remote sensing

MDPI

Article

Spatiotemporal Evolution Monitoring of Small Water Body Coverage Associated with Land Subsidence Using SAR Data: A Case Study in Geleshan, Chongqing, China

Tianhao Jiang [1], Faming Gong [2], Qiankun Kong [1] and Kui Zhang [1,*]

[1] School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China; 202412021010@stu.cqu.edu.cn (T.J.); 20241201030g@stu.cqu.edu.cn (Q.K.)
[2] Wuxi County Digitized City Operation and Governance Center, Chongqing 405800, China; 06639198@alu.cqu.edu.cn
* Correspondence: zk@cqu.edu.cn

Highlights

What are the main findings?

- Proposed is a two-stage clustering algorithm for dual-pol SAR water mapping: temporal-feature-based initial segmentation and cluster refinement with adaptive cluster counts.
- Applied to long-term monitoring (7 February 2017–24 August 2025) of reservoir coverage evolution, the algorithm shows that small reservoirs had cumulative desiccation of up to 24 months from land subsidence caused by tunnel excavation.

What are the implications of the main findings?

- Stage one: The K-S test is used to characterize temporal features, generating water candidate regions and mitigating shadow misclassification. Stage two: SVD is used to fuse dual-pol features into a high-discrimination water-non-water set. Both steps enhance water mapping accuracy.
- The proposed method accommodates water drying-flooding regimes and enables spatiotemporal evolution monitoring of water body coverage.

Abstract

Monitoring small water body coverage spatiotemporal evolution in karst areas of complex hydrogeology is pivotal for water resource management and disaster assessment. With recent infrastructure expansion, intensive tunnel excavation has occurred in Chongqing's Geleshan, a typical karst region with fragile aquifers. It has disrupted hydrogeological systems, triggering ground subsidence, groundwater leakage, and subsequent reservoir desiccation, as well as threatening regional water security and ecology. Thus, monitoring reservoir coverage evolution is critical to clarify dynamics and driving mechanisms. Synthetic Aperture Radar (SAR) is ideal for water body mapping, enabling data acquisition independent of illumination and weather. However, traditional SAR-based water extraction methods are hampered by low-scatter noise and poor adaptability to hydrological fluctuations. To address this, a two-stage dual-polarization SAR clustering algorithm (TSDPS-Clus) was developed using 452 time-series Sentinel-1 images (7 February 2017–24 August 2025). Specifically, the Kolmogorov–Smirnov test via pixel-wise time-series statistics screened core water areas, built candidate regions, and mitigated noise. Subsequently, dual-polarization and positional features were fused via singular value decomposition (SVD) to generate a high-discrimination low-dimensional feature set, followed by the Iterative Self-Organizing

Check for updates

Academic Editor: Assefa M. Melesse

Received: 26 January 2026
Revised: 14 February 2026
Accepted: 17 February 2026
Published: 19 February 2026

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

Data Analysis Techniques Algorithm (ISODATA) clustering for high-precision extraction. Results demonstrate that the algorithm suits reservoir storage-desiccation dynamics; dual-polarization complementarity boosts accuracy and clarifies six reservoirs' spatiotemporal evolution. Notably, post-2023, tunnel excavation-induced land subsidence increased drying frequency and duration, with a 24-month maximum cumulative desiccation period.

Keywords: water-body segmentation; water area; synthetic aperture radar (SAR); dual-polarization; data fusion; land subsidence

1. Introduction

Karst areas are fragile ecosystems and geological hazard-sensitive zones. Their surface and groundwater spatiotemporal evolution processes directly relate to ecosystem stability and geological hazard risks [1,2]. In particular, regions with intensive human engineering activities (e.g., tunnel construction, mineral mining) face groundwater drawdown risks. This easily triggers chain disasters, such as karst fracture expansion and land subsidence. Abnormal water body distribution, including reservoir drying and spring cutoff, acts as a core precursor signal for such hazards [3,4]. Therefore, long-term, high-precision monitoring of karst water bodies is essential. It holds crucial scientific value and practical significance for deciphering disaster evolution mechanisms and establishing early warning systems.

Remote sensing technology has become an indispensable tool for water body dynamic monitoring, thanks to its high spatiotemporal resolution and wide coverage [5]. Traditional optical remote sensing, such as multispectral and hyperspectral remote sensing, is easily constrained by external factors. Such factors include cloud cover, rainfall, illumination, and vegetation obstruction, resulting in obvious limitations for water body monitoring in mountainous areas with frequent cloud and rain and high vegetation coverage [6]. In contrast, synthetic aperture radar (SAR) relies on an active microwave imaging mechanism. It not only possesses core advantages of all-weather, day-and-night observation and strong penetration, which effectively overcome the application bottlenecks of optical remote sensing, but also leverages the weak backscattering characteristics of water surfaces [7,8]. Thus, it has become the preferred technical approach for high-precision water body monitoring in complex mountainous areas with frequent clouds and rain [9,10].

Over the past few decades, various SAR data sources have been widely applied in water body mapping. These include ENVISAT-ASAR [11–14], RADARSAT/RADARSAT-2 [15–18], JERS-1/ALOS-PALSAR/ALOS2 [19–22], and China's GF-3 satellite data [23,24]. Since the successive launches of Sentinel-1A/B satellites in 2014, they have gradually become the dominant SAR data source for water body mapping. Their core advantages lie in a short revisit cycle of 5–7 days, 10 m spatial resolution, and free open access [25–28]. For water body extraction from Sentinel-1 data, existing methods can be roughly categorized into two types: deep learning and traditional methods. Since Geoffrey Hinton proposed the concept of deep learning in 2006 [29], this technology has attracted widespread attention. Among its branches, deep convolutional neural networks show significant advantages in image semantic segmentation tasks [30,31]. For instance, Fernando used the U-Net neural network on Sentinel-1 SAR images to successfully extract floods in the Los Ríos region of Mexico [32]. However, the application of deep learning in SAR water body mapping remains limited. The core bottleneck is the scarcity of high-quality SAR image datasets, which hinders the full exertion of its feature capture potential. Additionally, deep learning relies on manually labeled training samples. This process is not only time-consuming and

labor-intensive but also subject to interference from subjective labeling. Given this, this study focuses on traditional methods to achieve unsupervised and automated water body extraction based on Sentinel-1 SAR data.

In SAR images, water bodies often exhibit low grayscale values and more uniform textures due to their specular reflection property. This feature facilitates accurate water body extraction in simple scenarios. For example, Jia et al. proposed a SAR dual-polarization water index (SDWI). It enhances water body features, suppresses soil and vegetation signals, and successfully extracts water information of Poyang Lake based on thresholding [33]. Additionally, Du et al. used the SDWI to extract large-scale water bodies in the permafrost regions of the Qinghai–Tibet Plateau [34]. The results outperformed those obtained from Vertical Transmit-Vertical Receive (VV) or Vertical Transmit-Horizontal Receive (VH) images alone. This further confirms that dual-polarization SAR data, with its complementary advantages, can effectively improve the accuracy of water body extraction.

Currently, SAR-based water body extraction studies focus more on large lakes [35,36], rivers [37], and other large water bodies. Research on small water bodies, such as ponds, small lakes, and fishponds, however, remains insufficient. For instance, Sun et al. incorporated SAR images with Digital Elevation Models for water body extraction [38]. This method facilitates the accurate capture of small and slender water bodies. Bao et al. proposed a dual-threshold graph-cut model. Utilizing Gabor filters, they generated multi-scale texture feature maps, enabling the model to handle both large and small water body extraction [39]. Small water body extraction is particularly challenging. Small water bodies have a small surface area and appear dark in images, easily being confused with surrounding building shadows [40]. Severe speckle noise further hinders accurate extraction [41]. Furthermore, small water bodies are prone to dynamic drying-impoundment alternations, driven by seasonal fluctuations and engineering impacts. Most traditional methods map water bodies based on the assumption of their existence and thus cannot satisfy the accuracy and stability requirements for long-term monitoring. Overall, identifying small water bodies in complex environments poses significant challenges.

Geleshan Mountain, located in the Shapingba District of Chongqing, is a typical high-altitude karst-sensitive area. Intensive tunnel construction in recent years has caused significant groundwater drawdown and karst land subsidence here [42]. Notably, the reservoirs in this region have exhibited abnormal water body evolution, providing an ideal research carrier to verify the effectiveness of water body area monitoring methods.

Accordingly, a two-stage clustering algorithm for water body extraction based on dual-polarization SAR images (TSDPS-Clus) is proposed in this study. Stage 1: Temporal water index maps are calculated, and the Kolmogorov–Smirnov (K-S) test is then applied based on their temporal statistical characteristics to construct water body candidate regions. Stage 2: Two steps are involved, where firstly, Singular Value Decomposition (SVD) is employed to fuse dual-polarization features with spatial location information, yielding a low-dimensional feature set with high discriminability, and secondly, Iterative Selforganizing Data Analysis Techniques Algorithm (ISODATA) clustering is applied to the feature set constrained by water body candidate regions to obtain accurate water body segmentation results. Overall, TSDPS-Clus not only fully leverages temporal information and dual-polarization features but also adapts to dynamic scenarios such as alternating water body drying and impoundment, thereby meeting the accuracy and stability requirements for long-term monitoring.

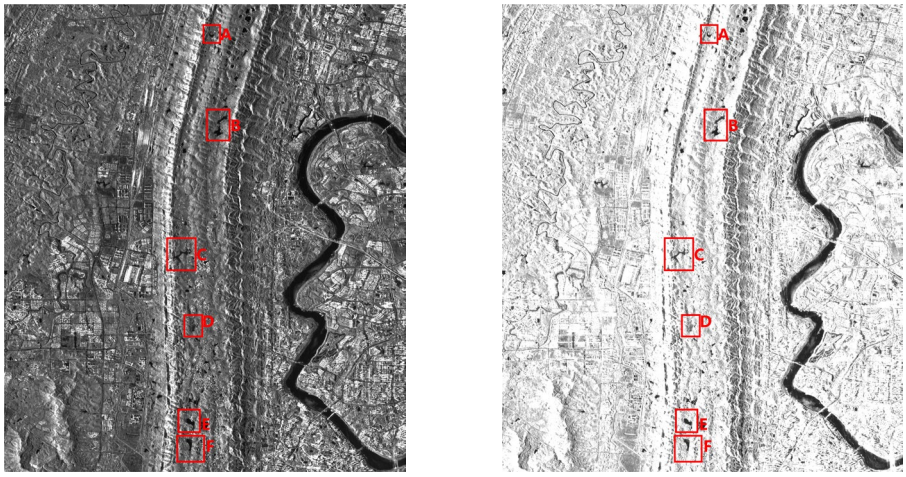
The remainder of this paper is structured as follows: Section 2 presents the study area and dataset; Section 3 elaborates on the relevant theories of the proposed method; Section 4 showcases the experimental results; Section 5 discusses and analyzes these results; and finally, Section 6 summarizes the conclusions drawn from this study.

