

1 **Beyond clouds: Seamless flood mapping using Harmonized Landsat and**
2 **Sentinel-2 time series imagery and water occurrence data**

3 Zhiwei Li^{a,b}, Shaofen Xu^a, Qihaoweng^{a,b,c,*}

4 ^a Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University,
5 Hung Hom, Hong Kong

6 ^b Research Centre for Artificial Intelligence in Geomatics, The Hong Kong Polytechnic University,
7 Hung Hom, Hong Kong

8 ^c Research Institute for Land and Space, The Hong Kong Polytechnic University, Hung Hom,
9 Hong Kong

10 * Corresponding author. E-mail address: qihao.weng@polyu.edu.hk

11 * Corresponding author. E-mail address: qihao.weng@polyu.edu.hk

13 **Abstract**

14 Floods are among the most devastating natural disasters, posing significant risks to life, property,
15 and infrastructure globally. Earth observation satellites provide data for continuous and extensive
16 flood monitoring, yet limitations exist in the spatial completeness of monitoring using optical
17 images due to cloud cover. Recent studies have developed gap-filling methods for reconstructing
18 cloud-covered areas in water maps. However, these methods are not tailored for and validated in
19 cloudy and rainy flooding scenarios with rapid water extent changes and limited clear-sky
20 observations, leaving room for further improvements. This study investigated and developed a
21 novel reconstruction method for time series flood extent mapping, supporting spatially seamless
22 monitoring of flood extents. The proposed method first identified surface water from time series
23 images using a fine-tuned large foundation model. Then, the cloud-covered areas in the water
24 maps were reconstructed, adhering to the introduced submaximal stability assumption, on the
25 basis of the prior water occurrence data in the Global Surface Water dataset. The reconstructed
26 time series water maps were refined through spatiotemporal Markov random field modeling for
27 the final delineation of flooding areas. The effectiveness of the proposed method was evaluated
28 with Harmonized Landsat and Sentinel-2 datasets under varying cloud cover conditions, enabling
29 seamless flood mapping at 2–3-day frequency and 30 m resolution. Experiments at four global
30 sites confirmed the superiority of the proposed method. It achieved higher reconstruction accuracy

31 with average F1-scores of 0.931 during floods and 0.903 before/after floods, outperforming the
32 typical gap-filling method with average F1-scores of 0.871 and 0.772, respectively. Additionally,
33 the maximum flood extent maps and flood duration maps, which were composed on the basis of
34 the reconstructed water maps, were more accurate than those using the original cloud-
35 contaminated water maps. The benefits of synthetic aperture radar images (e.g., Sentinel-1) for
36 enhancing flood mapping under cloud cover conditions were also discussed. The method proposed
37 in this paper provided an effective way for flood monitoring in cloudy and rainy scenarios,
38 supporting emergency response and disaster management. The datasets and experimental results
39 will be available online (<https://github.com/dr-lizhiwei/SeamlessFloodMapper>).

40
41 **Keywords:** Flood extent, Water body extraction, Cloud removal, HLS imagery, Cloudy and rainy
42 environments, Time series, Satellite mapping
43

44 **1. Introduction**

45 As one of the most devastating natural disasters, floods exacerbated by global climate change
46 pose significant risks to life, property, and infrastructure globally (Rentschler et al., 2022; IPCC,
47 2023). Recent estimates suggest that the proportion of the global population affected by floods has
48 increased by 20%–24% since the turn of the century, with projections indicating that up to 758
49 million people will be exposed to a 100-year flood event by 2030 (Tellman et al., 2021). Enhanced
50 flood monitoring techniques, specifically focusing on near-real-time flood extent mapping, are
51 critically important for effective emergency response and flood disaster management, to deal with
52 the growing threat of flood disasters efficiently.

