppressing the influence of weakly fluctuating acquisitions that resemble background noise. This strategy not only reduces the risk of misinterpreting noise as change but also enables better delineation of land cover boundaries that evolve over time. As a result, it significantly improves the accuracy and reliability of temporal edge detection and superpixel segmentation in dynamic scenes.

2.1.2. Edge constraint based SLIC clustering measure

While the temporal similarity measure between polarimetric covariance matrices captures changes in scattering behavior, it may still be sensitive to noise and boundary ambiguity. To address this, we introduce an edge constraint to improve spatial coherence in the clustering process. The edge features of an image can effectively represent

the boundary information of potential targets (Cour et al., 2005; Su et al., 2023; Sun et al., 2024b), which is especially important when the background clutter and target radiation intensity are similar. In edge detection frameworks based on bilateral window filters, the Gaussian window is more robust than rectangular windows as it avoids the false local maxima caused by strong noise, leading to more robust edge detection (Li et al., 2025). The Gaussian filter assigns higher weights to pixels near the center, and the horizontal Gaussian window function is defined as:

$$W(x, y) = \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{2\pi}\sigma_y} \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right) \quad (10)$$

where x, y represent the relative coordinates, and σ_x and σ_y control the rate of weight variation in the two directions of the window. For single-time PolSAR data, the average covariance matrix of the region on both sides of the central line is calculated as follows:

$$\hat{\mathbf{C}}_i = \frac{\sum_{(x,y) \in \Omega_i} W^{\theta_f}(x, y) \mathbf{C}(x, y)}{\sum_{(x,y) \in \Omega_i} W(x, y)}, i = 1, 2 \quad (11)$$

$$W^{\theta_f}(x, y) = W(x \cos \theta_f - y \sin \theta_f, x \sin \theta_f + y \cos \theta_f) \quad (12)$$

where θ_f denotes the orientation angle of the central line, and $\Omega_i, i = 1, 2$ represent the pixel sets on both sides of the central line. By combining Eqs. (9) and Eqs. (11), the final formulation for computing the temporal edge strength of PolSAR images is given as follows:

$$\text{EDGE} = \max_{i=1,\dots,N} D_{\text{JBLD}}(\hat{\mathbf{C}}_{i,1}, \hat{\mathbf{C}}_{i,2}). \quad (13)$$

Here, $\hat{\mathbf{C}}_{i,1}$ and $\hat{\mathbf{C}}_{i,2}$ denote the bilateral average covariance matrices at the i th temporal instance, computed using Eq. (11). By applying Equation (13) across the entire image, the edge strength map EDGE can be obtained.

Traditional SLIC-based superpixel generation methods exhibit a notable limitation in their use of edge information. Typically, edge features are only utilized during the initialization stage to avoid placing cluster centers on strong edges. However, these features are generally ignored during the subsequent iterative clustering process, which often results in superpixels crossing land cover boundaries. To overcome this limitation, we propose a dynamic edge constraint mechanism that explicitly incorporates edge strength information into the SLIC clustering similarity metric at every iteration. By embedding an edge-guided penalty term directly into the distance computation, the clustering process becomes more sensitive to structural boundaries throughout its evolution. This ensures that the resulting superpixels better adhere to actual image edges and significantly improves boundary-preserving segmentation quality. The edge constraint term is defined as follows:

$$D_{\text{EDGE}} = \max_{p \in \text{line}(i,j)} \text{EDGE}_{\text{norm}}(p) \quad (14)$$

where D_{EDGE} denotes the normalized edge strength map, $\text{line}(i, j)$ represents the line segment connecting any two pixels i and j , and p denotes the pixels lying on this line segment.

Fig. 2(c) and Fig. 2(d) illustrate the edge strength map of the image patch and the edge constraint term centered at pixel i , respectively. As shown in Fig. 2(c), the edges of land cover objects in temporal PolSAR images are accurately extracted. Fig. 2(d) further demonstrates that the edge constraint term remains low when two pixels lie within the same region, but increases significantly when the pixels span across an image edge. Notably, the proposed edge-constrained mechanism is not limited to the current framework and can be readily integrated into other SLIC-based superpixel segmentation methods to improve edge adherence and spatial regularity. Since it operates independently of specific distance metrics or clustering strategies, this mechanism can serve as a general enhancement module for a wide range of SLIC-derived algorithms. Leveraging this property and building upon our previous work (Li et al., 2025), the SLIC clustering similarity measure for temporal PolSAR images is defined as follows:

$$D_{\text{SLIC}} = D_{\text{JBLD}}^{\max} \cdot (1 + D_p) + \alpha \cdot D_{\text{EDGE}} + \beta \cdot (D_S / S) \quad (15)$$

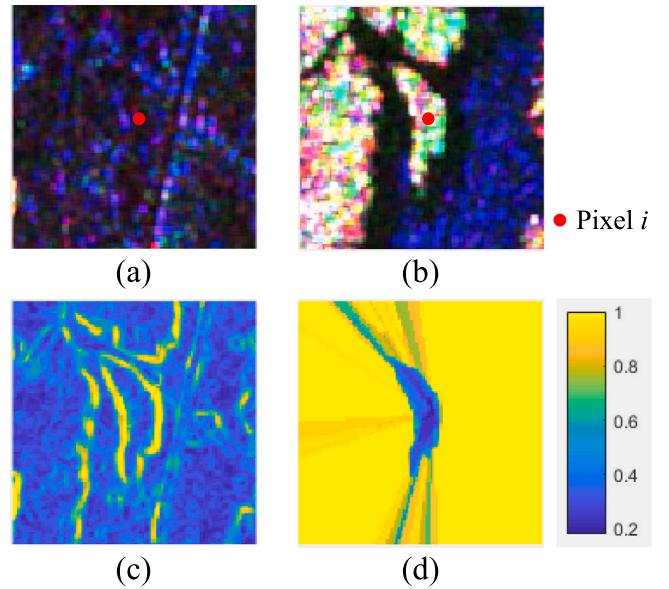


Fig. 2. Edge constraint construction. (a),(b) Pauli image patches from two temporal acquisitions; (c) edge strength map computed using the proposed method; (d) edge constraint term centered at pixel i .

The SLIC clustering measure D_{SLIC} consists of three components: feature similarity, edge constraint term, and spatial similarity. In Eq. (15), α and β are manually set coefficients that control the weights of the edge constraint term and spatial similarity, respectively. In the experiments presented in this paper, α is set to 1.5 and β is set to 1.0. Specifically, we found that setting α within the range of 0.7 to 2.0 ensures that the edge constraint mechanism effectively influences superpixel boundary adherence without overly dominating the similarity metric. Meanwhile, we recommend setting β within the range of 0.8 to 1.2 to maintain superpixel compactness while avoiding overemphasis on spatial proximity. The parameter S denotes the sampling interval, which is computed based on the desired number of superpixels. The feature similarity is primarily governed by the similarity measure defined in Eq. (9), which enhances computational efficiency while maximizing the utilization of the temporal differences in the most fluctuating sequence. D_p denotes the maximum absolute percentage power similarity, which is defined as:

$$D_p = \max_{i=1,\dots,N} \frac{|P_{i,1} - P_{i,2}|}{\max(P_{i,1}, P_{i,2})} \quad (16)$$

where $P_{i,1}$ and $P_{i,2}$ represent the total scattering power of the two pixels in the i th temporal image, respectively. The value of D_p ranges from [0, 1] and the factor $(1 + D_p)$ provides a maximum gain of two times for feature similarity. Finally, spatial similarity is incorporated to improve the compactness of the generated superpixels:

$$D_S = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (17)$$

where (x_1, y_1) and (x_2, y_2) denote the spatial coordinates of two pixels. When a prominent edge exists between the superpixel center and the pixel to be classified, even if the pixel exhibits high feature similarity with the superpixel center, the proposed edge-constrained term introduces a large penalty to effectively prevent the pixel from being misclassified into the current superpixel region. This ensures the structural integrity of object boundaries.

2.2. Graph-structured energy optimization-based change detection for PolSAR images

Based on the edge-constrained temporal superpixels generated in Section 2.1, we further construct a graph representation to facilitate

robust change detection. Conventional change detection methods primarily rely on pixel-level or region-level direct comparison strategies, identifying changed areas by analyzing feature differences at corresponding locations in bi-temporal images. However, such approaches are highly susceptible to imaging noise and radiometric inconsistencies. In contrast, graph-based representations offer greater robustness by capturing the topological relationships and spatial connectivity among internal image structures. Leveraging this advantage, the proposed method constructs a topological representation by treating superpixels as fundamental processing units. A graph-cut algorithm is then employed to globally optimize a defined energy function, enabling robust and noise-resistant change detection.

2.2.1. Topological representation of bi-temporal PolSAR images

Given their ability to preserve land cover boundaries, exploit contextual information, and substantially reduce computational complexity, superpixels are adopted in this study as the fundamental processing units. Pairwise similarity relationships are then established based on both spatial proximity and feature similarity, effectively capturing the topological structure of the image.

First, using the method from the previous section, the dual-temporal PolSAR data is segmented into N_S superpixels \mathbf{X}_i^{t1} and \mathbf{X}_i^{t2} , where $i = 1, 2, \dots, N_S$, with \mathbf{X}_i^{t1} and \mathbf{X}_i^{t2} representing the same geographic area in different temporal superpixel blocks. Since superpixels are typically homogeneous, it is assumed that each superpixel represents the same land cover class. To construct graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ that characterizes the topological structure of the PolSAR image, we first extract polarimetric scattering features $\mathbf{F} = [P_s, P_d, P_{od}, P_q, P_p, P_h]$ from the PolSAR data using our previously proposed polarimetric target decomposition method (Li et al., 2023a), followed by global normalization. It is worth noting that the proposed framework is compatible with various polarimetric decomposition methods. While we adopt our previous method in this work to demonstrate its effectiveness, commonly used alternatives such as Pauli or Yamaguchi decomposition can also be employed without requiring changes to the subsequent graph construction or optimization pipeline. Then, we take the average polarimetric scattering features of superpixels as the feature vectors of nodes and construct the graph nodes as follows:

$$\mathcal{V} = \{\mathbf{V}_i | i = 1, 2, \dots, N_S\}. \quad (18)$$

Here, $\mathbf{V}_i = [\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2}]$, where $\bar{\mathbf{X}}_i^{t1}$ and $\bar{\mathbf{X}}_i^{t2}$ represent the average polarimetric scattering features of the i th superpixel at two different time sequence, respectively. In the subsequent sections, these features will be used to refer to their corresponding superpixels. The edges \mathcal{E} of the graph are constructed based on spatial similarity and feature similarity:

$$\mathcal{E} = \mathcal{E}_F^{t1} \cup \mathcal{E}_F^{t2} \cup \mathcal{E}_S \quad (19)$$

where \mathcal{E}_F^{t1} and \mathcal{E}_F^{t2} are edges constructed based on feature similarity in two phases respectively, which are defined as follows:

$$\mathcal{E}_F^{t1} = \{(\mathbf{V}_i, \mathbf{V}_j) | i = 1, 2, \dots, N_S, j \in \mathcal{N}_i^{t1}\} \quad (20)$$

$$\mathcal{E}_F^{t2} = \{(\mathbf{V}_i, \mathbf{V}_j) | i = 1, 2, \dots, N_S, j \in \mathcal{N}_i^{t2}\} \quad (21)$$