2. Study Area and Data

2.1. Study Area

This study focuses on selected reservoirs in Geleshan Mountain, Shapingba District, Chongqing. Specifically, the study area is situated between 106°23′52″ and 106°26′42″E longitude and 29°29′30″ and 29°44′50″N latitude. As a branch of the Huaying Mountain Range, it lies in the Guanyin Gorge Anticline of the Eastern Sichuan Fold Belt and exhibits a low-mountain terrain with a total area of approximately 68.2 km^2. Notably, characterized by the "one mountain, two troughs, and three ridges" as the core landform, the region has well-developed karst troughs and is categorized as a high-altitude karst sensitive zone. Additionally, adjacent to the Jialing River, it features developed karst fissures, abundant surface and groundwater resources, and an integrated hydrological system. In recent years, intensive tunnel construction has caused significant groundwater depletion, thereby triggering numerous karst ground collapses [42,43]. Core manifestations of these collapses include abnormal water body distribution (e.g., variations in depression water accumulation, anomalous spring activities). Consequently, monitoring water body evolution in this region enables the capture of hydrological responses, which in turn provide a basis for analyzing the collapse process and formulating early warning, prevention, and control strategies for geological hazards in similar areas.

To accurately monitor the process of water body evolution, this study selects six reservoirs within the region as core experimental areas to conduct a dynamic analysis of water body evolution over the past eight years (7 February 2017–24 August 2025). The specific distribution of the experimental areas is shown in Figure 1 and Table 1.



(**a**) (**b**)

Figure 1. Average SAR magnitude. (**a**) VH image; (**b**) VV image. (A–F) represent the regions of interest, which are six reservoirs.

Table 1. Basic information about the study area.

NO.	Name	Size (Pixel)	Polarization	Longitude	Latitude	Numbers of SAR
Data A	Houfu Lake	150 × 140	VV + VH	106.41	29.74	452
Data B	Feiran Lake	260 × 190	VV + VH	106.42	29.70	452
Data C	Canruo Lake	270 × 240	VV + VH	106.40	29.65	452
Data D	Yujiawan Reservoir	180 × 150	VV + VH	106.41	29.62	452
Data E	Shangtianchi Reservoir	240 × 180	VV + VH	106.40	29.58	452
Data F	Xiatianchi Reservoir	268 × 226	VV + VH	106.40	29.57	452

2.2. Dataset Description

The Geleshan area in Chongqing is prone to dense and persistent cloud cover in winter, leading to a paucity of high-quality optical remote sensing imagery for water bodies to capture fine-scale subtle changes, which cannot meet the requirements for long-term time-series monitoring of water body spatiotemporal evolution. In contrast, Sentinel-1 provides C-band high spatiotemporal resolution products and global free data access with stable acquisition, featuring a revisit period of 12 days and a spatial resolution of approximately 5.5 × 5.5 m. Thus, to achieve the research goals, 452 high-quality SAR images acquired by the Sentinel-1A satellite are employed as the core input dataset after standard preprocessing. Notably, all these images effectively fully cover the Geleshan study area in Chongqing and were systematically and consistently acquired in the VV + VH dual-polarization imaging mode. Specifically, the imaging period ranges from 7 February 2017 to 24 August 2025, forming an approximately 8-year reliable continuous high-temporal-resolution time-series dataset. Detailed parameters are provided in Table 2.

Table 2. List of SAR datasets used in this work.

NO.	Date	NO.	Date	NO.	Date	NO.	Date	Sensor Mode	Polarization
1	7 February 2017	58	9 February 2019	115	24 December 2020	172	25 April 2023	IW	VV + VH
2	19 February 2017	59	21 February 2019	116	5 January 2021	173	7 May 2023	IW	VV + VH
3	3 March 2017	60	5 March 2019	117	17 January 2021	174	19 May 2023	IW	VV + VH
4	27 March 2017	61	17 March 2019	118	29 January 2021	175	31 May 2023	IW	VV + VH
5	8 April 2017	62	29 March 2019	119	10 February 2021	176	23 August 2023	IW	VV + VH
6	20 April 2017	63	10 April 2019	120	22 February 2021	177	4 September 2023	IW	VV + VH
7	2 May 2017	64	22 April 2019	121	6 March 2021	178	16 September 2023	IW	VV + VH
8	14 May 2017	65	4 May 2019	122	18 March 2021	179	28 September 2023	IW	VV + VH
9	26 May 2017	66	16 May 2019	123	30 March 2021	180	10 October 2023	IW	VV + VH
10	7 June 2017	67	28 May 2019	124	11 April 2021	181	22 October 2023	IW	VV + VH
11	19 June 2017	68	9 June 2019	125	23 April 2021	182	3 November 2023	IW	VV + VH
12	1 July 2017	69	21 June 2019	126	29 May 2021	183	15 November 2023	IW	VV + VH
13	13 July 2017	70	3 July 2019	127	10 June 2021	184	27 November 2023	IW	VV + VH
14	25 July 2017	71	15 July 2019	128	22 June 2021	185	9 December 2023	IW	VV + VH
15	6 August 2017	72	27 July 2019	129	16 July 2021	186	21 December 2023	IW	VV + VH
16	18 August 2017	73	8 August 2019	130	28 July 2021	187	2 January 2024	IW	VV + VH
17	30 August 2017	74	20 August 2019	131	9 August 2021	188	26 January 2024	IW	VV + VH
18	11 September 2017	75	1 September 2019	132	21 August 2021	189	7 February 2024	IW	VV + VH
19	23 September 2017	76	13 September 2019	133	2 September 2021	190	19 February 2024	IW	VV + VH
20	5 October 2017	77	25 September 2019	134	14 September 2021	191	14 March 2024	IW	VV + VH
21	29 October 2017	78	7 October 2019	135	26 September 2021	192	26 March 2024	IW	VV + VH
22	10 November 2017	79	19 October 2019	136	8 October 2021	193	7 April 2024	IW	VV + VH
23	22 November 2017	80	31 October 2019	137	20 October 2021	194	19 April 2024	IW	VV + VH

Table 2. *Cont.*

NO.	Date	NO.	Date	NO.	Date	NO.	Date	Sensor Mode	Polarization
24	4 December 2017	81	12 November 2019	138	1 November 2021	195	1 May 2024	IW	VV + VH
25	16 December 2017	82	24 November 2019	139	13 November 2021	196	25 May 2024	IW	VV + VH
26	28 December 2017	83	6 December 2019	140	25 November 2021	197	6 June 2024	IW	VV + VH
27	9 January 2018	84	18 December 2019	141	7 December 2021	198	18 June 2024	IW	VV + VH
28	21 January 2018	85	30 December 2019	142	19 December 2021	199	5 August 2024	IW	VV + VH
29	2 February 2018	86	11 January 2020	143	31 December 2021	200	17 August 2024	IW	VV + VH
30	14 February 2018	87	23 January 2020	144	12 January 2022	201	29 August 2024	IW	VV + VH
31	26 February 2018	88	4 February 2020	145	24 January 2022	202	10 September 2024	IW	VV + VH
32	10 March 2018	89	16 February 2020	146	5 February 2022	203	22 September 2024	IW	VV + VH
33	22 March 2018	90	28 February 2020	147	17 February 2022	204	4 October 2024	IW	VV + VH
34	3 April 2018	91	11 March 2020	148	13 March 2022	205	16 October 2024	IW	VV + VH
35	15 April 2018	92	23 March 2020	149	6 April 2022	206	28 October 2024	IW	VV + VH
36	27 April 2018	93	4 April 2020	150	18 April 2022	207	9 November 2024	IW	VV + VH
37	9 May 2018	94	16 April 2020	151	30 April 2022	208	21 November 2024	IW	VV + VH
38	21 May 2018	95	28 April 2020	152	24 May 2022	209	3 December 2024	IW	VV + VH
39	2 June 2018	96	10 May 2020	153	29 June 2022	210	15 December 2024	IW	VV + VH
40	14 June 2018	97	22 May 2020	154	11 July 2022	211	27 December 2024	IW	VV + VH
41	26 June 2018	98	3 June 2020	155	23 July 2022	212	8 January 2025	IW	VV + VH
42	8 July 2018	99	15 June 2020	156	4 August 2022	213	20 January 2025	IW	VV + VH
43	20 July 2018	100	27 June 2020	157	28 August 2022	214	1 February 2025	IW	VV + VH
44	1 August 2018	101	9 July 2020	158	9 September 2022	215	9 March 2025	IW	VV + VH
45	25 August 2018	102	21 July 2020	159	21 September 2022	216	21 March 2025	IW	VV + VH
46	6 September 2018	103	2 August 2020	160	20 November 2022	217	2 April 2025	IW	VV + VH
47	18 September 2018	104	14 August 2020	161	14 December 2022	218	14 April 2025	IW	VV + VH
48	30 September 2018	105	26 August 2020	162	26 December 2022	219	26 April 2025	IW	VV + VH
49	12 October 2018	106	7 September 2020	163	7 January 2023	220	8 May 2025	IW	VV + VH
50	24 October 2018	107	19 September 2020	164	19 January 2023	221	20 May 2025	IW	VV + VH
51	5 November 2018	108	1 October 2020	165	31 January 2023	222	1 June 2025	IW	VV + VH
52	29 November 2018	109	13 October 2020	166	12 February 2023	223	7 July 2025	IW	VV + VH
53	11 December 2018	110	25 October 2020	167	24 February 2023	224	19 July 2025	IW	VV + VH
54	23 December 2018	111	6 November 2020	168	8 March 2023	225	12 August 2025	IW	VV + VH
55	4 January 2019	112	18 November 2020	169	20 March 2023	226	24 August 2025	IW	VV + VH
56	16 January 2019	113	30 November 2020	170	1 April 2023				
57	28 January 2019	114	12 December 2020	171	13 April 2023				

3. Methods

3.1. Image Preprocessing

To ensure the consistency and reliability of time-series data, data preprocessing is required. The image acquired on 29 January 2021 was adopted as the master image to complete the co-registration of all time-series images. Simultaneously, basic preprocessing steps such as orbit correction, radiometric calibration, and atmospheric correction were performed, ultimately outputting normalized backscattering coefficient data. Subsequently, based on the scope of the experimental areas shown in Figure 1, the data of the target study area were clipped from the preprocessed full Sentinel-1 dataset, with detailed basic image information provided in Table 1. It should be noted that each experimental area contains 226 VV-polarized and 226 VH-polarized images. In addition, the backscattering coefficient distribution of SAR data exhibits an overall left-skewed characteristic, with a large number of abnormally high-value pixels in the images. To ensure the accuracy of subsequent extraction, outliers with a cumulative probability exceeding 99.9% in the histogram need to be removed via the threshold truncation method.