53 Satellite remote sensing offers a cost-effective means for continuous and extensive flood
54 monitoring. Optical and synthetic aperture radar (SAR) images, the primary data sources in
55 satellite-based flood monitoring, serve distinct roles given their unique characteristics. On the one
56 hand, SAR data have generally been widely used in flood monitoring because of their
57 independence from weather conditions and time of imaging. However, because of the limitations
58 of SAR imaging mechanisms, flood monitoring with SAR data in urban areas is relatively
59 constrained (Notti et al., 2018; Liang and Liu, 2020; McCormack et al., 2022). Additionally, the
60 publicly available sources of SAR data are relatively limited, and the use of single-source SAR

61 data for flood monitoring is restricted by the revisit intervals of satellites. For example, the
62 combined Sentinel-1A/1B satellites provide SAR data with a temporal resolution of 6 days, which
63 makes capturing short-term flood events, lasting less than 6 days, challenging. Such a situation
64 became even worse due to the operational failure of Sentinel-1B in December 2021, which led to
65 the temporal resolution of Sentinel-1 being reduced to 12 days. On the other hand, numerous
66 optical satellites, such as Landsat and Sentinel-2, can provide open-access images for surface
67 information extraction and flood mapping. However, they are inevitably affected by cloud cover,
68 leading to the missing ground surface information in images and reducing the frequency of valid
69 observations (Yang et al., 2020; Zeng et al., 2020; Goffi et al., 2020; Z. Li et al., 2022; Shastry et
70 al., 2023). In consideration of the revisit cycles of satellites and the duration of floods with their
71 respective intervals, the frequency of satellite observation is crucial for flood monitoring (Tulbure
72 et al., 2022).

73 **Flood extent mapping with satellite imagery:** Mapping the extent of flood inundated areas using
74 remotely sensed satellite imagery relies on water extraction. Existing water extraction and flood extent
75 mapping methods primarily consist of water index- and machine learning-based methods. Water
76 indexes, such as the normalized difference water index (NDWI) (McFeeters, 1996), modified NDWI
77 (Xu, 2005), and automated water extraction index (Feyisa et al., 2014), are widely used to separate
78 surface water from other ground objects because they are easy to use and computationally efficient.
79 Traditional machine learning methods, such as support vector machine (Li and Narayanan, 2003; Sun
80 et al., 2014, 2015) and random forest (Deng et al., 2017; Wang et al., 2018), are more effective in
81 water extraction than index-based methods. However, they require manual extraction of spectral and
82 spatial features, impacting prediction accuracy. The recent introduction of deep learning techniques,
83 such as convolutional neural networks (CNN), a subset of machine learning methods, has
84 revolutionized this field and significantly enhanced the accuracy of water extraction owing to their
85 strong feature representation ability (Wang et al., 2020; Konapala et al., 2021; Bentivoglio et al.,
86 2022; He et al., 2024; Valman et al., 2024). Despite the development of numerous water extraction
87 and flood mapping methods, the negative influence of cloud cover was often avoided in previous
88 studies only by using cloud-free optical imagery for flood monitoring (Benoudjit and Guida, 2019).
89 Alternatively, only cloud-free optical satellite images were used to interpret land cover types
90 before and during flood events for flood impact assessments (Huang and Jin, 2020; Psomiadis et

91 al., 2020). Therefore, exploring the potential of using optical satellite imagery that might be cloud-
92 covered in flood mapping to improve flood monitoring frequency is worth further investigation
93 (DeVries et al., 2020; S. Li et al., 2022). Developing a cloud reconstruction method for
94 multisensor optical satellite imagery in cloudy scenarios is essential for seamless flood extent
95 mapping at high temporal density.