In Eq. (20), the set \mathcal{N}_i^{t1} denotes the set of K-nearest neighbors of the i th superpixel in the image at time $t1$, determined based on feature similarity. The feature similarity between any two superpixels is defined as:

$$d_F(\bar{\mathbf{X}}_i, \bar{\mathbf{X}}_j) = 1 - 2 \times \text{mean}(\min(\bar{\mathbf{X}}_i, \bar{\mathbf{X}}_j) ./ (\bar{\mathbf{X}}_i + \bar{\mathbf{X}}_j)) \quad (22)$$

where $./$ denotes element-wise division between vectors. For the i th superpixel at time $t1$, feature similarities with all other superpixels in the $t1$ -temporal image are computed using Eq. (22). These similarity values are then sorted, and the $k = \sqrt{N_S}$ most similar superpixels are selected to form the neighbor set \mathcal{N}_i^{t1} . The procedure for time $t2$ is identical and thus omitted for brevity. The edge set \mathcal{E}_S , constructed based on spatial adjacency, is defined as follows:

$$\mathcal{E}_S = \{(\mathbf{V}_i, \mathbf{V}_j) | d_s(\mathbf{V}_i, \mathbf{V}_j) < 2 \cdot S\} \quad (23)$$

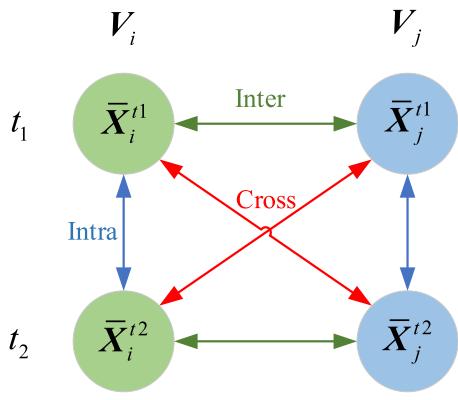


Fig. 3. Feature similarity among Superpixels.

where $d_s(\mathbf{V}_i, \mathbf{V}_j)$ denotes the spatial distance between the geometric centers of superpixels \mathbf{V}_i and \mathbf{V}_j , and S is the superpixel sampling interval.

2.2.2. Association between graph structure and change states

To support graph-based modeling of land cover change, it is necessary to establish the relationship between graph topology and change semantics. We first define the state $O(\mathbf{X}, \mathbf{Y}) \in \{0, 1\}$ between any two superpixels \mathbf{X} and \mathbf{Y} , where 0 indicates they belong to the same land cover type, and 1 indicates different types. After constructing the topological structure graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ of the dual-temporal PolSAR data, for each node \mathbf{V}_i , which contains two superpixels $\bar{\mathbf{X}}_i^{t1}$ and $\bar{\mathbf{X}}_i^{t2}$, the label of the node is denoted as $L_i = O(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t1}) \in \{0, 1\}$, where $L_i = 0$ represents unchange and $L_i = 1$ represents change. Accordingly, the label of any edge $(\mathbf{V}_i, \mathbf{V}_j)$ in the graph can be expressed as (L_i, L_j) .

As shown in Fig. 3, for any two nodes $\mathbf{V}_i = [\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2}]$ and $\mathbf{V}_j = [\bar{\mathbf{X}}_j^{t1}, \bar{\mathbf{X}}_j^{t2}]$, conventional change detection methods typically rely solely on the feature similarity of superpixels from different time phases but at the same spatial location, i.e., intra-node feature similarities $D_{Intra}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2})$ and $D_{Intra}(\bar{\mathbf{X}}_j^{t1}, \bar{\mathbf{X}}_j^{t2})$, to determine the node labels L_i and L_j . Clearly, this approach is highly susceptible to noise and insufficiently exploits the contextual information between nodes. To address this limitation, Sun et al. (2024a) proposed a method for heterogeneous image change detection. This method leverages inter-node feature similarities $D_{Inter}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t1})$ and $D_{Inter}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t2})$ (e.g., similarities between superpixels at same time but distinct spatial locations) to infer node states. Furthermore, it effectively leverages the features of the topologically adjacent nodes for each superpixel node. Building upon these two strategies, this study proposes cross-node temporal feature similarities $D_{cross}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t2})$ and $D_{cross}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t1})$, which can significantly aid in the discrimination of node states. A more detailed analysis and discussion of their contributions will be presented later. It is worth noting that the computations of D_{Intra} , D_{Inter} , and D_{Cross} are all based on Eq. (22).

Fig. 4 illustrates all possible states of four superpixels from any two connected nodes, where the same land cover type is represented using the same shape. For each pair of connected nodes, there are four possible state combinations, and each state is represented in Fig. 4 with the same color. These are analyzed as follows:

State 1#, $(L_i, L_j) = (0, 0)$. As illustrated in Figs. 4(a) and 4(b), there are only two cases in which both nodes remain unchanged. From this observation, the following conclusions can be drawn: (1#a) $D_{Intra}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2})$ and $D_{Intra}(\bar{\mathbf{X}}_j^{t1}, \bar{\mathbf{X}}_j^{t2})$ are expected to be small. (1#b) Since the states of $O_{Inter}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t1})$ and $O_{Inter}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t2})$ are identical, $D_{Inter}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t1})$ and $D_{Inter}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t2})$ will be close in value. (1#c) Similarly, $D_{cross}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t2})$ and $D_{cross}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t1})$ will also be close.

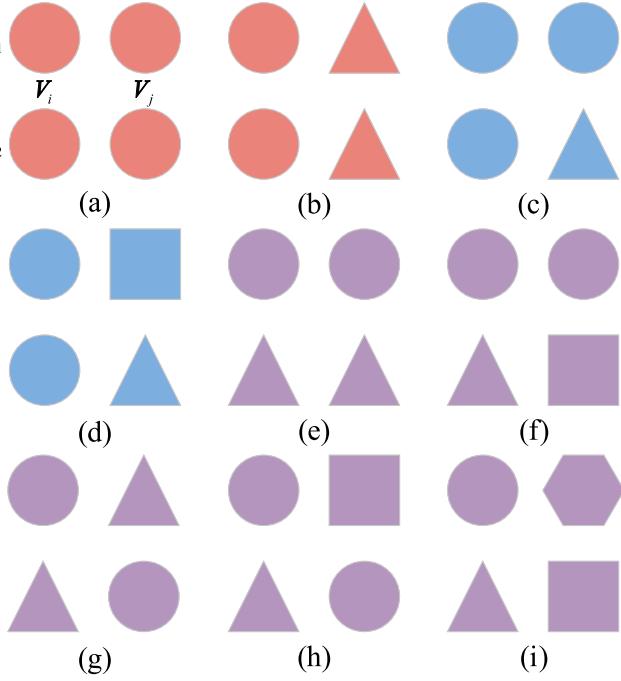


Fig. 4. All possible states of node pairs. (a)–(b) represent $(L_i, L_j) = (0, 0)$; (c)–(d) represent $(L_i, L_j) = (0, 1)$; (e)–(f) represent $(L_i, L_j) = (1, 1)$.

State 2#, $(L_i, L_j) = (0, 1)$. Figs. 4(c) and 4(d) illustrate the case where node i remains unchanged while node j undergoes change. Based on this configuration, the following conclusions can be drawn: **(2#a)** $D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2})$ is expected to be small, whereas $D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2})$ will be relatively large. Although the states of O_{Inter} and O_{Cross} cannot be directly determined, given that $L_i = 0$, it can be inferred that $O_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1})$ and $O_{Cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})$ share the same state. Similarly, $O_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2})$ and $O_{Cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})$ must also share the same state. This leads to: **(2#b)** $D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1})$ and $D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})$ will be close in value. **(2#c)** Likewise, $D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2})$ and $D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})$ will also be similar. Furthermore, considering that $L_j = 1$, it can be deduced that $O_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2})$ and $O_{Cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})$ must differ in state, and the same applies to $O_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1})$ and $O_{Cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})$. As a result, we can derive the following: **(2#d)** There will be a significant difference between $D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1})$ and $D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})$. **(2#e)** Similarly, $D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2})$ and $D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})$ will also differ substantially.

State 3#, $(L_i, L_j) = (1, 0)$. The analysis results are similar to **State 2#** and will not be repeated here.

State 4#, $(L_i, L_j) = (1, 1)$. Since the superpixels within both nodes exhibit different land cover states, the first conclusion is: **(4#a)** Both $D_{Intra}(\bar{X}_i, \bar{Y}_i)$ and $D_{Intra}(\bar{X}_j, \bar{Y}_j)$ will be large. Depending on the number of superpixel pairs with the same land cover type between the two nodes, **State 4#** can be further divided into three sub-cases: “2”, “1”, and “0”, indicating that there are two, one, or no same-type superpixel pairs, respectively. Fig. 4(e) illustrates the case where there are “2” pairs of same-type superpixels between nodes. Figs. 4(f)–(h) show the case where there is “1” such pair. Fig. 4(i) represents the case where there are “0” same-type superpixel pairs between the nodes. It is worth noting that when constructing the edge set \mathcal{E} of the topological graph $G = \{\mathcal{V}, \mathcal{E}\}$, both feature similarity between same-time superpixels and spatial proximity are considered. Given that spatially adjacent superpixels are highly likely to belong to the same land cover type, the scenarios depicted in Figs. 4(g), 4(h), and 4(i) are expected to occur with very low probability. Based on this analysis, an important

Table 1

Summary of change state analyzes and edge cost terms.

ID	Edge cost terms
(1#a)	$\max(D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2}), D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2}))$
(1#b)	$ D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) $
(1#c)	$ D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1}) $
(2#a)	$D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2}) + (1 - D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2}))$
(2#b)	$ D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1}) $
(2#c)	$ D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2}) $
(2#d)	$ D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) $
(2#e)	$ D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) $
(4#a)	$\max(1 - D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2}), 1 - D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2}))$
(4#b)	$1 - \max(D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2}), D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1}))$

conclusion can be drawn: **(4#b)** At least one of $D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})$ or $D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})$ will be large.

2.2.3. Energy model construction and change detection

Based on the topological representation of bi-temporal PolSAR images and the analysis of change states presented in the previous section, the change detection task in PolSAR imagery is reformulated as an energy minimization problem. The corresponding energy function is defined as follows:

$$E(x) = \theta_{const} + \sum_{i \in \mathcal{V}} \theta_i(L_i) + \sum_{(i,j) \in \mathcal{E}} \theta_{ij}(L_i, L_j) \quad (24)$$

Here, θ_{const} is a constant bias term in the energy function, which is set to 0 in this study. $\theta_i(L_i)$ denotes the cost associated with assigning a label L_i to node V_i , while $\theta_{ij}(L_i, L_j)$ represents the cost of assigning labels (L_i, L_j) to a pair of superpixels exhibiting either feature similarity or spatial adjacency. Therefore, the core of the proposed change detection method based on energy minimization lies in the design of the node cost and edge cost functions.