3.2. Water Extraction

The superposition of inherent speckle noise in SAR images and scattering from complex terrain features imposes significant limitations on traditional water extraction methods, including Otsu thresholding and conventional clustering. In high-noise scenarios, blurred edges and high misclassification rates are universal bottlenecks for these methods, which fail to meet the requirements of high-precision time-series monitoring. To address this limitation, this study proposes a two-step strategy of "rough screening of candidate areas—refined extraction" for the TSDPS-Clus algorithm, with its core focusing on the synergistic achievement of noise suppression and dynamic scene adaptation. Specifically, the first step involves the application of the K-S test for rapid screening of water body candidate areas. Non-water pixels and strong noise are eliminated through this test, and the processing scope is thereby reduced. The second step, designated as "SVD-Clus", achieves high-precision extraction under the constraint of the candidate areas. Given the complementarity between VV polarization (sensitive to water surface roughness) and VH polarization (enhancing the scattering difference between water and non-water features), dual-polarization features are first fused via SVD in the second step. This fusion separates effective information from residual noise, generating a low-dimensional feature set with high discriminability. Furthermore, ISODATA adaptive clustering is utilized for classification. Cluster centers and the number of clusters can be iteratively optimized by this algorithm, which enables adaptation to dynamic scenarios (e.g., seasonal fluctuations of water bodies, alternations between drying and impoundment). Under-extraction or over-extraction caused by a fixed number of clusters is thus avoided. Significant improvements in the edge accuracy and dynamic adaptability of water body distribution maps are achieved by this strategy, which provides reliable support for the time-series analysis of water body evolution in the study area over the past eight years.

3.2.1. Extraction of Water Body Candidate Areas

In this study, a two-sample K-S test was employed to extract the core water body areas and construct candidate areas. Its core advantage lies in the elimination of the need to prespecify data distribution patterns; instead, class discrimination can be directly achieved through differences in the cumulative distribution functions (CDFs) of samples, thereby accommodating the complex statistical characteristics of SAR backscattering coefficients. From a theoretical perspective, due to their inherent physical properties, water bodies in time-series SAR data exhibit generally low and stable backscattering coefficients. In contrast,

non-water features (e.g., bare land, roads, and buildings) demonstrate significantly higher backscattering coefficients. Influenced by speckle noise and atmospheric disturbances, these non-water features also exhibit greater numerical dispersion and larger fluctuation amplitudes. The significant difference in the temporal statistical distributions between the two categories provides a core basis for K-S test-based classification. Furthermore, water body distribution is susceptible to factors such as seasonal changes and geological collapses. Relying solely on temporal statistical characteristics makes it difficult to accurately capture dynamic edge changes. Therefore, the K-S test was first used to extract stable core water body areas, which were then expanded to form candidate areas to cover potential edge water bodies.

Subsequently, the SDWI is calculated by fusing VV and VH data, based on which initial water bodies are extracted. By integrating VV and VH data, the SDWI enhances water body features while mitigating interferences caused by soil and vegetation during the water extraction process [33]. The specific formula for the SDWI is presented:

$$\mathrm{K_{SDWI}} = \ln(10 \times VV \times VH) \tag{1}$$

Herein, VV and VH refer to the pixel values of the VV and VH bands in SAR images, respectively.

Subsequently, in the stacked N-dimensional time-series SDWI images, a sample P is randomly selected as the reference sample, and the observation sample based on pixel P can be expressed as follows:

$$\mathbf{x}_P = \left[x_P^1, x_P^2, \ldots, x_P^i, \ldots, x_P^N\right]^T \tag{2}$$

Herein, x_P^i denotes the SDWI value of pixel P in the i-th SAR acquisition, and T represents the transpose operator. A two-sample K-S test is performed between the reference sample and other potential water/non-water pixel samples within the study area, thereby partitioning the data into two sets. First, the CDFs of the reference sample and the sample to be tested are calculated separately, with the specific formulas provided below:

$$F_P(x) = \frac{1}{N}\sum_{i=1}^{N} I(x_P^i < x) \quad x \in \mathbf{x}_P \tag{3}$$

Herein, $I(\mathrm{A})$ is the indicator function, which equals 1 if event A holds and 0 otherwise; x_P^i represents the i-th element of the reference sample point P sorted by amplitude. The test statistic $D_{P,Q}$ is then computed, which is defined as the maximum vertical distance between the CDFs of the reference sample and the sample to be tested:

$$D_{P,Q} = \max\left\{\max_{i=1,2,\ldots,N}\left|F_P(x_P^i) - F_Q(x_P^i)\right|, \max_{i=1,2,\ldots,N}\left|F_P(x_Q^j) - F_Q(x_Q^j)\right|\right\} \tag{4}$$

Herein, Q is the sample to be tested, and F_Q is the CDF of the Q set. A larger $D_{P,Q}$ value indicates a more significant difference in the distribution between the two types of samples, implying that they belong to different categories. Ultimately, in practical applications, the p-value is typically calculated to assess the significance level, and its comparison with the threshold α is used to determine whether a statistically significant difference exists between two samples. The formula for the p-value is as follows:

$$p \approx 2e^{-N(D_{P,Q})^2} \tag{5}$$

Herein, N denotes the length of the sample. In the experiment, the threshold α was set to 0.05. If the p-value is less than the threshold, it indicates a significant difference between the two samples, which belong to two distinct sets (water bodies and non-water bodies); conversely, the two samples belong to the same set.

Notably, the K-S test is only capable of partitioning the sample set into two subsets but cannot explicitly associate these subsets with either water bodies or non-water bodies. Given that water bodies typically exhibit lower backscattering coefficients than non-water bodies, the average amplitude of each subset can be calculated to determine their categorical affiliation.

$$\begin{cases} K^1_{SDWI} \in water, K^2_{SDWI} \in non\text{-}water & \frac{1}{N_1}\sum_{i=1}^{N_1} K^{i,1}_{SDWI} \leq \frac{1}{N_2}\sum_{i=1}^{N_2} K^{i,2}_{SDWI} \\ K^2_{SDWI} \in non\text{-}water, K^2_{SDWI} \in water & \frac{1}{N_1}\sum_{i=1}^{N_1} K^{i,1}_{SDWI} < \frac{1}{N_2}\sum_{i=1}^{N_2} K^{i,2}_{SDWI} \end{cases} \quad (6)$$

Herein, 1 and 2 denote Set 1 and Set 2, respectively, while N_1 and N_2 represent their total sample sizes. i denotes the i-th element in a subset, while j denotes the j-th subset. The subset with a smaller average amplitude is classified as water bodies, and the other as non-water bodies.

Finally, to accurately capture subtle variation regions at water body edges that are easily overlooked, a 20-pixel buffer zone is outwardly expanded around the identified stable water body boundaries to form a more precise water body candidate region. This approach thereby achieves preliminary elimination of non-water background noise and effectively reduces the subsequent processing scope, laying a solid foundation for follow-up refined extraction.

3.2.2. Fine Extraction of Water Using SVD-Clus

a. Water Feature Fusion

However, despite the preliminary screening of water body candidate regions and the reduced spatial scope for subsequent processing enabled by the two-sample K-S test, its core limitations remain unresolved. Specifically, the candidate regions retain inherent speckle noise of SAR images, which exhibits a multiplicative distribution and tends to induce grayscale fluctuations in water pixels. They also contain a small number of non-water interfering pixels with scattering characteristics similar to water bodies—typically building shadows (low backscattering), moist bare soil whose scattering coefficients approach those of shallow water in the rainy season, and low-vegetation-covered areas with mixed scattering signals. These interferents show extremely low separability from water pixels in a single feature dimension, directly creating an accurate bottleneck for traditional single-feature extraction methods that prevents them from meeting the high requirements for edge precision and category purity in time-series water body evolution monitoring. Thus, enhancing the feature discriminability between water bodies and interfering ground objects and further suppressing residual noise are core prerequisites for refined extraction.

To address the aforementioned limitations, this study adopts a technical approach of "complementary fusion of multi-source features and SVD dimensionality reduction for enhancement," aiming to improve water body discrimination capability by capitalizing on the advantages of multi-dimensional features. The constructed multi-source feature set is not a mere superposition but a careful selection based on the scattering mechanisms and spatial characteristics of ground objects, comprising four core features: (1) VV-polarized backscattering coefficients from Sentinel-1A satellite, which is sensitive to the surface roughness and calmness of water bodies, thereby enabling effective differentiation between calm and disturbed water; (2) VH-polarized backscattering coefficients, which captures scatter-

ing differences between water and non-water bodies with greater effectiveness—offering distinct advantages in distinguishing shallow water from moist bare soil; (3) the SDWI, which enhances the initial water body identification signal through the combination of polarization features and mitigates residual atmospheric effects; and (4) pixel spatial position information (spatial texture features derived from coordinates (x,y)), which utilizes the spatial continuity constraint of water bodies to minimize misclassification of isolated noise pixels. These four features form a discrimination framework covering two dimensions (scattering characteristics and spatial structure), thereby achieving comprehensive characterization of water and non-water ground objects. The feature space of any pixel $S_{i,j}$ is therefore expressed as follows:

$$\mathbf{S}_{i,j} = (\mathbf{VV}_{i,j}, \mathbf{VH}_{i,j}, \mathbf{K}_{SDWI,i,j}, i, j) \tag{7}$$

Herein, i and j denote the row and column indices of the pixel, respectively; VV and VH represent the two types of original SAR polarization data; and $K_{SDWI,i,j}$ denotes the Water Body Index value of the corresponding pixel.