96 **Cloud removal for seamless flood mapping:** Seamless water mapping with cloudy optical
97 images is key for continuous flood extent mapping. Existing methods can achieve cloud removal
98 by integrating complementary information from neighboring temporal images into the cloud-
99 covered areas of target images after transformations (Shen et al., 2015; Zhang et al., 2021; Li et al.,
100 2024). However, accurately reconstructing the spatial details of ground surfaces in cloud-covered
101 areas, especially in images with land cover changes, such as flooding, is challenging. A cloud
102 removal method for optical satellite images, targeting the application scenarios for flood
103 monitoring, is particularly necessary to leverage multisensor satellite images for flood monitoring
104 fully. In this case, gap-filling methods for the reconstruction of cloud-contaminated water areas,
105 including methods based on ancillary data (e.g., inundation frequency) (Zhao and Gao, 2018;
106 Mullen et al., 2021), spatiotemporal neighborhood similarity (Li et al., 2021; Bai et al., 2022;
107 Huang et al., 2023), and spatiotemporal modeling (M. Li et al., 2022; Bai et al., 2023), have been
108 studied recently. Although these recent studies have developed gap-filling methods for seamless
109 water mapping, such methods are not tailored for and validated in flooding scenarios, leaving
110 room for further improvements. First, existing studies of index-based water extraction methods
111 require manual feature extraction and threshold determination for water segmentation, which
112 shows both sensitivity and limitations in accuracy. In contrast, deep learning methods, particularly
113 recent large foundation models such as RVSA (Wang et al., 2023), RemoteCLIP (Liu et al.,
114 2023), and Prithvi-100M (Jakubik et al., 2023), have significantly enhanced image processing
115 accuracy and offered potential for more precise water extraction. Second, existing gap-filling
116 methods can effectively reconstruct permanent or seasonal surface water bodies covered by clouds
117 (Mullen et al., 2021; Bai et al., 2022; Huang et al., 2023). However, they fall short in flooding
118 scenarios with significant surface water changes, and their performance is limited by the number
119 of valid images during the flooding periods. Third, water occurrence data generated based on
120 historical satellite observations show large uncertainties at low occurrence values and often ignore

121 low-frequency water dynamics like floods. This oversight and the lack of confidence-level
122 consideration in water occurrence make it a big challenge for a direct application of existing
123 methods ([Zhao and Gao, 2018](#); [Li et al., 2021](#)) for seamless flood mapping. Additionally, locally
124 adaptive threshold determination for water segmentation based on water occurrence data warrants
125 further consideration for accurate reconstruction under complex cloudy and rainy weather
126 conditions. In practical flood extent mapping, the ability to obtain near-real-time inundation maps
127 is crucial for an effective emergency response to flood disasters ([Notti et al., 2018](#); [DeVries et al.,](#)
128 [2020](#); [Tulbure et al., 2022](#)). In this context, ensuring the spatiotemporal consistency of time series
129 flood mapping results obtained based on multisensor satellite data and improving the precision of
130 dynamic flood monitoring are additional issues to be considered.

131 To overcome these limitations in flood extent mapping with optical images under varying cloud
132 cover conditions, this study introduces a robust seamless time series flood extent mapping method.
133 This method fine-tunes a large foundation model to achieve high-precision water extraction,
134 enables the reconstruction of cloud-covered flooding areas in optical satellite time series imagery,
135 and conducts time series refinement via spatiotemporal modeling to improve the spatiotemporal
136 consistency of water maps. We utilize Harmonized Landsat and Sentinel-2 (HLS) datasets
137 ([Claverie et al., 2018](#)), which have been proven promising for high-temporal-density flood
138 mapping ([Tulbure et al., 2022](#)), as our experimental data. HLS datasets, which harmonize images
139 captured by the Landsat-8/9 and Sentinel-2A/B satellites, offer image time series typically at 2–3
140 day intervals with 30 m resolution. Through the validation of our developed seamless flood extent
141 mapping method under varying conditions at four global sites, this study aims to answer whether
142 cloud-covered flooding areas can be effectively reconstructed in HLS image time series. The key
143 contributions of this study are summarized as follows:

144 We investigate the feasibility of reconstructing cloud-contaminated flooding areas by
145 introducing a method that facilitates spatially continuous flood monitoring with cloudy time series
146 imagery. Utilizing this approach, we can generate a time series of seamless flood extent maps,
147 significantly mitigating the impacts of clouds on flood mapping. Additionally, we demonstrate the
148 potential of this reconstruction method for seamless time series flood mapping by quantitatively
149 evaluating its effectiveness under varying cloud cover conditions and at different phases of flood
150 events. The evaluation results affirm the superiority of the proposed method in various scenarios.