Based on the analysis in Section 2.2.2, the relationships between the change state analyzes and their corresponding edge cost functions are summarized in Table 1. According to this summary, four edge cost functions are constructed to model different types of pairwise relationships. Drawing upon the analysis summarized in Table 1, four edge cost functions are constructed. The cost function θ_{ij} under condition (L_i, L_j) is defined as follows:

$$\begin{aligned} \theta_{ij}(0,0) &= 2 \cdot \max(D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2}), D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2})) \\ &+ |D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2})| \\ &+ |D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})| \end{aligned} \quad (25)$$

$$\begin{aligned} \theta_{ij}(0,1) &= D_{Intra}(\bar{X}_i^{t1}, \bar{X}_i^{t2}) + (1 - D_{Intra}(\bar{X}_j^{t1}, \bar{X}_j^{t2})) \\ &+ 0.5 \cdot |D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})| \\ &+ 0.5 \cdot |D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})| \\ &+ 1 - 0.5 \cdot |D_{Inter}(\bar{X}_i^{t1}, \bar{X}_j^{t1}) - D_{cross}(\bar{X}_i^{t1}, \bar{X}_j^{t2})| \\ &+ 0.5 \cdot |D_{Inter}(\bar{X}_i^{t2}, \bar{X}_j^{t2}) - D_{cross}(\bar{X}_i^{t2}, \bar{X}_j^{t1})| \end{aligned} \quad (26)$$

Table 2
Description of experimental datasets.

No.	Sensor	Time range	Location	Event type	Spatial resolution
#1	ALOS-2 L-band	20150324–20180812	San Francisco, USA	Water-land change	≈ 3 m
#2	ALOS-2 L-band	20150324–20180812	San Francisco, USA	Vegetation change	≈ 3 m
#3	ALOS-2 L-band	20150324–20180812	San Francisco, USA	Water-land change	≈ 3 m
#4	UAVSAR L-band	20090423–20150511	Los Angeles, USA	Building change	≈ 7.5 m
#5	UAVSAR L-band	20090423–20150511	Los Angeles, USA	Building change	≈ 7.5 m

$$\begin{aligned} \theta_{ij}(1,0) = & (1 - D_{Intra}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2})) + D_{Intra}(\bar{\mathbf{X}}_j^{t1}, \bar{\mathbf{X}}_j^{t2}) \\ & + 1 - 0.5 \cdot |D_{Inter}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t1}) - D_{cross}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t1})| \\ & + 0.5 \cdot |D_{Inter}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t2}) - D_{cross}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t2})| \\ & + 0.5 \cdot |D_{Inter}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t1}) - D_{cross}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t2})| \\ & + 0.5 \cdot |D_{Inter}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t2}) - D_{cross}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t1})| \end{aligned} \quad (27)$$

$$\begin{aligned} \theta_{ij}(1,1) = & 2 \cdot \max(1 - D_{Intra}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_i^{t2}), 1 - D_{Intra}(\bar{\mathbf{X}}_j^{t1}, \bar{\mathbf{X}}_j^{t2})) \\ & + 2 \cdot (1 - \max(D_{cross}(\bar{\mathbf{X}}_i^{t1}, \bar{\mathbf{X}}_j^{t2}), D_{cross}(\bar{\mathbf{X}}_i^{t2}, \bar{\mathbf{X}}_j^{t1}))) \end{aligned} \quad (28)$$

Subsequently, considering the fact that in practical change detection scenarios only a small portion of the image typically undergoes change while the majority remains unchanged (Sun et al., 2024a), this study designs a change prior-based sparse penalty. Specifically, the node cost function is defined as:

$$\theta_i(0) = 0, \theta_i(1) = 1 \quad (29)$$

To balance the influence of the node cost function and the edge cost function on the energy optimization process, this study normalizes the four edge cost functions. Specifically, the normalization is performed as follows:

$$\theta_{ij}^{norm}(L_i, L_j) = \alpha_{L_i, L_j} \cdot \frac{\theta_{ij}(L_i, L_j) \cdot N_{SP}}{\sum_{(i,j) \in \mathcal{E}} \theta_{ij}(L_i, L_j)} \quad (30)$$

where α_{L_i, L_j} represents the gain coefficient when the state of the edge is (L_i, L_j) . Through the normalization process, the edge cost and node cost are ensured to be on the same order of magnitude. We further analyze the impact of different settings of α_{L_i, L_j} in detail through ablation experiments presented in Section 3.3.2.

Through the joint modeling of node-level change priors and edge-level topological constraints, the proposed energy model not only enforces label consistency among spatially or temporally correlated regions, but also maintains strong adaptability when dealing with diverse types of land cover. In particular, the incorporation of multiple temporal distance metrics (D_{Inter} , D_{Intra} and D_{cross}) within the edge cost function enables the algorithm to effectively distinguish structural change from radiometric inconsistency. By leveraging the underlying graph topology information, this method enhances the contextual coherence of labeling decisions, thereby suppressing the propagation of noise introduced by temporal radiometric inconsistencies. As a result, subtle yet meaningful changes — especially those with weak scattering signatures — can be more reliably identified. This property is particularly beneficial in natural land cover change scenarios, where conventional methods often struggle to effectively detect low-intensity changes.

Finally, various mature methods have been developed to solve the optimization problem in Eq. (24). In this study, we adopt the Quadratic Pseudo-Boolean Optimization (QPBO) method (Kolmogorov and Rother, 2007), which is based on the max-flow/min-cut algorithm. QPBO method demonstrates excellent time complexity performance in graph optimization tasks, enabling efficient solutions for large-scale problems, and it guarantees a globally optimal result.

3. Experiment and analysis

3.1. Experimental data and evaluation criteria

As shown in Fig. 5, extensive experiments were conducted on five real PolSAR datasets to validate the effectiveness of the proposed method. The detailed information of each dataset is summarized in Table 2. Among them, datasets #1–#3 were acquired by the ALOS-2 L-band system over San Francisco, California, USA, with major land cover changes involving natural variations such as water-land transitions and vegetation changes. Datasets #4 and #5 were acquired by the UAVSAR L-band system over Los Angeles, California, USA, with dominant land cover changes related to urban constructions.

This paper adopts Boundary Recall (BR) and Achievable Segmentation Accuracy (ASA) (Li et al., 2025) to quantitatively evaluate the performance of the superpixel generation algorithm. BR is used to assess the accuracy of boundary detection in the segmentation results, measuring the degree of correspondence between the detected boundaries and the ground truth boundaries. It is calculated as follows:

$$BR = \frac{\sum_{p \in B(GT)} I(\min_{p \in B(SP)} \|p - q\| < \epsilon)}{|B(GT)|} \quad (31)$$

where $B(SP)$ and $B(GT)$ denote the sets of boundary pixels from the superpixel segmentation and ground truth, respectively. $|B(\cdot)|$ represents the number of pixels in the set $B(\cdot)$. When a superpixel boundary pixel lies within a tolerance threshold ϵ (set to 2 in this study) from the ground truth boundary, the indicator function $I(\cdot)$ returns a boolean value of 1. A higher BR indicates better alignment of the detected boundaries with the actual ones.

ASA evaluates the highest possible classification accuracy achievable under ideal conditions, reflecting the extent to which the segmented regions match the ground truth. It is computed as follows:

$$ASA = \frac{\sum_k \max_i |s_k \cap g_i|}{\sum_i |g_i|} \quad (32)$$

where g_i denote the ground truth segmentation, s_k denote the segmentation result produced by the superpixel algorithm, and $|a|$ represent the size (i.e., the number of pixels) of a segmented region. A higher ASA value indicates a closer match between the segmentation result and the ground truth, serving as an indicator of both the quality of the segmentation and its potential for subsequent classification tasks.

The quantitative performance evaluation metrics of the change detection algorithm include Overall Accuracy (OA), F1 Score (F1), and Kappa Coefficient (KC). Their calculation formulas are as follows:

$$OA = \frac{TP}{TP + FP}, \quad (33)$$

$$F1 = \frac{2TP}{2TP + FP + FN}, \quad (34)$$

$$KC = \frac{OA - PRE}{1 - PRE}, \quad (35)$$

$$PRE = \frac{(TP + FN)(TP + FP) + (TN + FP)(TN + FN)}{(TP + TN + FP + FN)^2} \quad (36)$$

where TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives, respectively. In addition, this study also adopts the detection map approach from (Sun et al., 2024a) to evaluate the visual performance of the change detection results.

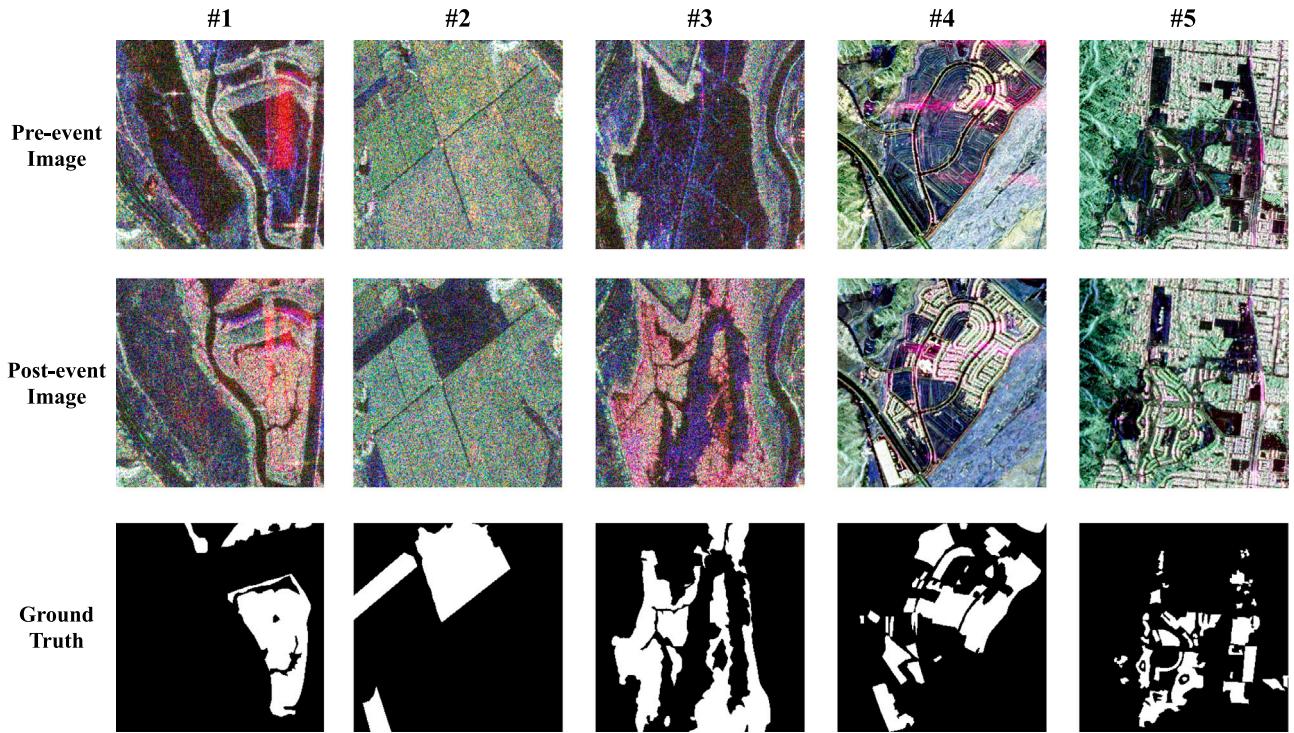


Fig. 5. Experiment datasets.