However, direct fusion of multi-source features leads to a significant increase in feature dimensionality. Specifically, this not only elevates the computational cost of subsequent clustering algorithms but also risks introducing the "curse of dimensionality" due to feature redundancy, ultimately compromising the model's generalization performance. To mitigate this issue, SVD is employed as an unsupervised tool for feature dimensionality reduction and information enhancement. It offers distinct core advantages: it requires no assumptions about data distribution, operates without labeled data, and is well-suited to the complex statistical characteristics of multi-source heterogeneous features in this study. Notably, it can effectively remove redundant components and noise while preserving core discriminative information. The core principle of SVD is to decompose any dimensional matrix into the product of three orthogonal matrices. Its mathematical expression is given as follows:

$$\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{8}$$

Herein, S denotes the $m \times 5$ candidate region feature matrix, with m being the total number of pixels in the candidate region; T denotes the transpose operation. U is a $m \times m$ left singular matrix, whose column vectors correspond to pixel sets with similar feature patterns in the candidate region, thus enabling characterization of pixel spatial distribution. Σ is a $m \times 5$ diagonal matrix, where diagonal elements are the singular values of S. The magnitude of a singular value directly reflects the information energy of its corresponding singular vector, with larger values indicating higher importance of the associated feature dimension. V is a 5×5 right singular matrix, whose column vectors capture the linear correlations among the original features. These definitions establish a theoretical basis for subsequent SVD-based feature selection and dimensionality reduction, thereby ensuring the validity of the dimensionality reduction process.

For feature fusion in this study, the candidate region feature matrix is first constructed and subjected to standardization. This mitigates the effects of dimensionality discrepancies among features, thereby ensuring balanced feature weights during fusion. SVD is then performed on matrix **S**, followed by the selection of core feature dimensions based on the cumulative contribution rate of singular values. Finally, the feature matrix is reconstructed using the selected singular values and their corresponding vectors to obtain the low-dimensional feature matrix $\mathbf{S}_r$:

$$\mathbf{S}_r = \begin{pmatrix} \mathbf{U}_r & \mathbf{U}_{m-r} \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_r^T \\ \mathbf{V}_{5-r}^T \end{pmatrix} \tag{9}$$

Herein, r denotes the number of retained singular values, and D_r represents the first r retained singular values. Specifically, only the largest singular value is retained, with coefficients that have no impact on clustering omitted: $\mathbf{S}_r = \mathbf{U}_1$. In this expression, $\mathbf{S}_r$ corresponds to the first column vector of the left singular matrix U. This column vector serves as the input for clustering, encapsulating the primary features of the image.

The low-dimensional feature set fused via SVD exhibits three key advantages: first, dimensionality is substantially reduced, which significantly improves the computational efficiency of subsequent ISODATA clustering; second, core discriminative information is highly concentrated, leading to a significant increase in the inter-class distance between water bodies and interfering ground objects, thus substantially enhancing separability; third, residual speckle noise and interference signals are effectively mitigated, resulting in a notable improvement in the signal-to-noise ratio of the subsequent feature matrix. This fused feature set delivers high-quality input data for subsequent ISODATA clustering, thus fundamentally safeguarding the accuracy and stability of refined water body extraction.

b. ISODATA

Within the low-dimensional, high-discriminability feature space derived from SVD-based feature fusion, the ISODATA is further adopted for refined water body extraction in the candidate region. While SVD effectively suppresses residual noise and enhances feature separability, feature overlap persists between water bodies and interfering pixels such as building shadows and moist bare soil in the candidate region. Furthermore, water bodies in the study area display prominent dynamic evolutionary characteristics such as seasonal fluctuations and drying-impoundment alternations. Traditional fixed-cluster algorithms, such as K-Means, lack the ability to adaptively adjust cluster counts, often resulting in under-extraction or over-extraction of water bodies. Leveraging its core logic of "iterative optimization-adaptive splitting-merging," ISODATA can dynamically adjust cluster centers and counts, enabling adaptation to the aforementioned complex scenarios. Details of its core principles and implementation workflow are provided below.

ISODATA is a classic unsupervised clustering algorithm. Compared to fixed-cluster counterparts such as K-Means, its key advantage is the capacity to dynamically adjust cluster centers and counts via a "splitting-merging" mechanism. No predefinition of fixed cluster counts is required, offering superior adaptability to dynamic changes in data distribution. Its streamlined workflow is summarized as follows: initialization of cluster centers and key parameters; assignment of samples to the nearest clusters using distance metrics; iterative optimization—assessing intra-class dispersion to decide on cluster splitting, evaluating inter-class distances for cluster merging, and updating cluster centers simultaneously; and finally, output of clustering results once convergence criteria are satisfied such as cluster center variation below a threshold or maximum iteration count reached.

In the ISODATA clustering algorithm, Euclidean distance is used to calculate the similarity between pixels in the candidate region and each cluster center, thereby determining the class to which each pixel belongs. For the low-dimensional feature vectors $\mathbf{x}_{ii} = (x_{i1}, x_{i2}, \ldots, x_{id})$ after feature fusion and the cluster centers $\mathbf{z}_j = (z_{j1}, z_{j2}, \ldots, z_{jd})$, their Euclidean distance is expressed as follows:

$$d(\mathbf{x}_i, \mathbf{z}_j) = \sqrt{\sum_{k=1}^{d} (x_{ik} - z_{jk})^2} \tag{10}$$

In this expression, x_{ik} denotes the value of the i-th pixel on the k-th feature dimension, and z_{jk} denotes the value of the j-th cluster center on the k-th feature dimension. The smaller $d(\mathbf{x}_i, \mathbf{z}_j)$ is, the higher the feature similarity between the pixel and the cluster center.

In the clustering algorithm, after each iteration, the cluster centers are recalculated based on the features of pixels belonging to each cluster to achieve center optimization. If, in the t-th iteration, the j-th cluster $G_j^{(t)}$ contains $N_j^{(t)}$ pixels, its updated cluster center is given by:

$$\mathbf{z}_j^{(t+1)} = \frac{1}{N_j^{(t)}} \sum_{\mathbf{x}_i \in G_j^{(t)}} \mathbf{x}_i = \left[\frac{1}{N_j^{(t)}} \sum_{\mathbf{x}_i \in G_j^{(t)}} x_{i1}, \ldots, \frac{1}{N_j^{(t)}} \sum_{\mathbf{x}_i \in G_j^{(t)}} x_{id} \right] \tag{11}$$

Notably, splitting and merging are the key processes differentiating this clustering algorithm from K-Means. The core objective of splitting is to address the issue where a single cluster contains heterogeneous samples. In the candidate region for water body extraction, certain clusters may concurrently contain water pixels such as calm water and disturbed water, as well as interfering pixels with similar scattering properties, such as building shadows and moist bare soil. This leads to excessively high intra-cluster feature dispersion, which directly impairs classification purity. To accurately quantify this dispersion, the intra-cluster feature standard deviation is adopted as the core criterion. It essentially quantifies the deviation of all pixels within a cluster from the cluster center across each feature dimension; a larger standard deviation indicates a more scattered sample distribution in that dimension and stronger intra-cluster heterogeneity. For the d-dimensional feature space derived from SVD fusion, in the t-th iteration, the formula for the standard deviation of the j-th cluster in the k-th feature dimension is expressed as follows:

$$\sigma_{jk}^{(t)} = \sqrt{\frac{1}{N_j^{(t)}} \sum_{i=1}^{N_j^{(t)}} \left(x_{ik} - z_{jk}^{(t)} \right)^2} \tag{12}$$

In this formula, $N_j^{(t)}$ denotes the total number of pixels in the j-th cluster, x_{ik} denotes the value of the i-th pixel within this cluster on the k-th feature dimension, and $z_{jk}^{(t)}$ denotes the center value of this cluster for the k-th feature dimension. When the standard deviation of a specific feature dimension exceeds the preset threshold θ_S, this dimension is identified as the core factor inducing cluster heterogeneity, and the cluster should be split into two new clusters along this dimension. Specifically, the cluster center for this dimension serves as the dividing point to partition intra-cluster pixels into two subsets, each forming a new cluster with its center recalculated. Let the maximum standard deviation of cluster j correspond to the k^*-th feature dimension; then $\mathbf{z}_j^{(t+1)}$ is split into two new centers, $\mathbf{z}_{j,1}^{(t+1)}$ and $\mathbf{z}_{j,2}^{(t+1)}$:

$$\begin{cases} \mathbf{z}_{j,1}^{(t+1)} = \mathbf{z}_j^{(t+1)} + \gamma \cdot \sigma_{jk^*}^{(t)} \cdot \mathbf{e}_{k^*} \\ \mathbf{z}_{j,2}^{(t+1)} = \mathbf{z}_j^{(t+1)} - \gamma \cdot \sigma_{jk^*}^{(t)} \cdot \mathbf{e}_{k^*} \end{cases} \tag{13}$$

Herein, $\gamma \in (0,1)$ is the splitting coefficient (typically set to 0.5), and $\mathbf{e}_{k^*}$ denotes the unit vector of the k^*-th dimension (i.e., the k^*-th element of $\mathbf{e}_{k^*}$ is 1, with all others being 0). For the context of this study, θ_S is determined via pre-experiments. It is set to avoid over-splitting, such as redundant small clusters caused by an excessively small threshold, and to prevent insufficient splitting, such as mixed heterogeneous samples due to an excessively large threshold—ultimately achieving effective separation between water pixels and interfering pixels.