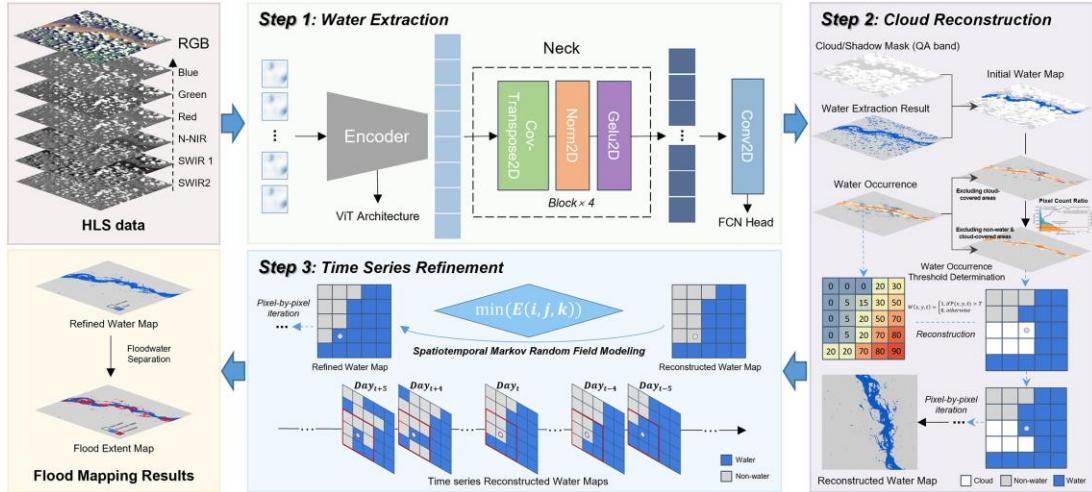
151 The advantages of combining Landsat and Sentinel-2 images for flood mapping are confirmed.
152 The benefits of further integrating SAR images (e.g., Sentinel-1) to augment flood mapping are
153 also discussed, which we consider a promising new paradigm for extensive application in this
154 field.

155 By leveraging a combined time series of Landsat and Sentinel-2 images, this study aims to
156 enhance the completeness, frequency, and precision of flood extent mapping in cloudy and rainy
157 conditions, thereby improving emergency response and contributing to effective flood
158 management.

159 The remaining contents are organized as follows. [Section 2](#) introduces the proposed method and
160 provides the implementation details. The experimental data and results are described in [Section 3](#).
161 In [Section 4](#), we highlight the benefits of cloud reconstruction for flood mapping and the
162 incorporation of SAR images for enhanced flooding area reconstruction, and then discuss the
163 efficiency and limitations of the proposed method. Our conclusions are drawn in [Section 5](#).

164 **2. Methodology**

165 This study proposes to achieve seamless flood extent mapping at 30 m resolution based on HLS
166 imagery to improve the completeness, frequency, and precision of flood extent mapping,
167 especially in cloud-prone areas, and to benefit flood emergency response and management. The
168 flowchart of the proposed method is illustrated in [Fig. 1](#). Three main steps are involved.
169 Specifically, in order to achieve optimal performance, we first fine-tuned a large foundation
170 model to enable accurate identification of water extents from HLS images. Then, auxiliary
171 water occurrence data from the Global Surface Water (GSW) dataset ([Pekel et al., 2016](#)) were
172 utilized to reconstruct cloud- and cloud-shadow-covered areas in the time series water maps based
173 on the introduced submaximal stability assumption. The cloud/shadow mask used was defined by
174 the Quality Assessment (QA) band included in the HLS dataset, which was generated using the
175 Fmask algorithm ([Zhu et al., 2015](#)). Lastly, we conducted spatiotemporal modeling to minimize
176 potential errors in water extraction and cloud reconstruction, consequently enhancing the
177 spatiotemporal consistency of time series water maps and ensuring more accurate identification of
178 flood extents.