3.2. Evaluation of superpixel segmentation performance

In this section, the effectiveness of the proposed temporal superpixel generation algorithm is systematically evaluated through extensive experiments, including visual comparisons, quantitative assessments, and computational efficiency analysis.

3.2.1. Visual comparison

To verify the proposed algorithm's superior capability in extracting detailed information, a visual comparison is first conducted with two state-of-the-art (SOTA) multi-temporal superpixel segmentation methods: the HADS method (Li et al., 2023b) and the TSPol-ASLIC method (Gao et al., 2021). Although the HADS method was originally designed for single-temporal data, it is adapted in this study for temporal polarimetric analysis by modifying the joint covariance matrix, as defined in Eq. (5). In the edge detection stage, the bilateral window parameters σ_x , σ_y , θ_f and d_f in the proposed method are set to 2, 2, $\pi/8$, and 1, respectively, while in the TSPol-ASLIC method, they are set to 1.5, 1, $\pi/8$, and 1, respectively. For the parameters in the iterative clustering step, the proposed method sets α and β to 1.5 and 1.0, respectively; the weight parameter m_{DRT} in the HADS method is set to 1.4; and the weight parameter β in the TSPol-ASLIC method is set to 1.

Fig. 6 shows the superpixel segmentation results of the HADS method, the TSPol-ASLIC method, and the proposed method on five bi-temporal PolSAR datasets. In the figure, each pixel is replaced by the average value of its corresponding superpixel. All methods use the same sampling interval $S = 10$. It can be clearly observed that, compared to the two SOTA methods, the proposed method more accurately captures the detailed boundaries of land cover objects, thereby effectively reducing boundary blurring and category confusion. Moreover, by fully leveraging temporal information, the proposed method maintains segmentation consistency across multi-temporal images, which contributes to improved reliability in change detection.

To more systematically evaluate the performance of each method, three typical land cover areas were selected, as indicated by the red boxes A, B, and C in Fig. 6: the vegetation–bare land boundary area

A, the land–water mixed area B, and the urban building area C. The corresponding results of each method are enlarged and shown in Figs. 7–9. Specifically, the first and third rows present the zoomed-in regions corresponding to those in Fig. 6, while the second and fourth rows overlay red lines indicating superpixel boundaries on the Pauli RGB images.

As shown in Fig. 7, both HADS method and the TSPol-ASLIC method fail to preserve key information of narrow strip-shaped roads, resulting in the loss of road details. Furthermore, in the boundary region between vegetation and bare land, both methods suffer from blurred boundaries and land cover confusion. The zoomed-in results of the land–water mixed area in Fig. 8 demonstrate that the proposed method better preserves the continuity and detail of the land–water boundary, effectively maintaining the spatial structure information of water bodies and land areas. Fig. 9 presents segmentation details in the urban building area. It is evident that the proposed method shows significant advantages in this region, providing more precise delineation of building edges compared to the other two methods. It effectively avoids boundary blurring and structural damage caused by narrow roads, resulting in more reliable building segmentation outcomes.

The superior edge adherence performance of the proposed algorithm can be attributed to two main factors. First, compared with the similarity measures used in the baseline methods, the proposed temporal similarity measure is capable of capturing differences in the most fluctuating temporal phase, thereby highlighting the most significant temporal changes. This leads to more effective extraction of land cover boundary information in change regions. For example, the upper-left edge of the road shown in Figure Fig. 9. Second, the edge constraint mechanism introduces a penalty term that effectively prevents pixels from being assigned to superpixel centers across distinct boundaries. As a result, the generated superpixels do not span land cover edges, ensuring the accuracy and integrity of the target boundary structures.

3.2.2. Quantitative assessment

To quantitatively validate the superiority of the proposed algorithm, the sampling interval S was set to a series of different values [8, 10, 12, 15, 18], corresponding to superpixel numbers of [256, 400, 625, 900, 1369]. Three groups of experiments were conducted on five datasets:

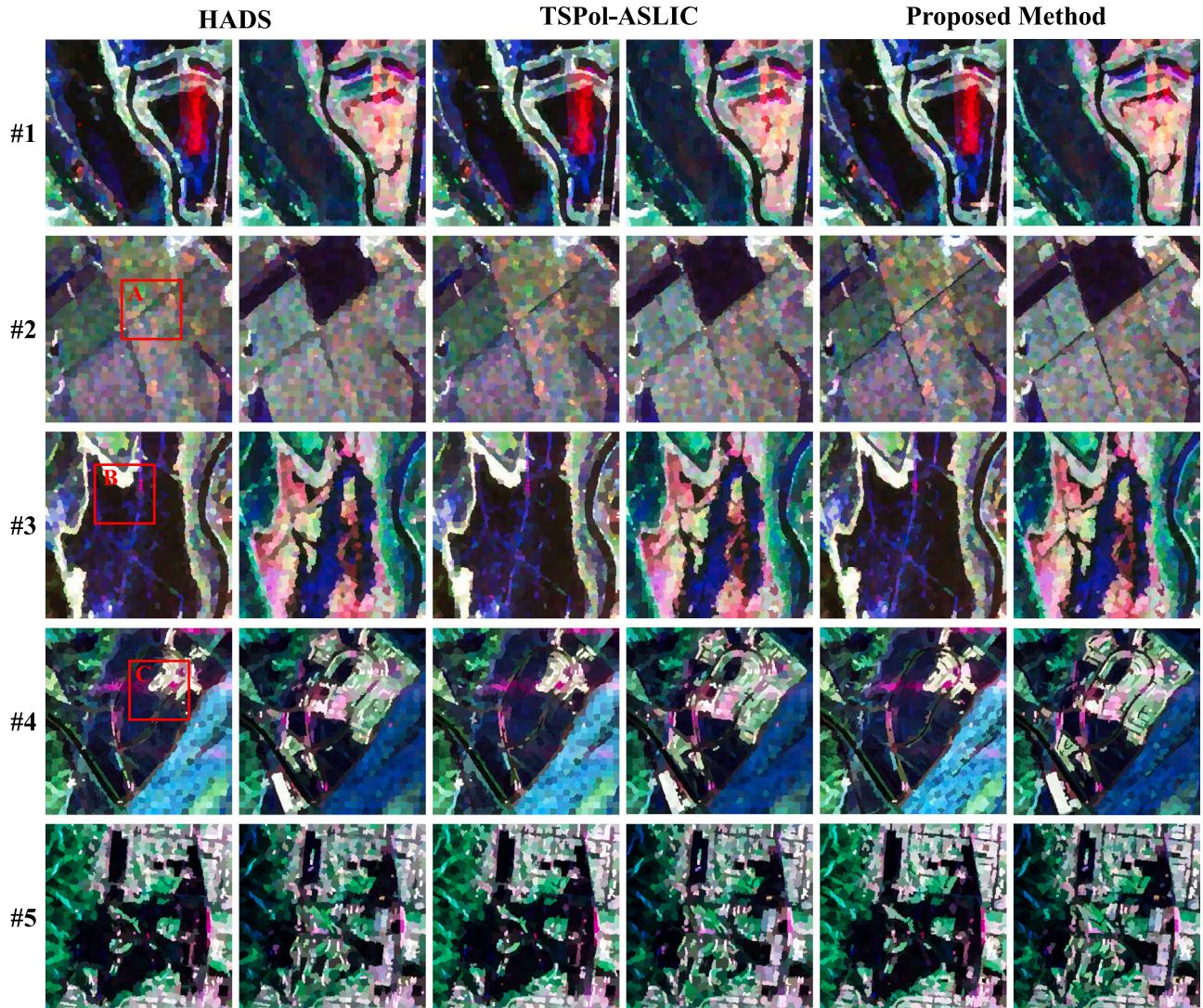


Fig. 6. Bi-temporal superpixel segmentation results on five datasets. In each image, every pixel is replaced by the mean value of its corresponding superpixel. The superpixel sampling interval is set to $S = 10$.

1. A horizontal comparison with HADS Method (Li et al., 2023b) and the TSPol-ASLIC method (Gao et al., 2021).
2. Within the framework of the proposed method, the proposed similarity measure was replaced with the HLT distance (Akbari et al., 2013) (Proposed-HLT), the DRT distance (Bouhlel et al., 2020) (Proposed-DRT), and the SRW distance (Yin et al., 2021) (Proposed-SRW) to analyze the effectiveness of the proposed similarity measure. Although these similarity measures were originally designed for single-temporal PolSAR data, we adapted them to the temporal setting by computing each distance on both temporal images and selecting the maximum value, as defined in Eq. (13).
3. An evaluation of the impact of removing the edge constraint term on superpixel segmentation quality (Proposed-NoEDG).

Fig. 10 presents the BR and ASA results of each method across the five datasets. Overall, the proposed method consistently outperforms the competing approaches under different dataset conditions and varying numbers of superpixels. As the number of superpixels increases, all methods tend to achieve higher BR and ASA scores. However, the performance improvement becomes less significant beyond a certain point. Notably, when the number of superpixels is relatively small,

Table 3

Average BR and ASA improvements of the proposed method compared to other methods.

Method	Superpixels	256	400	625	900	1369	Average
TSPol-ASLIC	BR	10.49%	9.06%	6.04%	4.09%	1.41%	6.22%
	ASA	1.75%	1.57%	1.48%	1.39%	1.13%	1.46%
HADS	BR	12.92%	11.42%	7.12%	5.17%	2.34%	7.79%
	ASA	1.76%	1.84%	1.58%	1.55%	1.31%	1.61%
Proposed-HLT	BR	5.27%	3.92%	3.00%	1.78%	0.56%	2.90%
	ASA	1.06%	0.88%	0.96%	0.95%	0.79%	0.93%
Proposed-SRW	BR	4.56%	4.22%	3.65%	2.06%	0.79%	3.06%
	ASA	1.02%	1.04%	1.21%	1.13%	0.90%	1.06%
Proposed-DRT	BR	5.46%	4.08%	3.07%	2.10%	0.68%	3.08%
	ASA	1.13%	1.00%	1.02%	1.00%	0.84%	1.00%
Proposed-NoEDG	BR	3.15%	3.15%	0.98%	0.85%	0.24%	1.67%
	ASA	0.56%	0.43%	0.42%	0.53%	0.40%	0.47%

the proposed method exhibits more pronounced advantages, delivering substantial gains in both boundary adherence and segmentation accuracy.