The merging operation aims to mitigate excessive clustering redundancy arising from over-splitting. The splitting process may split the same type of ground object, for instance, calm water across different regions, into multiple clusters or produce small clusters consisting solely of a few noise pixels, a phenomenon that not only elevates computational

costs but also impairs the spatial continuity of ground objects. The core criterion for merging is the inter-cluster center distance. Specifically, by quantifying the similarity of core features across different clusters, highly similar clusters are merged to ensure the simplicity and accuracy of category partitioning. For the *i*-th and *j*-th clusters, their center distance is computed using the Euclidean distance defined previously (Equation (10)). A smaller distance indicates greater similarity in feature patterns between the two clusters, which are thus likely to belong to the same ground object type. When the center distance between two clusters is smaller than the preset threshold θ_C, the clusters are marked for merging, with the center of the merged cluster being the feature mean of all pixels from both clusters. Cluster *i* and cluster *j* are merged into cluster m, as given by the following formula:

$$\mathbf{z}_m^{(t+1)} = \frac{N_i^{(t)} \cdot \mathbf{z}_i^{(t+1)} + N_j^{(t)} \cdot \mathbf{z}_j^{(t+1)}}{N_i^{(t)} + N_j^{(t)}} \tag{14}$$

In the water body extraction for this study, the value of θ_C must be tailored to the spatial distribution characteristics of water bodies. It must not only ensure that scattered small water body clusters, such as scattered water areas at reservoir edges, are merged into a single complete water body category but also prevent incorrect merging of water bodies with adjacent low-scattering interference clusters (e.g., small, shadowed regions). Ultimately, the dynamic balance between splitting and merging enables the accurate partitioning of water body categories.

In this study's experiments, the ISODATA parameter settings carry clear physical meanings: the preset expected number of clusters is set to 2 (initially corresponding to water and non-water categories) and is adaptively adjusted through the splitting-merging mechanism; the minimum number of cluster pixels θ_N is set according to the pixel density in the candidate region, preventing the formation of spurious clusters by isolated noise; the splitting threshold θ_S and merging threshold θ_C are determined via multiple pre-experiments, ensuring compatibility with the feature differences between water bodies and interfering pixels in the study area. Via iterative optimization, the algorithm achieves accurate differentiation between water bodies and residual interfering pixels within the SVD-fused feature space, while accommodating dynamic scenarios such as water body area fluctuations and drying-impoundment alternations. The final refined water body extraction results significantly enhance the edge integrity and category purity of water bodies, laying a high-quality data foundation for subsequent temporal water body evolution analysis.

Notably, ISODATA is only capable of partitioning pixels in the water body candidate region into two clusters without establishing the correspondence between these clusters and the water/non-water classes. To resolve this, leveraging the inherent characteristic of water bodies, i.e., small and stable backscattering coefficients, Equation (11) is employed to identify the cluster with a smaller average amplitude as the water body, with the remaining cluster designated as non-water.

4. Results

4.1. Water Body Extraction Results

To verify the effectiveness and superiority of the proposed TSDPS-Clus algorithm for water body extraction in karst areas, experiments were designed and implemented systematically. Hardware configuration: Intel Core i9-10900K CPU, 32 GB RAM; all programs involved in the experiments were independently developed using C++ (Microsoft Visual Studio 2022).

In this study, the TSDPS-Clus algorithm was utilized for high-precision water body extraction across the dataset, with comparative validation performed against water body

extraction results obtained via the SVD-Clus algorithm applied directly to the entire image. Using the experimental parameters established in the preceding sections, the water body extraction results acquired on 23 September 2017 are presented in Figure 2, where the extracted water body regions and water body candidate regions are overlaid in red and blue, respectively.

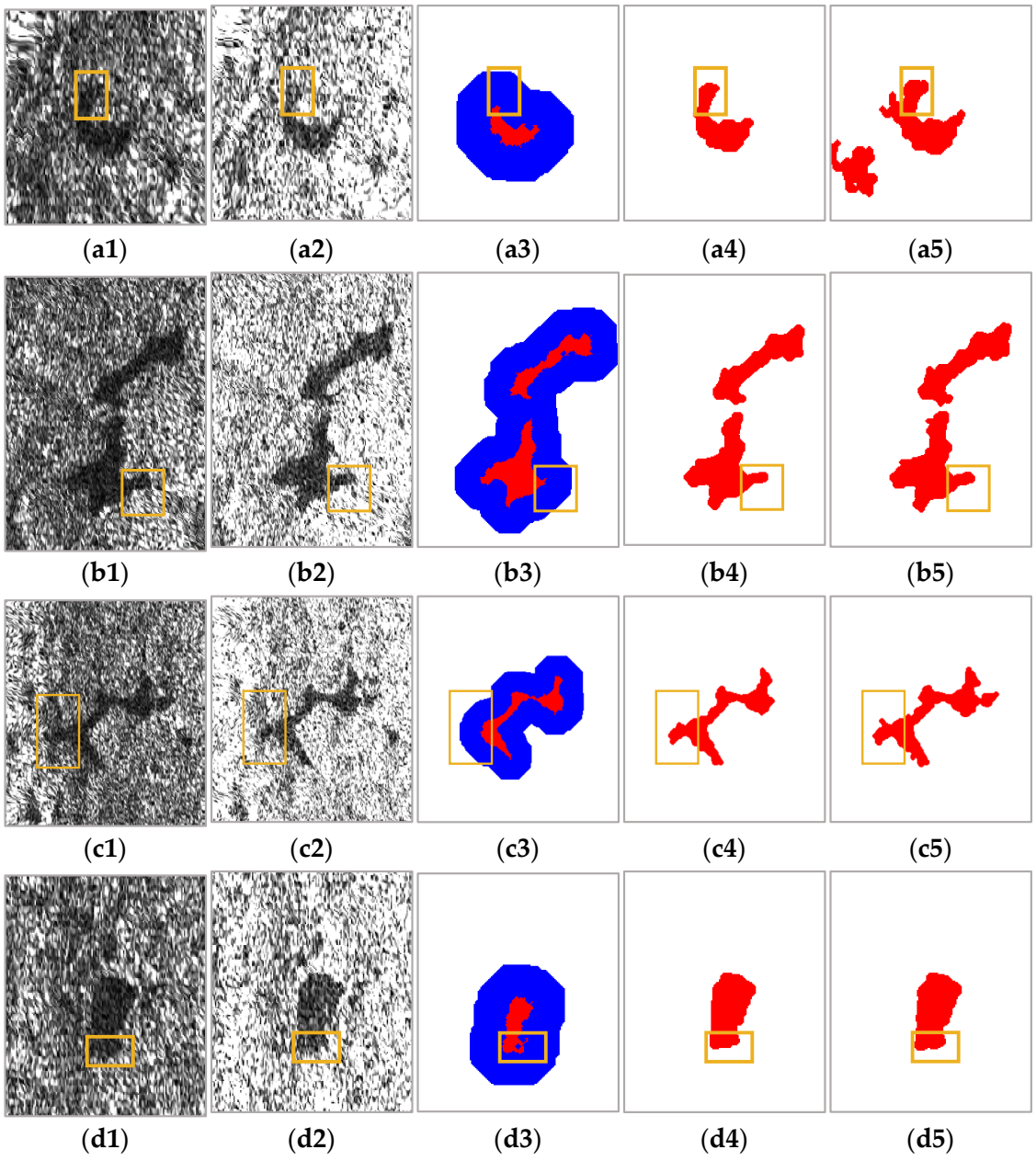


Figure 2. *Cont.*

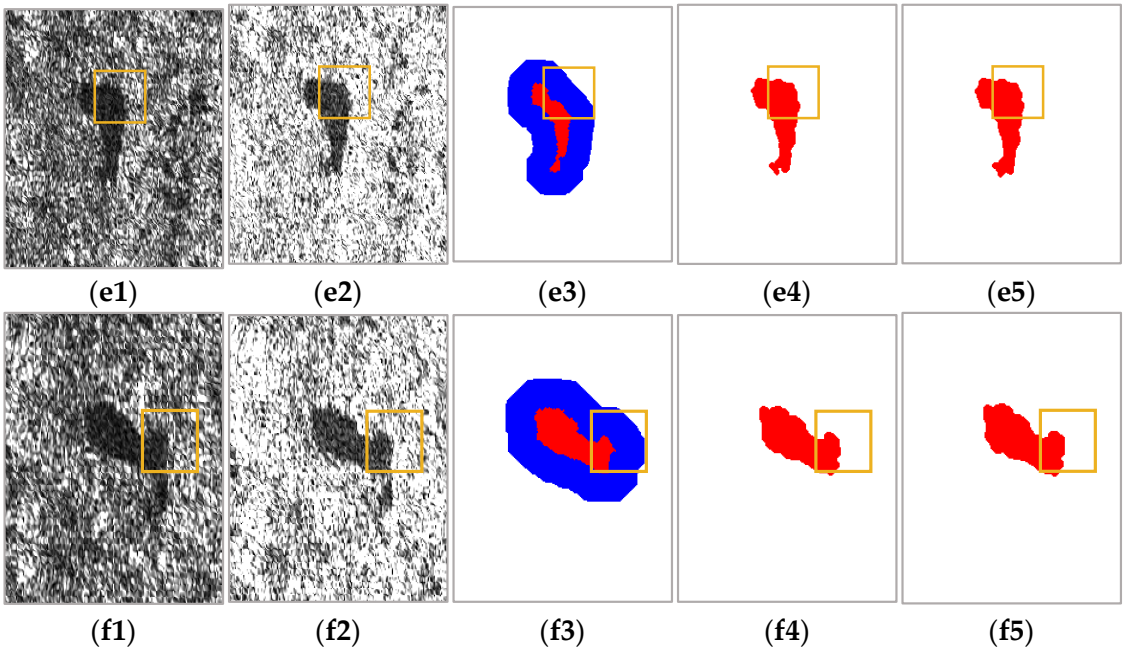


Figure 2. In the ISODATA clustering using different methods on the 23 September 2017 dataset. (**a**–**f**): Reservoirs A–F; (**a1**–**f1**): the original VH images; (**a2**–**f2**): the original VV images; (**a3**–**f3**): the water body candidate regions; (**a4**–**f4**): TSDPS-Clus results; (**a5**–**f5**): SVD-Clus results (without water body candidate regions). The highlighted colors: water (red), expanded water body candidate area (blue). The yellow box highlights the differences among the results of different methods.