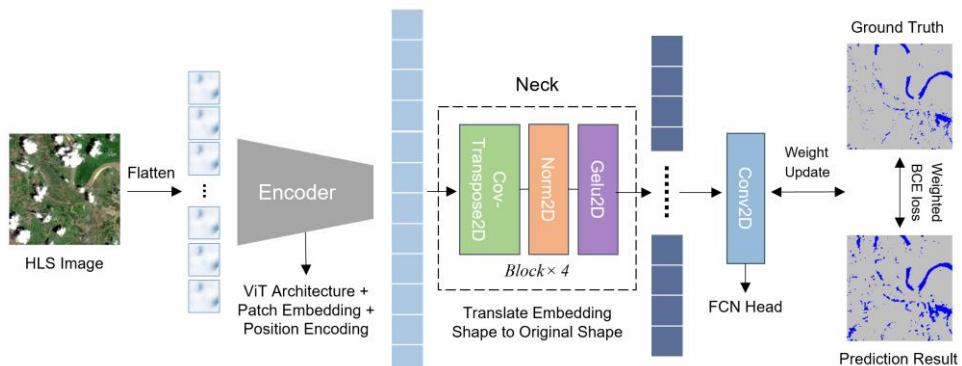
179
180 **Fig. 1.** Flowchart of the proposed seamless time series flood extent mapping method. The
181 proposed method involves three main steps. Step 1: water extraction for each HLS image via a
182 fine-tuned large foundation model; Step 2: reconstruction of cloud- and cloud-shadow-covered
183 areas, as defined by the QA band of the HLS image, in each water map; Step 3: refining time
184 series water maps via the spatiotemporal Markov random field (MRF) modeling.

185 **2.1 Water extraction via a fine-tuned large foundation model**

186 In this step, we harnessed the potential of a state-of-the-art deep learning model for water
187 extraction. Specifically, we utilized the large foundation model Prithvi-100M ([Jakubik et al., 2023](#))
188 as the baseline, which was constructed via the temporal Vision Transformer (ViT)
189 ([Dosovitskiy et al., 2021](#)) and pre-trained with a vast amount of harmonized Landsat-8/9 and
190 Sentinel-2 images in HLS datasets by the NASA and IBM team ([Jakubik et al., 2023](#)). HLS
191 datasets encompassed global coverage and yielded 30 m surface reflectance products. The baseline
192 foundation model, pre-trained with HLS images from the contiguous United States, served as an
193 effective feature extractor and was extended for water extraction in Landsat-8/9 and Sentinel-2
194 images included in HLS datasets. Highly accurate water maps are expected to be obtained by fine-
195 tuning the foundation model with existing labeled water and flood datasets featuring Landsat and
196 Sentinel-2 images.

197 The Prithvi-100M model adopted a self-supervised encoder based on ViT and employed a
198 learning strategy involving the masked autoencoder (MAE) ([He et al., 2022](#)). It utilized the mean
199 squared error loss function to guide model training. Operating on 3D patch data, the model was
200 allowed to receive the input data comprising multi-temporal images. For each batch of data, the
201 Query-Key-Value attention mechanism was introduced, enabling the model to capture both
202 temporal and spatial features ([Vaswani et al., 2017](#)). The images used for model training

203 comprised six bands: blue, green, red, narrow NIR, SWIR 1, and SWIR 2. This study concentrated
 204 on the task of flood segmentation for a mono-temporal image, utilizing the Prithvi-100M-
 205 Sen1Floods11 model architecture ([Jakubik et al., 2023](#)). This architecture was pre-trained on the
 206 Sen1Floods11 dataset ([Bonafilia et al., 2020](#)) for flood mapping. The model architecture is
 207 illustrated in Fig. 2. Starting with a multispectral HLS image of dimensions $M \times N$, the model
 208 normalized the six required bands based on their corresponding mean value and variance of
 209 images in the training dataset, and performed operations such as cropping and flattening. In the
 210 encoding phase, the ViT encoder was utilized to encode the data, generating patch embeddings
 211 with positional encoding. These embeddings underwent multiple convolutions for feature
 212 extraction, and the shape of the embeddings was then transformed back to the original image size.
 213 In the decoding phase, a fully convolutional network (FCN), applying the binary cross entropy
 214 loss function, was used to implement the identification of flood-specific pixels.