To systematically evaluate the performance of each method, Table 3 provides a detailed comparison of the average improvements in BR and ASA metrics for the proposed method relative to the comparison

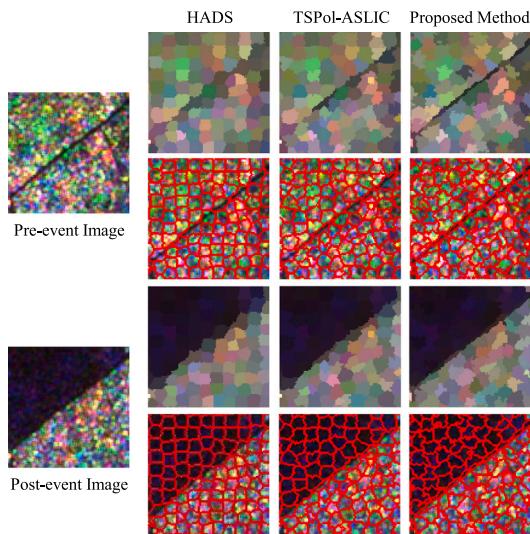


Fig. 7. Magnified result of Region A as shown in Fig. 6.

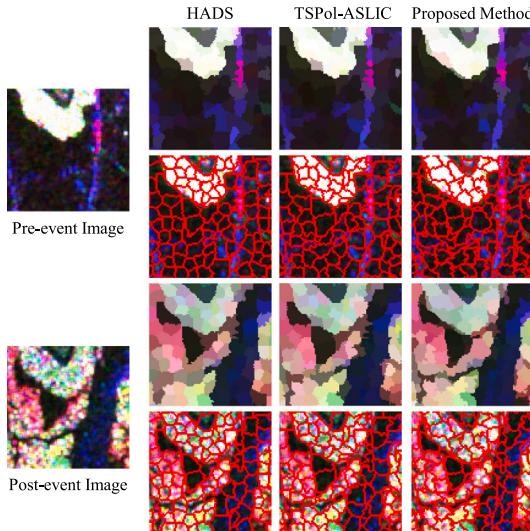


Fig. 8. Magnified result of Region B as shown in Fig. 6.

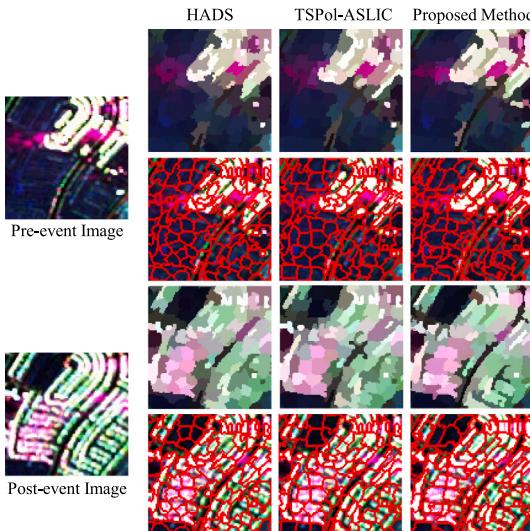


Fig. 9. Magnified result of Region C as shown in Fig. 6.

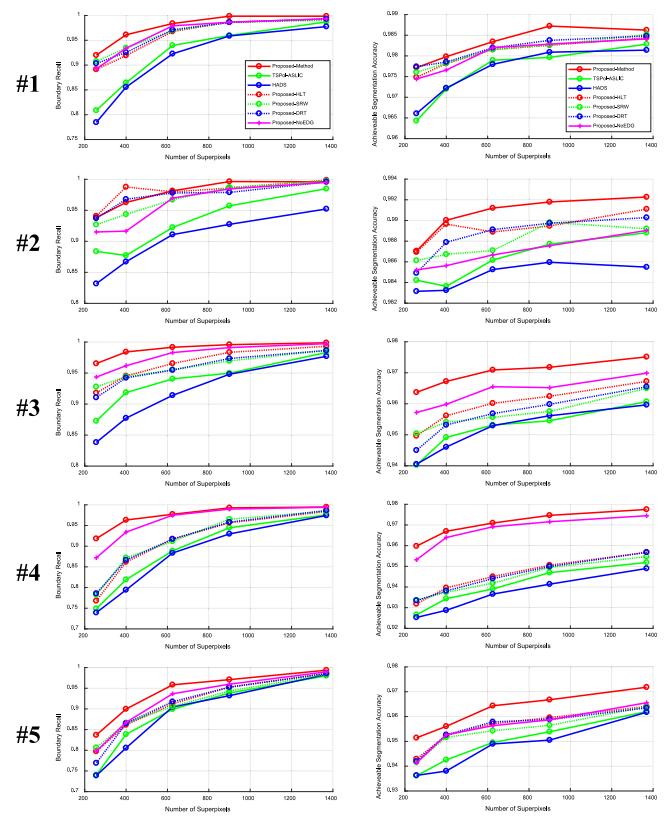


Fig. 10. BR and ASA performance of different methods on five datasets.

methods. Based on the comprehensive analysis, the following important conclusions can be drawn.

First, compared to the HADS method and TSPol-ASLIC method, the proposed method demonstrates significant advantages. In terms of BR improvement, the proposed method shows an average advantage of 7.79% and 6.22% over these two methods, with the most significant advantage observed when the number of superpixels is 256. In terms of ASA improvement, the average advantage is 1.61% and 1.46%, respectively.

Second, by analyzing the comparison results of different distance metrics, the superiority of the proposed temporal similarity measure is verified. The experimental data show that, compared to the Proposed-HLT method, the proposed method achieves an average improvement of 2.90% and 0.93% in BR and ASA, respectively, with the maximum improvements of 5.27% (BR) and 1.06% (ASA) when using 256 superpixels. Compared to the Proposed-SRW method, the proposed method achieves an average improvement of 3.06% (BR) and 1.06% (ASA). Compared to the Proposed-DRT method, the proposed method achieves an average improvement of 3.08% (BR) and 1.00% (ASA), with the maximum improvements of 5.46% (BR) and 1.13% (ASA).

Third, The comparison with the Proposed-NoEDG method strongly validates the effectiveness of the proposed edge constraint term. Experimental data show that after introducing the edge constraint, the proposed method achieves an average improvement of 1.67% and 0.47% in BR and ASA, respectively, with the maximum improvements of 3.15% (BR) and 0.56% (ASA) when using 256 superpixels.

In summary, the proposed method, through its innovative temporal similarity measure and edge constraint mechanism, outperforms existing advanced methods in both BR and ASA metrics.

3.2.3. Computational efficiency analysis

The practicality of the proposed superpixel segmentation method is further verified through a comparison of computational efficiency.

Table 4

Runtime of 5 clustering iterations for each method (Unit: seconds)

Dataset	#1	#2	#3	#4	#5	Avg.
Proposed-Method	13.58	13.42	13.35	13.31	13.30	13.39
TSPol-ASLIC	14.64	14.62	15.08	15.29	14.70	14.87
HADS	13.36	13.09	13.24	13.17	13.10	13.19
Proposed-HLT	16.49	16.12	16.04	16.82	16.08	16.31
Proposed-SRW	16.93	16.75	16.59	16.78	16.60	16.74
Proposed-DRT	14.02	13.76	14.53	13.97	14.28	14.11
Proposed-NoEDG	11.45	11.63	11.69	11.41	11.61	11.56

In the experiment, the main loops of each method were implemented using MATLAB, while the similarity measure calculations were implemented in C language and compiled as library functions using MATLAB's MEX functionality. All experiments were conducted on a desktop computer equipped with an Intel(R) Core(TM) i7-10700K CPU @ 3.80 GHz, 32 GB of RAM, and an NVIDIA RTX 3060 GPU with 12 GB memory. Table 4 shows the time taken to complete 5 clustering iterations for each method when the sampling interval $S = 10$. The experimental results indicate the following: First, the proposed method is more computationally efficient than TSPol-ASLIC method, but slightly less efficient than HADS Method. Second, the edge constraint term introduces an additional computational overhead of about 1.83 s per iteration. Third, the proposed distance measure exhibits higher computational efficiency compared to HLT, DRT, and SRW.

An analysis of the similarity measures used by each method reveals that Proposed-HLT, Proposed-SRW, and TSPol-ASLIC all involve two matrix inversion operations, while HADS and Proposed-DRT involves three determinant computations. In contrast, the proposed method requires only two determinant computations. As a result, the proposed method achieves relatively high computational efficiency.

3.3. Evaluation of change detection performance

3.3.1. Comparison with state-of-the-art methods

To evaluate the effectiveness of the proposed change detection method, we compared it with six SOTA PolSAR image change detection approaches: HLT (Akbari et al., 2013), DRT (Bouhlel et al., 2020), SE (Nascimento et al., 2018), LPEM (Sun et al., 2024a), ITSA (Lv et al., 2023), and CODE (Lin and Lin, 2023). Among them, HLT, DRT, and SE are traditional pixel-level PolSAR change detection methods. LPEM is a method based on superpixels and energy models. ITSA is a CNN-based supervised learning approach that starts with a small number of randomly selected labeled samples and progressively expands the training set using an iterative training sample augmentation strategy. CODE is a semi-supervised method that combines the structure-awareness of graph neural networks with the robustness of convex deep learning, enabling accurate change detection with only a small number of labeled samples.

During the experiments, the parameter α_{L_i, L_j} for the proposed method was configured as [0.7, 1.2, 0.7, 0.7, 0.6] for the five datasets, and the number of superpixels was set to 900. The three traditional methods computed the final change map using the Otsu thresholding method (Otsu et al., 1975) after calculating the difference map. The parameter α^* for the LPEM method was calculated based on the spatial sparsity adaptation provided by its source code, and the parameter β^* was set to 5. The number of superpixels was also set to 900. Moreover, it was observed that using the 9-channel polarization coherence matrix directly as input for LPEM, ITSA, and CODE yielded unsatisfactory results. Therefore, the three components derived from the Pauli decomposition method were used as input for these methods instead. Finally, both ITSA and CODE were implemented using their publicly available source codes with default parameter settings. Fig. 11 displays the change detection results of seven methods on five datasets.

Firstly, it can be observed that the three traditional pixel-level PolSAR image change detection methods suffer significantly from noise

interference, with large areas of FP in datasets #1, #2, #3, and #5 (shown in red in the figure). In contrast, the proposed superpixel-based method and the LPEM method do not exhibit such errors, demonstrating the effectiveness of superpixel preprocessing in suppressing noise in PolSAR image change detection applications. Furthermore, the experimental results show that under the influence of speckle noise in PolSAR images, the traditional change detection methods result in significant FN areas (shown in green in the figure), manifesting as irregular hole structures within the change regions. In many cases, morphological post-processing is applied to these methods in an attempt to fill such holes and improve the continuity of the detection results. However, this type of post-processing introduces several inherent limitations. First, the choice of morphological parameters is typically based on manual experience and lacks adaptability. Second, over-processing may lead to shape distortion and boundary blurring in true change regions. Most importantly, such operations only address superficial artifacts and cannot fundamentally improve robustness to speckle noise. This limitation is particularly evident in the weak scattering change regions of dataset #2, where true changes are often overwhelmed by noise interference. In contrast, our proposed method inherently suppresses noise during the segmentation and optimization stages, without requiring any morphological post-processing, thus maintaining the structural integrity and boundary accuracy of the detected change regions.