As illustrated in Figure 2(a1–f1,a2–f2), the core characteristics of water bodies are highly consistent in the original VV and VH polarization data. In contrast, non-water regions show notable discrepancies. This provides complementary information for multi-dimensional water body extraction. Figure 2(a3–f3) shows the water body candidate regions extracted via the K-S test in the first stage. Red regions represent core water bodies with stable long-term distribution. The outer blue regions cover water expansion/contraction zones, driven by factors like seasonal variations. Detailed analysis of these zones enables accurate capture of water body edge details. We compared Figure 2(a3–f3,a4–f4). The second-stage results retain richer water body edge details than the first-stage outputs. They also align more closely with the actual water body distribution in the original data. This validates the effectiveness of the TSDPS-Clus algorithm's second stage. Figure 2(a5–e5) shows water body extraction results from the direct application of the SVD-Clus algorithm. Compared with the TSDPS-Clus algorithm, its extraction results are generally consistent in Regions B and D–F. However, non-water regions are misclassified as water bodies in Figure 2(a5). Additionally, notable inaccuracies in water body edge extraction are seen in the orange boxes of Figure 2(c5). This highlights the value of the TSDPS-Clus algorithm's first stage (water body candidate region extraction). It effectively removes shadows and noise from non-water regions. This significantly improves water body extraction accuracy. In summary, Figure 2 fully confirms the necessity of the TSDPS-Clus algorithm's two-stage design. The two stages work synergistically. This facilitates the development of a more versatile water body extraction scheme. It also enables the precise capture of water body spatial distribution information.

4.2. Feature Fusion Result

To validate the effectiveness of the feature fusion module in the TSDPS-Clus algorithm, Figure 3 presents the histogram of original SAR data and the 1D/2D distribution characteristics of fused features for Reservoir F on 14 May 2017. As shown in Figure 3a,b, both

the original VH and VV polarization data exhibit a unimodal distribution. This makes it difficult to effectively distinguish between water bodies and non-water bodies directly from the histograms. Furthermore, the overall data shows a left-skewed trend with extremely high-value outliers. These outliers tend to interfere with the determination of classification thresholds and reduce water extraction accuracy. In terms of numerical distribution, VH polarization data are mainly concentrated in the [0, 200] interval. VV polarization data, by contrast, cluster in the [0, 400] interval. Compared with single-polarization data, the complementarity of dual-polarization data provides richer ground object scattering information. This lays a foundation for enhancing water body discrimination capabilities. Figure 3c shows the histogram of 1D feature vectors when only the maximum singular value is retained. It exhibits a distinct bimodal distribution. This indicates that this feature can effectively support the clustering algorithm in separating water bodies from non-water bodies. Figure 3d depicts the 2D feature space distribution when the first two singular values are retained. Notably, the two types of samples (water and non-water) form clear bilateral clustering patterns in the feature space. The significant inter-class gaps greatly reduce the difficulty of clustering classification. This facilitates the accurate extraction of the water body spatial distribution. Collectively, Figure 3 fully confirms that fusing dual-polarization SAR data via SVD and condensing core discriminative information can transform the original indivisible unimodal distribution into a cluster-friendly bimodal distribution. This provides high-quality feature input for the subsequent precise classification of the clustering algorithm and lays a solid data foundation.

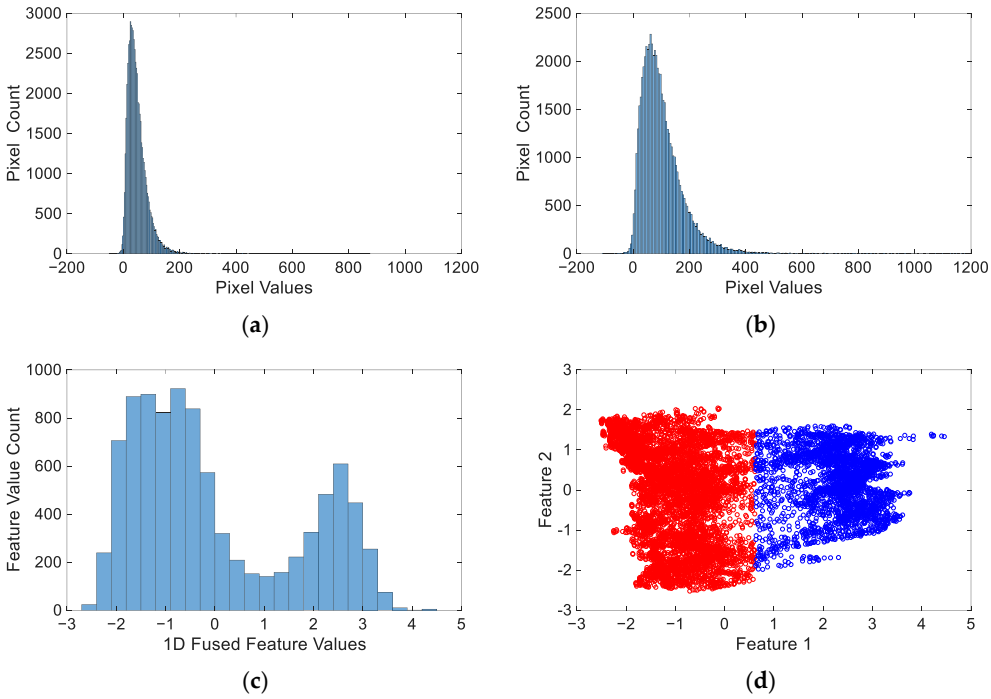


Figure 3. Histogram of the 14 May 2017 dataset for Reservoir F. (**a**) Original VH polarization data; (**b**) Original VV polarization data; (**c**) 1D fused feature histogram (maximum singular value); (**d**) 2D fused feature distribution (first two singular values). Colors in the figure: water (red), non-water (blue).

5. Discussion

5.1. Dual-Polarization Data

Compared with single-polarization SAR data, dual-polarization (VV + VH) SAR offers significant advantages and unique value in water body extraction within complex karst

terrain. This is primarily stemming from its polarization complementarity and ability to characterize multi-dimensional ground object information. Specifically, the VV and VH polarization channels correspond to different ground object scattering mechanisms. They can simultaneously depict the surface roughness, dielectric properties, and geometric structure of ground targets. This inherently overcomes the inherent limitation of single-polarization data—its ability to only characterize a single scattering mechanism—laying a core foundation for enhancing the separability between water and non-water features.

Taking the measured data of Reservoir A on 1 August 2018 as an example, Figure 4 intuitively verifies the complementary value of dual-polarization data. VV polarization data, affected by the coupling of minor water surface perturbations and scattering from surrounding karst landforms, struggle to effectively distinguish water bodies from background features. In contrast, VH polarization data can significantly enhance the scattering difference between water bodies and background features, enabling accurate delineation of water body boundaries. This demonstrates that the combined application of dual-polarization data can effectively avoid water body omission issues that are common with single-polarization data. It is thus well-suited for the complex scattering scenarios in the Geleshan karst area.

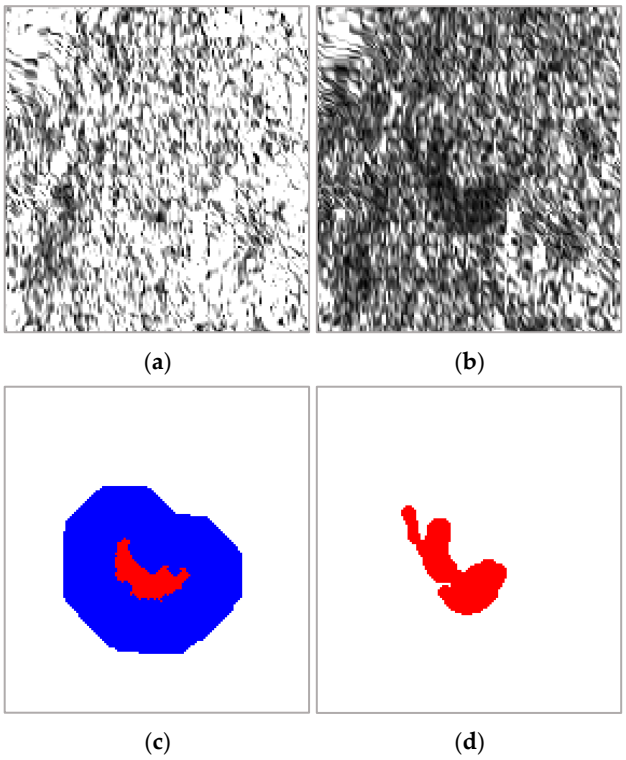


Figure 4. 1 August 2018 Reservoir A data. (**a**) VV; (**b**) VH; (**c**) water body candidate regions; (**d**) TSDPS-Clus water extraction results. The highlighted colors: water (red), expanded water body candidate area (blue).

Notably, the TSDPS-Clus algorithm proposed in this study achieves in-depth adaptation and efficient exploitation of dual-polarization features through a two-stage design. Stage 1 fuses VV and VH polarization features based on water indices and combines the K-S test to accurately extract core water body areas and construct candidate regions. This creates initial noise suppression and reduces the processing scope. Furthermore, Stage 2 leverages SVD technology to fuse dual-polarization features with spatial location information. This condenses core discriminative information and improves feature distinguishability. This hierarchical fusion strategy fully unleashes the complementary potential of dual-polarization

data, significantly enhancing the algorithm's robustness and anti-interference capability in complex scattering scenarios. It thereby provides core technical support for the accuracy and stability of long-term water body monitoring.

5.2. Long-Term Evolution Monitoring of Reservoirs

To thoroughly investigate the spatiotemporal evolution characteristics of reservoirs in the Geleshan area over an approximately 8-year long-term period (7 February 2017–24 August 2025), this study employed the previously proposed TSDPS-Clus algorithm. The algorithm was used to continuously monitor and quantitatively analyze the water persistence status of six core reservoirs in the study area. Monitoring results are presented in Figure 5, where "0" indicates no water signal and "1" denotes the detection of water.

Temporal monitoring results in Figure 5 show that Reservoirs A and B maintained a stable impounded state throughout the entire monitoring period, with no drying events observed. Combined with their geological settings and field survey data, both reservoirs are equipped with comprehensive artificial anti-seepage facilities. Moreover, their water supply is dominated by stable surface runoff, with minimal impact from groundwater dynamics. This characteristic indicates stable operational performance, making them suitable as baseline references for regional water body evolution monitoring.