215
 216 **Fig. 2.** Architecture of the Prithvi-100M-Sen1Floods11 model for flood mapping (Adapted from
 217 [Jakubik et al., 2023](#)).

218 To achieve optimal performance in water extraction, we fine-tuned the pre-trained Prithvi-
 219 100M-Sen1Floods11 model. During the process, we monitored the fine-tuning and evaluated the
 220 model's performance to determine the optimal model. Consequently, we selected the model with
 221 the highest overall accuracy on the Sen1Floods11 test dataset as the final model for subsequent
 222 water extraction. With the fine-tuned model, time series water maps could be obtained on the basis
 223 of the harmonized Landsat-8/9 and Sentinel-2 images. Cloud-covered areas in these images,
 224 including the clouds and the cloud shadows labeled in the quality assessment (QA) band of the HLS
 225 products, were also labeled in the time series water maps.

226 **2.2 Reconstructing cloud-covered areas in time series water maps**

227 To reconstruct cloud-covered areas in time series water maps, we introduced an approach that

leveraged the water occurrence data included in the GSW dataset (Pekel et al., 2016). The water occurrence data, in which each pixel denotes the frequency of water presence, were crafted by aggregating monthly water coverage data using historical Landsat imagery spanning from 1984 to 2021. The assumption is that in a local area, if most cloud-free pixels possessing a certain water occurrence value are identified as water, adjacent cloud-contaminated pixels with water occurrence exceeding this value are likely to be water (Zhao and Gao, 2018; Mullen et al., 2021). This assumption strengthens as the proximity between cloud-free and cloud-contaminated pixels increases.

In addition, considering the biases inherent in the water occurrence data, to determine optimal water occurrence thresholds with high confidence automatically for the robust reconstruction of cloud-contaminated pixels under varying cloud cover conditions and water dynamics, we introduced a “**submaximal stability**” assumption. This assumption was based on our observations, which guided the threshold determination. The three observations are as follows: 1) water extents under regular conditions usually fall within the range of maximum water extents; 2) water dynamics often occur in areas with low water occurrence; and 3) random cloud cover has a greater impact on the occurrence calculation for surface water with lower occurrence than those for surface water with higher occurrence, thus indicating a trend in which low water occurrence implies low confidence. From these observations, the greatest overlap between regular and maximum water extents occurs in areas with high occurrence. Furthermore, the ratio of pixel counts in histogram bins for water occurrence within regular and maximum extents generally tends to increase across the occurrence range and can generally indicate the confidence level of water occurrence. Therefore, optimal water occurrence thresholds exceeding certain confidence levels can be determined on the basis of this generally increasing ratio and robust trend for the reconstruction of cloud-contaminated water pixels.

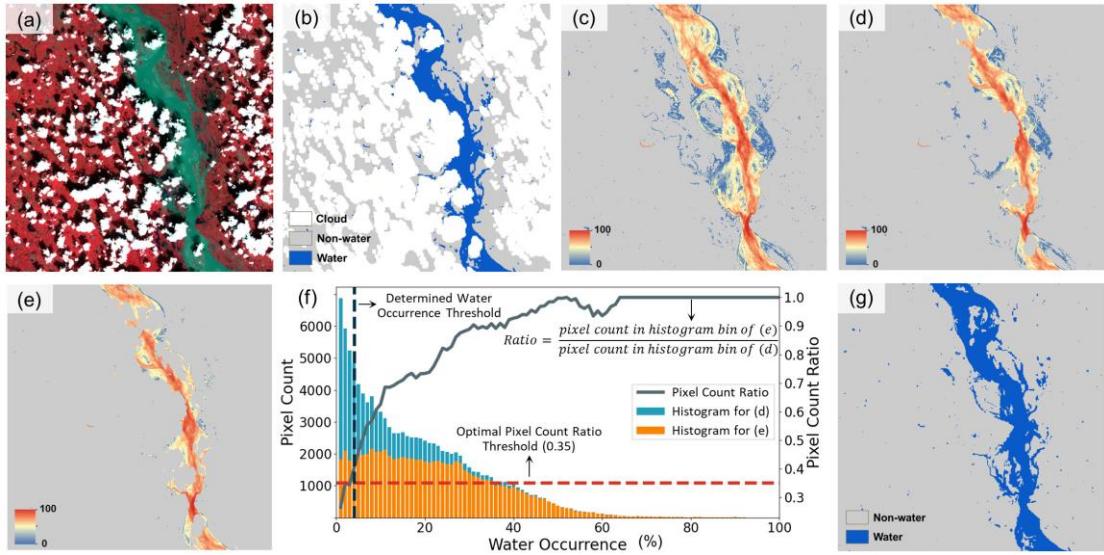
Consequently, we implemented the reconstruction of cloud-covered areas in time-series water maps over a local sliding window, iteratively in a pixel-by-pixel and image-by-image manner. All cloud-contaminated pixels were assigned a new class: 1 for water or 0 for non-water. The reconstruction process was applied only to cloud-contaminated pixels in the water maps, while the identified water and non-water pixels remained unchanged. The reconstructed binary water maps $W(x, y, t)$ can be formalized as follows:

258

$$W(x, y, t) = \begin{cases} 1, & \text{if } P(x, y, t) > T \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

259 where $P(x, y, t)$ denotes the water occurrence of pixel (x, y) at date t . T is the determined
260 water occurrence threshold over the sliding local window centered around the target pixel.

261 3The local window size was adaptively determined based on the cloud cover conditions within
262 the window. To ensure that enough valid pixels were involved for reconstruction, the initial
263 window size was set at 50×50 and was stepwise increased if the cloud-free pixels within the
264 window were not predominantly water. If the window size reached the maximum height or width
265 of the image, it was reset to the entire image size. 3 illustrates the detailed steps for the
266 reconstruction of cloud-covered areas in an example water map. The determination of the water
267 occurrence threshold T , which was based on the introduced “submaximal stability” assumption,
268 involved three steps. First, the histogram of water occurrence excluding cloud-covered areas
269 shown in 3(d) and the histogram of water occurrence excluding non-water and cloud-covered
270 areas shown in 3(e) were computed. Then, the pixel count ratio, which is defined as the ratio of the
271 pixel count in the histogram bin of 3(e) to that of 3(d), was obtained. This ratio indicates the
272 proportion of observed water occurrence pixel count to the actual water occurrence pixel count
273 and can reflect the confidence level of the water occurrence. The introduced “submaximal
274 stability” assumption was also applicable to the water occurrence excluding cloud-covered areas
275 in 3(e) and 33(d) because of the random distribution of clouds in images. Finally, the water
276 occurrence threshold T was determined as the water occurrence corresponding to the first
277 histogram bin in the range of [0,100], in which the pixel count ratio equaled or exceeded the
278 optimal threshold of 0.35. The optimal pixel count ratio threshold was carefully selected by
279 parameter sensitivity analysis, in which iterative tests were conducted for the pixel count ratio
280 threshold within a typical range of [0, 0.6] at an interval of 0.01. The optimal threshold was
281 determined on the basis of the reconstruction accuracy assessments through a series of simulation
282 experiments, as described in Section 3.3.2. Once the water occurrence threshold within the local
283 window was determined, if the water occurrence value of a cloud-contaminated target pixel
284 exceeded this threshold, then the pixel was classified as water; otherwise, it was classified as non-
285 water, as shown in 3(g). This process continued until all cloud-covered pixels in each cloud-
286 covered water map were reconstructed.



287

288 **Fig. 4.** Illustration of reconstructing cloud-covered areas in the water map. (a) Cloudy HLS image;
 289 (b) Initial cloud-covered water map; (c) Water occurrence in the GSW dataset; (d) Water occurrence
 290 excluding cloud-covered areas labeled in (b); (e) Water occurrence excluding non-water and cloud-
 291 covered areas labeled in (b); (f) Water occurrence histograms for (d) and (e) and determination of
 292 the water occurrence threshold. The threshold is determined as the water occurrence corresponding
 293 to the first histogram bin in the range of [0,100], in which the pixel count ratio equals or exceeds
 294 the optimal threshold of 0.35; (g) Reconstructed binary water map.