Secondly, a comparison with the LPEM method — another superpixel-and energy-based approach — further highlights the advantages of the proposed algorithm in detecting weak scattering changes. The LPEM method constructs its energy function using only local similarity information between adjacent nodes within the same temporal domain. This design limits its ability to capture subtle or gradual changes across time, especially in large-scale weak scattering regions. As observed in the experimental results on dataset #2, LPEM tends to produce extensive false negatives when confronted with broad but low-magnitude change areas. Nevertheless, thanks to its superpixel-based representation, LPEM performs well in maintaining region continuity and preserving boundary integrity. To address its limitations, our method introduces a cross-temporal feature similarity metric, which not only captures the correlations between temporally corresponding nodes more accurately, but also leverages the self-consistency of features within the same node across different time points. This dual-level similarity constraint significantly enhances the sensitivity to weak scattering changes, while retaining the structural advantages of superpixel-based modeling. Experimental results on datasets #2 and #3 — which include extensive gradual change patterns — demonstrate that the proposed method achieves substantially higher detection accuracy compared to LPEM, particularly in weak scattering change scenarios where traditional approaches often fall short.

Thirdly, we analyze the performance of deep learning-based methods in the comparative experiments, focusing on the supervised ITSA method and the semi-supervised CODE method. The results of ITSA indicate that its performance is highly sensitive to the quality and representativeness of the training data. In relatively simple natural land cover change scenarios (datasets #1, #2, and #3), where land cover types are clearly distinguishable and change patterns are relatively stable, representative samples are easier to obtain, enabling the model to achieve satisfactory performance. However, in more complex urban building change scenarios (datasets #4 and #5), ITSA shows a noticeable decline in accuracy, primarily characterized by extensive false positives. This is mainly due to the high heterogeneity and intricate scattering mechanisms associated with urban changes, which are difficult to capture using a limited number of training samples. The CODE method, a recent semi-supervised approach based on graph neural networks, relies heavily on two key factors: the availability of only 0.1% labeled training data and the effectiveness of its superpixel segmentation algorithm. However, the superpixel method used in CODE was not originally developed for PolSAR imagery and therefore struggles to accommodate its unique polarimetric scattering behavior

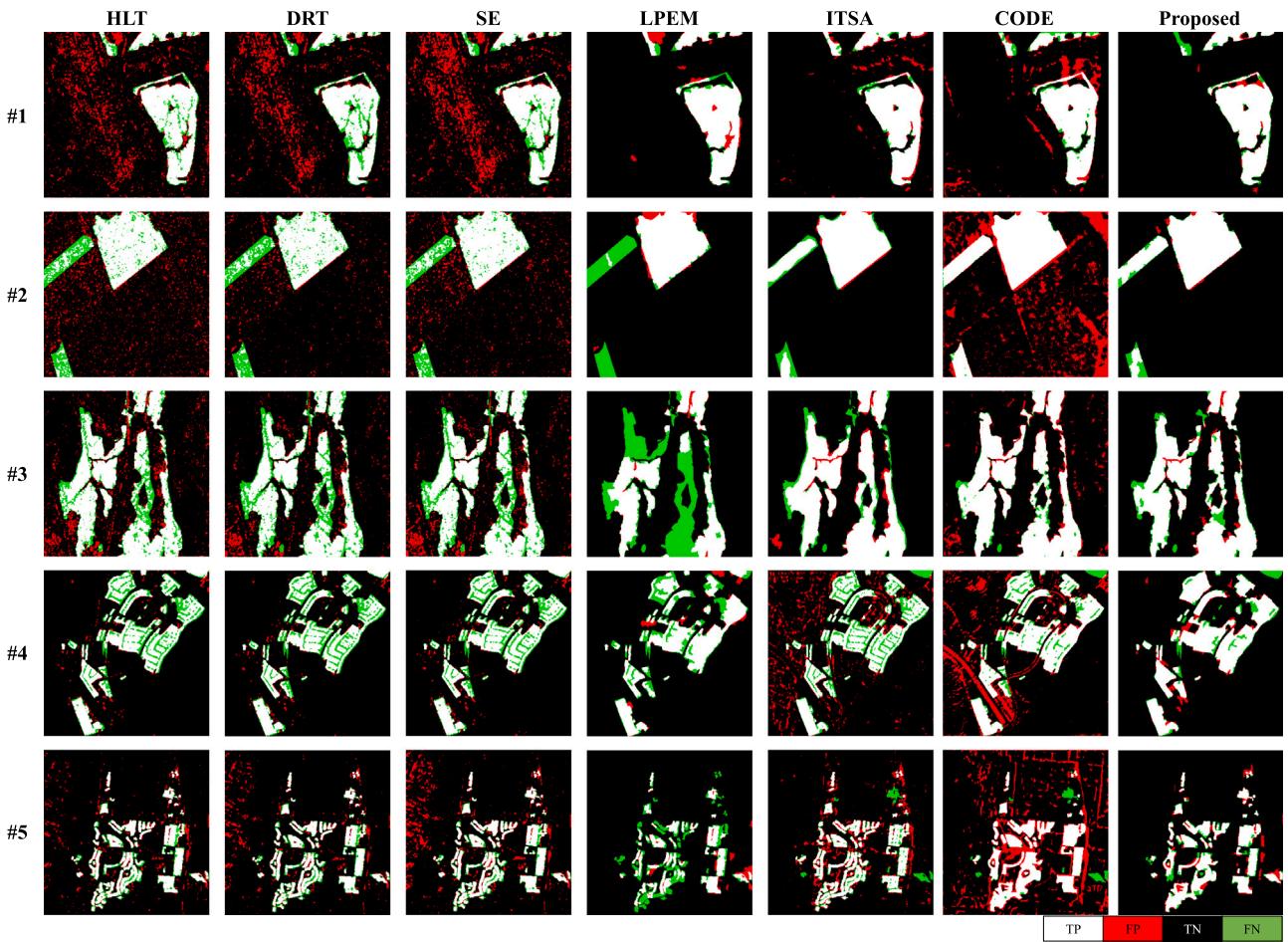


Fig. 11. Change detection results of different methods on five datasets.

and severe speckle noise, resulting in inaccurate region delineation and reduced reliability. For instance, in dataset #2, which involves subtle vegetation changes and is heavily affected by speckle noise, the limited labeled samples are easily contaminated, leading to widespread false detections. In the urban datasets (#4 and #5), false positives frequently occur along road regions, likely due to inaccurate superpixel boundaries and the lack of discriminative polarimetric features between roads and adjacent structures. These observations suggest that deep learning methods originally developed for non-PolSAR data cannot be directly applied to PolSAR change detection without substantial adaptation. In contrast, the unsupervised method proposed in this paper eliminates the need for labeled data and consistently demonstrates stable and robust performance across both natural and urban scenarios. Experimental results confirm that our method significantly reduces false alarms in complex environments while maintaining high detection precision.

Fig. 12 shows the zoomed-in results of representative areas from Fig. 11. Through visual comparison, it becomes more evident that the three traditional methods (HLT, DRT, and SE) are severely affected by speckle noise, leading to noisy and fragmented detection results. The LPEM method suffers from extensive missed detections, especially in regions with weak scattering changes. The two deep learning-based methods (ITSA and CODE) exhibit only moderate performance, largely limited by the quality and representativeness of the training samples. In contrast, the proposed method leverages the edge-constrained temporal superpixel segmentation, which facilitates more accurate region delineation and preserves the structural consistency of land cover boundaries. This contributes to a more robust representation of temporal changes and effectively suppresses the influence of speckle noise

in PolSAR imagery. Compared with LPEM, our approach further constructs a graph-structured model that captures the relationship between superpixel nodes and change semantics from three perspectives: intra-node, inter-node, and cross-node. This comprehensive utilization of topological structure enables more reliable detection of changes, particularly in complex urban scenarios and weak scattering change regions. Due to the scarcity of labeled data, most PolSAR change detection tasks adopt few-shot learning paradigms, where the selection of training samples can significantly affect the results. In contrast, our method is training-free and demonstrates stable performance across diverse datasets with different land cover types, highlighting its generalizability and robustness.

Tables 5–7 present a comprehensive performance analysis of different methods across five datasets, clearly demonstrating that the method proposed in this paper exhibits significant advantages in OA, F1, and KC.

First, as shown in Table 5, the proposed method achieves the highest overall accuracy (OA) on all five datasets. In particular, it attains OA values of 0.9802 and 0.9560 on dataset #2 and #5, respectively—significantly outperforming all other comparison methods. Notably, in the challenging scenarios involving urban building changes (datasets #4 and #5), the performance of the supervised ITSA method drops considerably, with OA values of only 0.8787 and 0.9233. In contrast, the proposed method effectively handles the complexity of such scenes, boosting OA to 0.9507 and 0.9560, respectively. These results clearly demonstrate its robustness and generalization capability in complex change detection tasks.

Second, as shown in Table 6, the proposed method also leads in terms of the F1 score, performing well in both natural land cover

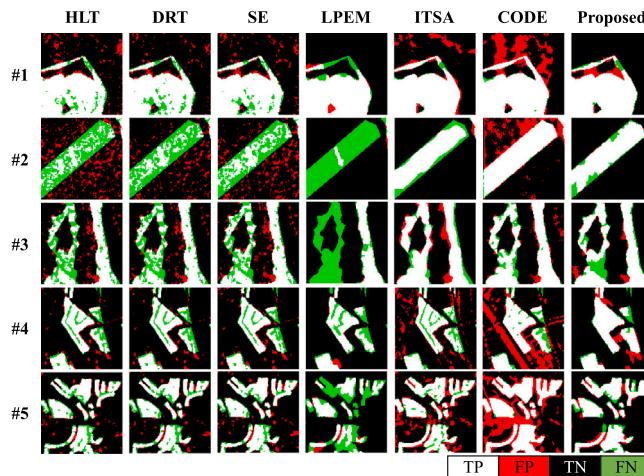


Fig. 12. Enlarged views of typical regions for visual comparison, corresponding to [Fig. 11](#).

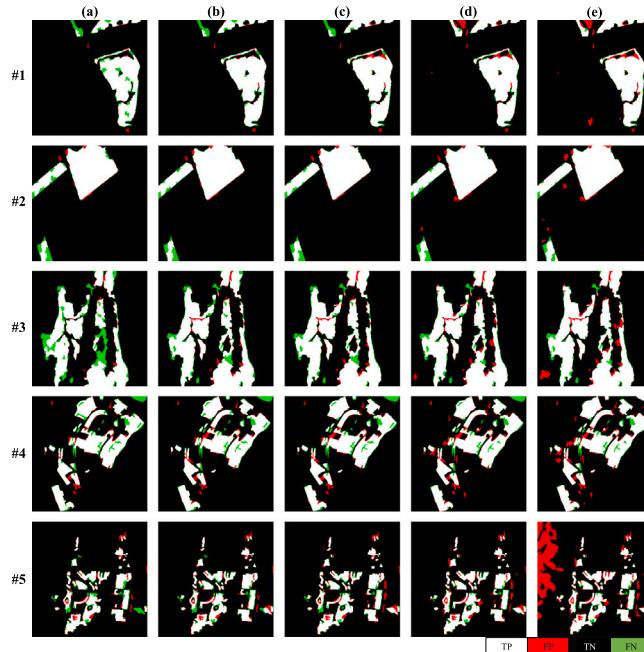


Fig. 13. Ablation results of the gain coefficient α_{L_i,L_j} . (a)–(e) correspond to adjustment factors α of [0.8, 0.9, 1.0, 1.2, 1.5], respectively.