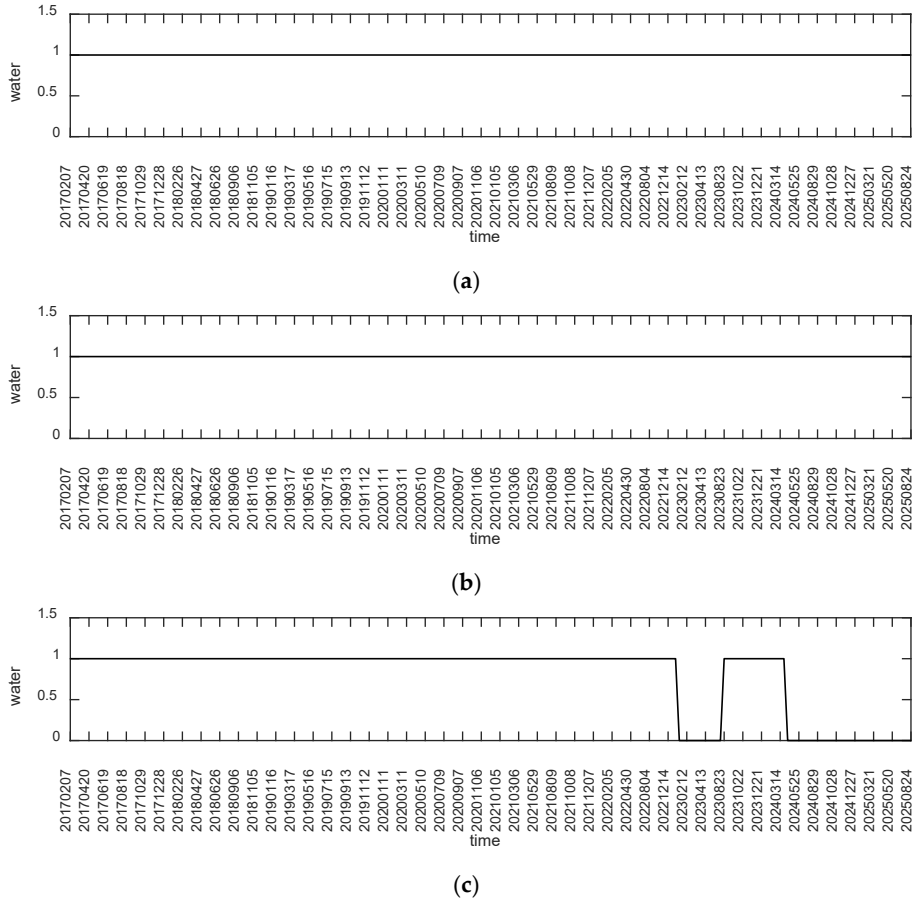


Figure 5. *Cont.*

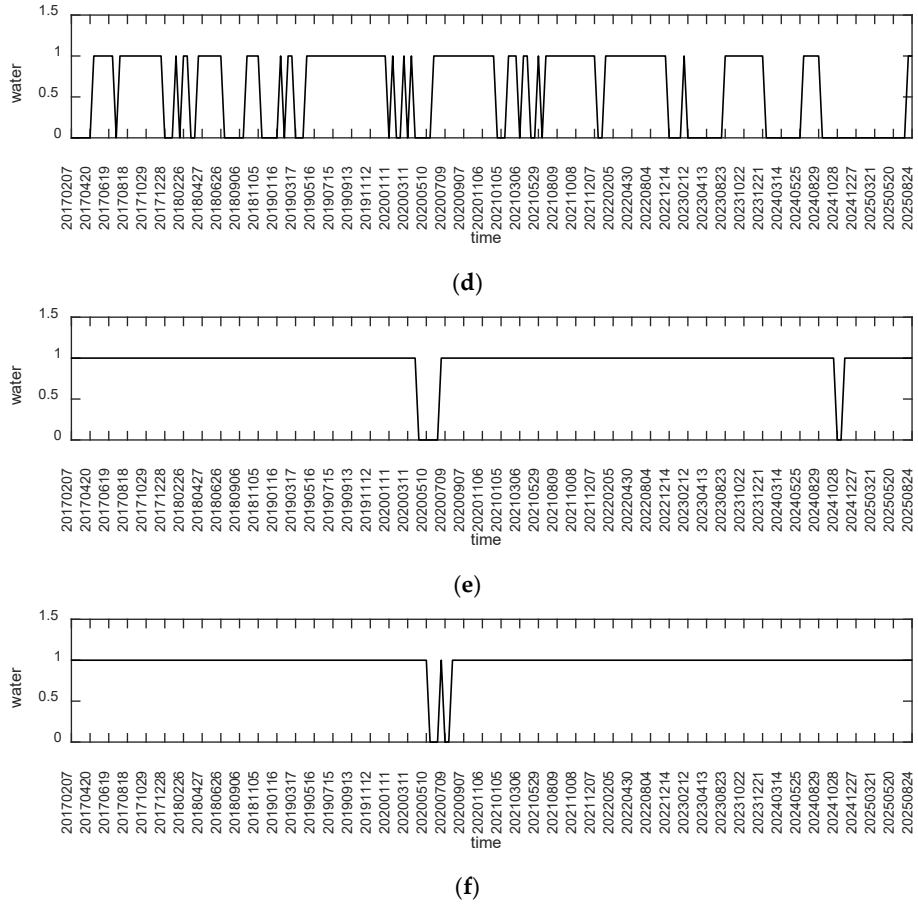


Figure 5. Drying status of each reservoir during 7 February 2017–24 August 2025. 0 = no water; 1 = water present. (**a**–**f**) Reservoirs A–F.

In sharp contrast, Reservoir C exhibited significant phased drying characteristics. It experienced two distinct drying periods: 31 January 2023–23 August 2023 (lasting approximately 7 months) and 7 April 2024–24 August 2025 (having persisted for 17 months as of the end of monitoring). Reservoir D, however, displayed high-frequency periodic drying. It recorded a total of 22 drying events during the monitoring period, with drying periods highly concentrated between November and February each year. This pattern reflects distinct seasonal periodicity, which is highly consistent with the climatic features of the Geleshan area—sharply reduced precipitation and insufficient groundwater recharge during the winter dry season.

Notably, the drying characteristics of Reservoirs E and F showed strong spatial correlation and synchronization. Both underwent sudden drying between May and July 2022, with similar drying durations. This phenomenon is not accidental. As indicated by the spatial distribution in Figure 1, the two reservoirs are less than 1 km apart. Their hydrological processes are regulated by the same meteorological factors and geological conditions, thus exhibiting highly consistent drying responses. Additionally, Reservoir E experienced its second drying event during the monitoring period in November 2024.

Reservoirs C–F are confirmed to be all within the influence radius of the Geleshan tunnel group. Large-scale excavation activities since 2023 have induced both strong groundwater drawdown and land subsidence, which jointly drive the expansion of karst fractures and increased reservoir bottom seepage [42,43]. This conclusion is highly consistent with the temporal monitoring results of this study: Before 2023, Reservoirs C, E, and F had an average annual drying frequency of less than once, with each drying event lasting no more than 3 months. After 2023, the number of drying events increased significantly, and

the duration of individual drying events prolonged. Among these, Reservoirs C and D were most severely affected by the tunnel project, with a cumulative drying duration of 24 months each after 2023.

6. Conclusions

Monitoring the spatiotemporal evolution of small water bodies, especially abnormal changes induced by land subsidence, is crucial for water resource management and disaster assessment. Focusing on the Geleshan karst sensitive area, this study proposes an unsupervised TSDPS-Clus algorithm for water body mapping using 452 time-series dual-polarization Sentinel-1 SAR images acquired from 2017 to 2025, aiming to explore the correlation between long-term water body evolution and engineering disturbances (i.e., tunnel excavation-induced land subsidence).

The TSDPS-Clus algorithm achieves efficient extraction in two phases. First, it combines the K-S test with dual-polarization time-series characteristics to screen core water body areas, construct candidate regions, eliminate interferences, and significantly reduce the overall processing scope; second, it fuses dual-polarization and spatial features via SVD and couples with ISODATA clustering optimization, effectively distinguishing water bodies from low-scatter interfering features and providing a new paradigm for water body extraction in complex karst areas. Verification results demonstrate that TSDPS-Clus can accurately extract water bodies and adapt to dynamic scenarios of drying and water storage. Furthermore, dual-polarization SAR significantly improves extraction accuracy through feature complementarity.

Applying this algorithm to monitor six reservoirs in the study area reveals that Reservoirs A and B maintain stable water storage due to favorable geological conditions and anti-seepage measures, serving as regional monitoring benchmarks. Reservoirs C and D undergo periodic drying, with notably increased frequency and duration post-2023; the cumulative drying time has reached 24 months due to land subsidence induced by tunnel construction. Reservoir D, in particular, dries periodically during the winter dry season, consistent with regional climate patterns. Reservoirs E and F share a karst recharge network, resulting in highly synchronized drying characteristics.

Looking ahead, research will integrate topographic and surface deformation data to further analyze the drivers of water body changes, thereby enhancing the scientific value of this study.

Author Contributions: Conceptualization, K.Z.; methodology, F.G.; Writing—original draft preparation, T.J.; writing—review and editing, Q.K. and T.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 42074034.

Data Availability Statement: The Sentinel-1 SAR data listed in Table 1 are available in https://dataspace.copernicus.eu/ (accessed on 13 September 2025). The other data presented in this study are available on request from the corresponding author.

Acknowledgments: The author would like to express gratitude to ESA for providing Sentinel-1 SAR data free of charge.