Table 5
OA performance of different methods on five datasets.

Method	dataset#1	dataset#2	dataset#3	dataset#4	dataset#5
HLT	0.8765	0.8783	0.8808	0.9267	0.9391
DRT	0.8792	0.9192	0.8879	0.9237	0.9355
SE	0.8521	0.9083	0.8844	0.9262	0.9214
LPEM	0.9674	0.9376	0.8623	0.9336	0.9310
ITSA	0.9661	0.9734	0.9365	0.8787	0.9233
CODE	0.9398	0.8495	0.9431	0.8754	0.8557
Proposed	0.9768	0.9802	0.9509	0.9507	0.9560

change scenarios (datasets #1–#3) and urban change scenarios (datasets #4–#5). For example, on dataset #2, the proposed method achieves an F1 score of 0.9431, which is 2.07% higher than the second-best method, ITSA (0.9224). On the more complex dataset #4, it achieves 0.8750, outperforming ITSA (0.7058) by nearly 17%, indicating superior preservation of change region structure and boundary accuracy.

Table 6

F1 performance of different methods on five datasets.

Method	dataset#1	dataset#2	dataset#3	dataset#4	dataset#5
HLT	0.6825	0.6893	0.7940	0.7971	0.7668
DRT	0.6712	0.7637	0.7944	0.7848	0.7544
SE	0.6338	0.7512	0.7948	0.7697	0.7246
LPEM	0.8960	0.8041	0.6905	0.8167	0.6687
ITSA	0.8954	0.9224	0.8870	0.7058	0.7235
CODE	0.8230	0.7013	0.9040	0.7286	0.6110
Proposed	0.9235	0.9431	0.9115	0.8750	0.8306

Table 7

KC performance of different methods on five datasets.

Method	dataset#1	dataset#2	dataset#3	dataset#4	dataset#5
HLT	0.6100	0.6143	0.7102	0.7530	0.7319
DRT	0.5996	0.7151	0.7175	0.7395	0.7173
SE	0.5476	0.6951	0.7144	0.7517	0.6795
LPEM	0.8767	0.7676	0.6106	0.7768	0.6321
ITSA	0.8752	0.9064	0.8429	0.6294	0.6794
CODE	0.7872	0.6118	0.8637	0.6493	0.5333
Proposed	0.9098	0.9311	0.8776	0.8443	0.8054

Third, as shown in [Table 7](#), the proposed method also demonstrates excellent performance in terms of KC, especially on datasets #2 and #4, where it achieves 0.9311 and 0.8443, respectively—substantially higher than competing methods. By contrast, although the LPEM method performs well on dataset #1, its KC values drop significantly on datasets #3 and #5 (0.6106 and 0.6321), revealing limited adaptability to diverse scene types.

In conclusion, the method proposed in this paper not only maintains high precision but also effectively addresses issues such as the sensitivity of traditional methods to noise and the dependency of supervised methods on training data. It demonstrates stable and superior performance across diverse change detection tasks.

3.3.2. Ablation study and analysis

The main parameters of the proposed method are the gain coefficient α_{L_i,L_j} used in constructing the cost function and the number of superpixels N_S . Ablation experiments were conducted for these two parameters.

In the ablation experiment for the gain coefficient α_{L_i,L_j} , the number of superpixels N_S was fixed at 900, and different values of α_{L_i,L_j} were configured for the five datasets ($\alpha_{L_i,L_j} = [0.7, 1.2, 0.7, 0.7, 0.6]$). By introducing a tuning coefficient α ($\alpha = [0.8, 0.9, 1.0, 1.2, 1.5]$), the parameter α_{L_i,L_j} for each dataset was weighted, and the corresponding comparison experiment results were obtained, as shown in [Fig. 13](#). The results clearly demonstrate that reducing α_{L_i,L_j} leads to an increase in missed detection (green areas in the figure), while increasing α_{L_i,L_j} leads to an increase in false detection (red areas in the figure). Overall, within a wide range of values for α_{L_i,L_j} , the proposed method achieves good detection results.

[Figs. 15\(a\)–15\(c\)](#) show how the OA, F1, and KC metrics change with the tuning coefficient α . The experimental results indicate that both excessively large and small values of α_{L_i,L_j} cause a decrease in these three metrics. Notably, datasets #1, #2, and #4 maintain relatively stable values for OA, F1, and KC across the entire experimental range. In conclusion, the optimal value of parameter α_{L_i,L_j} for the five datasets lies within a certain magnitude, and results remain stable within a large range. Therefore, parameter α_{L_i,L_j} can be optimized using the following simple iterative strategy. When high false positive rates and low missed detection rates are observed, the value of α_{L_i,L_j} should be reduced appropriately. Conversely, when high missed detection rates and low false positive rates occur, α_{L_i,L_j} should be increased.

In the ablation experiment on the number of superpixels N_S , this study used fixed parameter settings for α_{L_i,L_j} ($\alpha_{L_i,L_j} = [0.7, 1.2, 0.7, 0.7, 0.6]$, corresponding to the five datasets, respectively), while varying the values of N_S ($N_S = [256, 400, 625, 900, 1369]$) to evaluate its

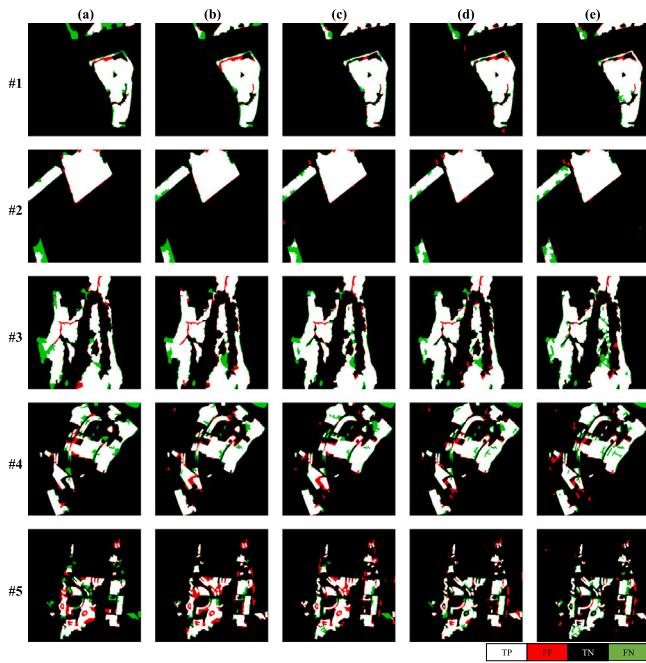


Fig. 14. Ablation results of the number of superpixels N_S . (a)–(e) correspond to superpixel quantities of [256, 400, 625, 900, 1369], respectively.

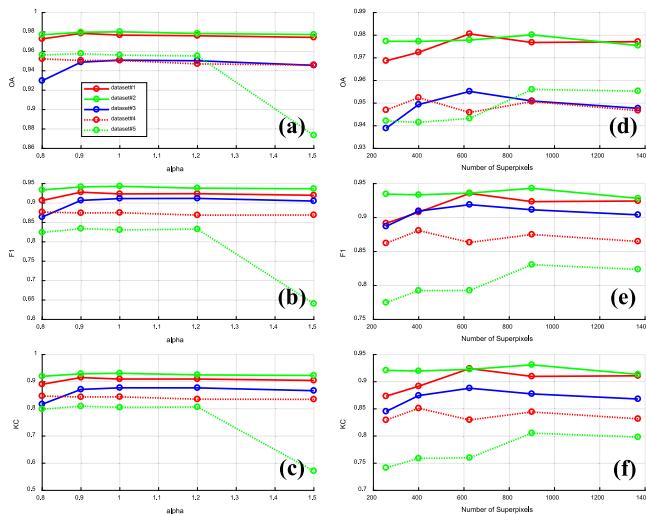


Fig. 15. Variations of OA, F1, and KC metrics in the ablation study.

impact. The experimental results are shown in Fig. 14. Overall, a larger number of superpixels N_S leads to finer over-segmentation of land cover objects, which helps distinguish small targets — such as urban buildings in dataset #5 — from surrounding areas more effectively. However, this also increases the risk of false detection and missed detection caused by image noise. As illustrated in Figs. 15(d)–15(f), which plot the OA, F1, and KC metrics against different values of N_S , the optimal number of superpixels varies with different land cover types.

In general, natural land cover targets tend to have lower signal-to-noise ratios, and a smaller number of superpixels helps suppress the influence of noise more effectively. For instance, in dataset #2, increasing the number of superpixels resulted in more missed detection near the boundaries of the change region in the upper-left corner of the image. In contrast, in urban building regions, a larger number of superpixels enables more detailed delineation of buildings, improving

detection accuracy. The determination of the optimal number of superpixels N_S is closely related to the size of the target land cover objects. Additionally, the experiments revealed that the optimal value of parameter α_{L_i, L_j} also changes when N_S is adjusted, indicating that these two parameters are interdependent and should be jointly optimized for best performance.

4. Conclusion

In this study, we proposed a robust change detection framework for temporal PolSAR images, addressing key challenges such as noise sensitivity, edge misalignment, and missed weak scattering changes. The proposed method integrates edge-constrained temporal superpixels with graph-structured energy optimization, offering the following main contributions. First, a novel temporal Jensen–Bregman LogDet divergence combined with a dynamic edge constraint mechanism was developed to emphasize salient temporal variations while effectively suppressing misclassification across object boundaries. This significantly improves boundary recall and segmentation accuracy compared to existing SOTA approaches. Second, a graph topology model was constructed by combining feature similarity, spatial adjacency, and cross-node temporal correlations. This enhances the detection of weak changes and demonstrates superior performance over traditional pixel-based and supervised learning methods. Extensive experiments on diverse datasets confirm that the proposed method achieves high accuracy and strong robustness across various scenarios. In particular, it shows great potential in complex applications such as urban structure change detection and disaster assessment, providing a practical and effective solution for high-resolution remote sensing-based dynamic monitoring.

CRediT authorship contribution statement

Nengcai Li: Writing – original draft, Methodology, Investigation. **Deliang Xiang:** Writing – review & editing, Supervision, Methodology. **Huaiyue Ding:** Software, Resources, Formal analysis. **Yuzhen Xie:** Validation, Investigation, Formal analysis. **Yi Su:** Visualization, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by [National Natural Science Foundation of China] under Grant ID [62171015].

References

- Akbari, V., Anfinsen, S.N., Doulgeris, A.P., Eltoft, T., 2013. The hotelling-lawley trace statistic for change detection in polarimetric SAR data under the complex wishart distribution. In: 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS. IEEE, pp. 4162–4165.
- Bai, T., Wang, L., Yin, D., Sun, K., Chen, Y., Li, W., Li, D., 2023. Deep learning for change detection in remote sensing: a review. *Geo-Spatial Inf. Sci.* 26 (3), 262–288.
- Bao, J., Yin, J., Yang, J., 2017. Superpixel-based segmentation for multi-temporal polsars images. In: 2017 Progress in Electromagnetics Research Symposium-Fall. PIERS-FALL, IEEE, pp. 654–658.
- Barcelos, I.B., Belém, F.D.C., João, L.D.M., Patrocínio, Jr., Z.K.D., Falcão, A.X., Guimarães, S.J.F., 2024. A comprehensive review and new taxonomy on superpixel segmentation. *ACM Comput. Surv.* 56 (8), 1–39.
- Bouhlel, N., Akbari, V., Méric, S., 2020. Change detection in multilook polarimetric SAR imagery with determinant ratio test statistic. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15.
- Cheng, G., Huang, Y., Li, X., Lyu, S., Xu, Z., Zhao, H., Zhao, Q., Xiang, S., 2024. Change detection methods for remote sensing in the last decade: A comprehensive review. *Remote. Sens.* 16 (13), 2355.

- Cherian, A., Sra, S., Banerjee, A., Papanikolopoulos, N., 2012. Jensen-bregman logdet divergence with application to efficient similarity search for covariance matrices. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (9), 2161–2174.
- Conradsen, K., Nielsen, A.A., Schou, J., Skriver, H., 2003. A test statistic in the complex wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* 41 (1), 4–19.
- Cour, T., Benezit, F., Shi, J., 2005. Spectral segmentation with multiscale graph decomposition. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2. CVPR'05, IEEE, pp. 1124–1131.
- Deng, J., Wang, W., Zhang, H., Zhang, T., Zhang, J., 2024. Polsar ship detection based on superpixel-level contrast enhancement. *IEEE Geosci. Remote. Sens. Lett.*
- Erten, E., Chesnokova, O., Rossi, C., Hajnsek, I., 2011. A polarimetric temporal scene parameter and its application to change detection. In: 2011 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 1091–1094.
- Faroog, B., Manocha, A., 2024. Satellite-based change detection in multi-objective scenarios: A comprehensive review. *Remote. Sens. Appl.: Soc. Environ.* 34, 101168.
- Feng, J., Cao, Z., Pi, Y., 2014. Polarimetric contextual classification of polsar images using sparse representation and superpixels. *Remote. Sens.* 6 (8), 7158–7181.
- Gao, H., Wang, C., Xiang, D., Ye, J., Wang, G., 2021. Tspol-ASLIC: Adaptive superpixel generation with local iterative clustering for time-series quad-and dual-polarization SAR data. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15.
- Gong, M., Zhan, T., Zhang, P., Miao, Q., 2017. Superpixel-based difference representation learning for change detection in multispectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 55 (5), 2658–2673.
- Gueguen, L., Soille, P., Pesaresi, M., 2011. Change detection based on information measure. *IEEE Trans. Geosci. Remote Sens.* 49 (11), 4503–4515.
- Guo, Z., Zhang, H., Ge, J., Shi, Z., Xu, L., Tang, Y., Wu, F., Wang, Y., Wang, C., 2024. Built-up area extraction in polsar imagery using real-complex polarimetric features and feature fusion classification network. *Int. J. Appl. Earth Obs. Geoinf.* 134, 104144.
- Han, C., Wu, C., Hu, M., Li, J., Chen, H., 2024. C2F-SemiCD: A coarse-to-fine semi-supervised change detection method based on consistency regularization in high-resolution remote-sensing images. *IEEE Trans. Geosci. Remote Sens.*
- Kolmogorov, V., Rother, C., 2007. Minimizing nonsubmodular functions with graph cuts—a review. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (7), 1274–1279.
- Lei, Y., Liu, X., Shi, J., Lei, C., Wang, J., 2019. Multiscale superpixel segmentation with deep features for change detection. *Ieee Access* 7, 36600–36616.
- Li, N., Hu, C., Wang, W., Quan, S., Xiang, D., 2023a. Polarimetric SAR target decomposition method based on independent polarization orientation angle integration. *Acta Geod. et Cartogr. Sin.* 52 (12), 2141–2153.
- Li, N., Xiang, D., Sun, X., Hu, C., Su, Y., 2025. Multiscale adaptive PolSAR image superpixel generation based on local iterative clustering and polarimetric scattering features. *ISPRS J. Photogramm. Remote Sens.* 220, 307–322.
- Li, M., Zou, H., Dong, Z., Qin, X., Liu, S., Zhang, Y., 2024a. Unsupervised semantic segmentation of polsar images based on multiview similarity. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 17, 5317–5331.
- Li, M., Zou, H., Qin, X., Dong, Z., Sun, L., Wei, J., 2023b. Superpixel generation for polarimetric SAR images with adaptive size estimation and determinant ratio test distance. *Remote. Sens.* 15 (4), 1123.
- Li, H., Zou, B., Zhang, L., Qin, J., 2024b. CausalCD: A causal graph contrastive learning framework for self-supervised SAR image change detection. *IEEE Trans. Geosci. Remote Sens.*
- Li, H., Zou, B., Zhang, L., Qin, J., 2024c. Semi-supervised SAR image change detection via structure-optimized complex-valued graph contrastive learning. *IEEE Geosci. Remote. Sens. Lett.*
- Lin, T.-H., Lin, C.-H., 2023. Hyperspectral change detection using semi-supervised graph neural network and convex deep learning. *IEEE Trans. Geosci. Remote Sens.* 61, 1–18.
- Lv, Z., Huang, H., Sun, W., Jia, M., Benediktsson, J.A., Chen, F., 2023. Iterative training sample augmentation for enhancing land cover change detection performance with deep learning neural network. *IEEE Trans. Neural Netw. Learn. Syst.*
- Nascimento, A.D., Frery, A.C., Cintra, R.J., 2018. Detecting changes in fully polarimetric SAR imagery with statistical information theory. *IEEE Trans. Geosci. Remote Sens.* 57 (3), 1380–1392.
- Otsu, N., et al., 1975. A threshold selection method from gray-level histograms. *Automatica* 11 (285–296), 23–27.
- Qin, F., Guo, J., Lang, F., 2014. Superpixel segmentation for polarimetric SAR imagery using local iterative clustering. *IEEE Geosci. Remote. Sens. Lett.* 12 (1), 13–17.
- Qin, X., Zhang, Y., Li, Y., Cheng, Y., Yu, W., Wang, P., Zou, H., 2022. Distance measures of polarimetric SAR image data: A survey. *Remote. Sens.* 14 (22), 5873.
- Shi, J., He, T., Ji, S., Nie, M., Jin, H., 2023. Cnn-improved superpixel-to-pixel fuzzy graph convolution network for polsar image classification. *IEEE Trans. Geosci. Remote Sens.* 61, 1–18.
- Song, H., Yang, W., Bai, Y., Xu, X., 2015. Unsupervised classification of polarimetric SAR imagery using large-scale spectral clustering with spatial constraints. *Int. J. Remote Sens.* 36 (11), 2816–2830.
- Su, Z., Zhang, J., Wang, L., Zhang, H., Liu, Z., Pietikäinen, M., Liu, L., 2023. Lightweight pixel difference networks for efficient visual representation learning. *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (12), 14956–14974.
- Sun, Y., Lei, L., Guan, D., Kuang, G., Li, Z., Liu, L., 2024a. Locality preservation for unsupervised multimodal change detection in remote sensing imagery. *IEEE Trans. Neural Netw. Learn. Syst.*
- Sun, X., Li, X., Xiang, D., Hu, C., 2024b. SAR vehicle image generation with integrated deep imaging geometric information. *Int. J. Appl. Earth Obs. Geoinf.* 132, 104028.
- Wang, R., Nie, Y., Geng, J., 2024a. Multiscale superpixel-guided weighted graph convolutional network for polarimetric SAR image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 17, 3727–3741.
- Wang, L., Zhang, M., Gao, X., Shi, W., 2024b. Advances and challenges in deep learning-based change detection for remote sensing images: A review through various learning paradigms. *Remote. Sens.* 16 (5), 804.
- Wu, S., Wang, W., Deng, J., Quan, S., Ruan, F., Guo, P., Fan, H., 2024. Nearshore ship detection in PolSAR images by integrating superpixel-level GP-PNF and refined polarimetric decomposition. *Remote. Sens.* 16 (6), 1095.
- Xiang, D., Ban, Y., Wang, W., Su, Y., 2017. Adaptive superpixel generation for polarimetric SAR images with local iterative clustering and SIRV model. *IEEE Trans. Geosci. Remote Sens.* 55 (6), 3115–3131.
- Xiang, D., Ding, H., Sun, X., Cheng, J., Hu, C., Su, Y., 2024. PolSAR image registration combining siamese multiscale attention network and joint filter. *IEEE Trans. Geosci. Remote Sens.*
- Xiang, D., Xie, Y., Cheng, J., Xu, Y., Zhang, H., Zheng, Y., 2022. Optical and SAR image registration based on feature decoupling network. *IEEE Trans. Geosci. Remote Sens.* 60, 1–13.
- Xie, J., Gao, F., Zhou, X., Dong, J., 2024. Wavelet-based bi-dimensional aggregation network for SAR image change detection. *IEEE Geosci. Remote. Sens. Lett.*
- Xie, L., Zhang, H., Wang, C., Liu, M., Zhang, B., 2015. Superpixel-based PolSAR images change detection. In: 2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar. APSAR, IEEE, pp. 792–796.
- Yang, W., Yang, X., Yan, T., Song, H., Xia, G.-S., 2016. Region-based change detection for polarimetric SAR images using wishart mixture models. *IEEE Trans. Geosci. Remote Sens.* 54 (11), 6746–6756.
- Yin, J., Wang, T., Du, Y., Liu, X., Zhou, L., Yang, J., 2021. SLIC superpixel segmentation for polarimetric SAR images. *IEEE Trans. Geosci. Remote Sens.* 60, 1–17.
- Zhang, T., Du, Y., Yang, Z., Quan, S., Liu, T., Xue, F., Chen, Z., Yang, J., 2021. PolSAR ship detection using the superpixel-based neighborhood polarimetric covariance matrices. *IEEE Geosci. Remote. Sens. Lett.* 19, 1–5.
- Zhang, X., Yin, J., Yang, J., Guo, X., 2022. Region-based change detection for polarimetric SAR images. In: IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 675–678.
- Zhao, M., Hu, X., Zhang, L., Meng, Q., Chen, Y., Bruzzone, L., 2024. Beyond pixel-level annotation: Exploring self-supervised learning for change detection with image-level supervision. *IEEE Trans. Geosci. Remote Sens.*
- Zhou, N., Zhou, M., Sui, H., 2025. DepthCD: Depth prompting in 2D remote sensing imagery change detection. *ISPRS J. Photogramm. Remote Sens.* 227, 145–169.