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

References

1. Jiang, G.; Chen, Z.; Siripornpibul, C.; Haryono, E.; Nguyen, N.X.; Oo, T.; Manzano, L.S.J.; Vongphachanh, S.; Kong, S.; Guo, F. The karst water environment in Southeast Asia: Characteristics, challenges, and approaches. *Hydrogeol. J.* **2021**, *29*, 123–135. [CrossRef]

2. Lai, Y.; Li, S.; Guo, J.; Zhu, Z.; Huang, X. Analysis of Seepage and Displacement Field Evolutionary Characteristics in Water Inrush Disaster Process of Karst Tunnel. *Geofluids* **2021**, *2021*, 5560762. [CrossRef]
3. Wu, Y. Mechanism analysis of hazards caused by the interaction between groundwater and geo-environment. *Environ. Geol.* **2003**, *44*, 811–819. [CrossRef]
4. Gui, H.; Qiu, H.; Chen, Z.; Ding, P.; Zhao, H.; Li, J. An overview of surface water hazards in China coal mines and disaster-causing mechanism. *Arab. J. Geosci.* **2020**, *13*, 67. [CrossRef]
5. Farhadi, H.; Esmaeily, A.; Najafzadeh, M. Flood monitoring by integration of Remote Sensing technique and Multi-Criteria Decision Making method. *Comput. Geosci.* **2022**, *160*, 105045. [CrossRef]
6. Markert, K.N.; Chishtie, F.; Anderson, E.R.; Saah, D.; Griffin, R.E. On the merging of optical and SAR satellite imagery for surface water mapping applications. *Results Phys.* **2018**, *9*, 275–277. [CrossRef]
7. Guo, Z.; Wu, L.; Huang, Y.; Guo, Z.; Zhao, J.; Li, N. Water-Body Segmentation for SAR Images: Past, Current, and Future. *Remote Sens.* **2022**, *14*, 1752. [CrossRef]
8. Mason, D.C.; Davenport, I.J.; Neal, J.C.; Schumann, G.J.-P.; Bates, P.D. Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3041–3052. [CrossRef]
9. Chen, L.; Zhang, P.; Xing, J.; Li, Z.; Xing, X.; Yuan, Z. A multi-scale deep neural network for water detection from SAR images in the mountainous areas. *Remote Sens.* **2020**, *12*, 3205. [CrossRef]
10. Naboureh, A.; Li, A.; Bian, J.; Lei, G.; Nan, X.; Zhang, Z. A Comprehensive Analysis of the Different Remote Sensing-Based Approaches for Surface Water Mapping in Mountainous Areas. *Earth Syst. Environ.* **2025**, 1–21. [CrossRef]
11. Ding, X.; Li, X. Monitoring of the water-area variations of Lake Dongting in China with ENVISAT ASAR images. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 894–901. [CrossRef]
12. Bartsch, A.; Trofaier, A.; Hayman, G.; Sabel, D.; Schlaffer, S.; Clark, D.; Blyth, E. Detection of open water dynamics with ENVISAT ASAR in support of land surface modelling at high latitudes. *Biogeosciences* **2012**, *9*, 703–714. [CrossRef]
13. Santoro, M.; Wegmüller, U.; Lamarche, C.; Bontemps, S.; Defourny, P.; Arino, O. Strengths and weaknesses of multi-year Envisat ASAR backscatter measurements to map permanent open water bodies at global scale. *Remote Sens. Environ.* **2015**, *171*, 185–201. [CrossRef]
14. Liebe, J.R.; van de Giesen, N.; Andreini, M.S.; Steenhuis, T.S.; Walter, M.T. Suitability and limitations of ENVISAT ASAR for monitoring small reservoirs in a semiarid area. *IEEE Trans. Geosci. Remote Sens.* **2008**, *47*, 1536–1547. [CrossRef]
15. Bolanos, S.; Stiff, D.; Brisco, B.; Pietroniro, A. Operational surface water detection and monitoring using Radarsat 2. *Remote Sens.* **2016**, *8*, 285. [CrossRef]
16. Zakhvatkina, N.; Korosov, A.; Muckenhuber, S.; Sandven, S.; Babiker, M. Operational algorithm for ice–water classification on dual-polarized RADARSAT-2 images. *Cryosphere* **2017**, *11*, 33–46. [CrossRef]
17. Li, J.; Wang, S. An automatic method for mapping inland surface waterbodies with Radarsat-2 imagery. *Int. J. Remote Sens.* **2015**, *36*, 1367–1384. [CrossRef]
18. Hong, S.; Jang, H.; Kim, N.; Sohn, H.-G. Water area extraction using RADARSAT SAR imagery combined with landsat imagery and terrain information. *Sensors* **2015**, *15*, 6652–6667. [CrossRef]
19. Hess, L.L.; Melack, J.M.; Affonso, A.G.; Barbosa, C.; Gastil-Buhl, M.; Novo, E.M. Wetlands of the lowland Amazon basin: Extent, vegetative cover, and dual-season inundated area as mapped with JERS-1 synthetic aperture radar. *Wetlands* **2015**, *35*, 745–756. [CrossRef]
20. Xie, C.; Shao, Y.; Xu, J.; Wan, Z.; Fang, L. Analysis of ALOS PALSAR InSAR data for mapping water level changes in Yellow River Delta wetlands. *Int. J. Remote Sens.* **2013**, *34*, 2047–2056. [CrossRef]
21. Bendib, A. High-resolution Alos Palsar for the Characterization of Water Storage at the Fountaine Des Gazelles Dam in Biskra, Eastern Algeria. *J. Indian Soc. Remote Sens.* **2021**, *49*, 1927–1938. [CrossRef]
22. Plank, S.; Jüssi, M.; Martinis, S.; Twele, A. Mapping of flooded vegetation by means of polarimetric Sentinel-1 and ALOS-2/PALSAR-2 imagery. *Int. J. Remote Sens.* **2017**, *38*, 3831–3850. [CrossRef]
23. Gu, X.; Zeng, Q.; Shen, H.; Chen, E.; Zhao, L.; Yu, F.; Tu, K. Study on water information extraction using domestic GF-3 image. *Natl. Remote Sens. Bull.* **2021**, *23*, 555–566. [CrossRef]
24. Liu, X.; Liu, L.; Shao, Y.; Zhao, Q.; Zhang, Q.; Lou, L. Water Detection in Urban Areas from GF-3. *Sensors* **2018**, *18*, 1299. [CrossRef]
25. Bioresita, F.; Puissant, A.; Stumpf, A.; Malet, J.-P. A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery. *Remote Sens.* **2018**, *10*, 217. [CrossRef]
26. Chen, Z.; Zhao, S. Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *113*, 103010. [CrossRef]
27. Wieland, M.; Fichtner, F.; Martinis, S.; Groth, S.; Krullikowski, C.; Plank, S.; Motagh, M. S1S2-Water: A global dataset for semantic segmentation of water bodies from Sentinel-1 and Sentinel-2 satellite images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *17*, 1084–1099. [CrossRef]

28. Huang, W.; DeVries, B.; Huang, C.; Lang, M.W.; Jones, J.W.; Creed, I.F.; Carroll, M.L. Automated extraction of surface water extent from Sentinel-1 data. *Remote Sens.* **2018**, *10*, 797. [CrossRef]
29. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]
30. Hertel, V.; Chow, C.; Wani, O.; Wieland, M.; Martinis, S. Probabilistic SAR-based water segmentation with adapted Bayesian convolutional neural network. *Remote Sens. Environ.* **2023**, *285*, 113388. [CrossRef]
31. Ren, Y.; Li, X.; Yang, X.; Xu, H. Development of a dual-attention U-Net model for sea ice and open water classification on SAR images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [CrossRef]
32. Pech-May, F.; Aquino-Santos, R.; Delgadillo-Partida, J. Sentinel-1 SAR Images and Deep Learning for Water Body Mapping. *Remote Sens.* **2023**, *15*, 3009. [CrossRef]
33. Jia, S.; Xue, D.; Li, C.; Zheng, J.; Li, W. Research on Water Body Information Extraction Method Based on Sentinel-1 Data. *Yangtze River* **2019**, *50*, 213–217. [CrossRef]
34. Du, Q.; Li, G.; Chen, D.; Qi, S.; Zhou, Y.; Wang, F.; Cao, Y. Extracting water body data based on SDWI and threshold segmentation: A case study in permafrost area surrounding Salt Lake in Hoh Xil, Qinghai-Tibet Plateau, China. *Res. Cold Arid. Reg.* **2023**, *15*, 202–209. [CrossRef]
35. Wang, Z.; Xie, F.; Ling, F.; Du, Y. Monitoring Surface Water Inundation of Poyang Lake and Dongting Lake in China Using Sentinel-1 SAR Images. *Remote Sens.* **2022**, *14*, 3473. [CrossRef]
36. Vickers, H.; Malnes, E.; Høgda, K.-A. Long-Term Water Surface Area Monitoring and Derived Water Level Using Synthetic Aperture Radar (SAR) at Altevatn, a Medium-Sized Arctic Lake. *Remote Sens.* **2019**, *11*, 2780. [CrossRef]
37. Kseňak, Ľ.; Pukanská, K.; Bartoš, K.; Blišťan, P. Assessment of the Usability of SAR and Optical Satellite Data for Monitoring Spatio-Temporal Changes in Surface Water: Bodrog River Case Study. *Water* **2022**, *14*, 299. [CrossRef]
38. Sun, Q.; Li, J. A method for extracting small water bodies based on DEM and remote sensing images. *Sci. Rep.* **2024**, *14*, 760. [CrossRef]
39. Bao, L.; Lv, X.; Yao, J. Water Extraction in SAR Images Using Features Analysis and Dual-Threshold Graph Cut Model. *Remote Sens.* **2021**, *13*, 3465. [CrossRef]
40. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [CrossRef]
41. Yang, S.; Wang, L.; Yuan, Y.; Fan, L.; Wu, Y.; Sun, W.; Yang, G. Recognition of small water bodies under complex terrain based on SAR and optical image fusion algorithm. *Sci. Total Environ.* **2024**, *946*, 174329. [CrossRef]
42. Shapingba: Flood Prevention and Drought Resistance in Action, Fully Managing Risks, Reservoirs and Ensuring Water Supply. Available online: https://nyncw.cq.gov.cn/zwxx_161/qxlb/202305/t20230525_11996324.html (accessed on 20 August 2025).
43. Chongqing Reservoir Collapse Pit: There Have Been Multiple Collapses in the Surrounding Area, and the Local Government Has Invested Nearly Ten Million Yuan to Reinforce the Dam and Eliminate Risks. Available online: https://news.qq.com/rain/a/20241106A0A67400 (accessed on 20 August 2025).

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