295 The proposed reconstruction approach offered two primary advantages: 1) It used locally
 296 adaptive window sizes that accounted for the spatial proximity between local areas, enhancing the
 297 accuracy and detail of reconstruction—especially beneficial for small water bodies. It also
 298 accommodated the heterogeneity in flooding area changes across different regions, including
 299 events such as dam breaches, in which the surface water in reservoirs may shrink instead of
 300 expanding like flooded areas. 2) The method employed the pixel count ratio to determine the water
 301 occurrence threshold on the basis of the introduced “submaximal stability” assumption, which
 302 offered greater stability compared with using direct pixel counts. This stability ensured that
 303 variations in the water occurrence histogram, which may arise from differing cloud cover
 304 conditions and water dynamics, leading to variable pixel counts, do not affect the threshold
 305 determination.

306 2.3 Time series refinement of water maps via spatiotemporal MRF modeling

307 Inconsistencies may exist in reconstructed time series water maps owing to differences among
 308 images captured under varying cloud cover conditions. In this study, spatiotemporal MRF
 309 modeling was employed to refine the reconstructed time series water maps and enhance their

310 spatiotemporal consistency. Specifically, each pixel in time series water maps belonged to only
 311 two possible classes, i.e., water and non-water, denoted as 1 and 0, respectively. Given that water
 312 areas are likely to be connected patches and that the water body at a given pixel is likely to persist
 313 for certain periods even during flood events, the time series water maps were assumed to satisfy
 314 the MRF properties with a Gibbs potential function in spatial and temporal domains ([Moussouris,](#)
 315 [1974; Li, 2009; Kasetkasem et al., 2014](#)). The Gibbs potential function defined the energy of a
 316 configuration, i.e., the sum of the potential (or cost) for spatial and temporal local neighborhoods.
 317 Thus, a spatiotemporal MRF model was constructed to minimize an energy function to find the
 318 most likely state configuration, indicating the classes of water and non-water for each pixel. The
 319 energy function $E(i, j, k)$ for a given target pixel at location (i, j) and date k can be formalized
 320 as follows:

$$321 \quad E(i, j, k) = \sum_{s \in S} \sum_{t \in T} w_{s,t} \cdot I(x_{s,t} \neq x_{i,j,k}) \quad (2)$$

322 where $I(\cdot)$ checks whether the class of pixels in spatial ($s \in S$) and temporal ($t \in T$)
 323 neighborhoods is different from the center pixel $x_{i,j,k}$. The weight $w_{s,t}$ is determined on the basis
 324 of spatiotemporal proximity to the center pixel, in which a smaller spatiotemporal proximity leads
 325 to a larger weight, defined as follows:

$$326 \quad w_{s,t} = \gamma \frac{\frac{1}{D_s}}{\sum_{s \in S} \frac{1}{D_s}} + \beta \frac{\frac{1}{D_t}}{\sum_{t \in T} \frac{1}{D_t}} \quad (3)$$

327 where γ and β are the balanced parameters applied to the spatial and temporal terms,
 328 respectively. D_s is the 2D Euclidean distance from the pixel in neighborhood S to the center
 329 pixel $x_{i,j,k}$. D_t refers to the temporal distance from the pixel in adjacent timeslots T to the target
 330 date k . In this study, the size of neighborhood S is empirically set to 3×3 pixels to preserve
 331 spatial detail efficiently; the time range of adjacent timeslots T is set to be $[k - 5, k + 5]$, which
 332 ensures that only timely information is involved during reconstruction, considering the typical
 333 duration of flood events.

334 A smaller weight of 0.75 was specifically assigned to the reconstructed pixel compared with
 335 that for the clear pixel during the energy computation process to mitigate the uncertainty in
 336 refinement due to potential reconstruction errors. The optimal classes (i.e., water and non-water)
 337 were assigned to each pixel in the reconstructed time series water maps, determined by achieving

338 the minimum energy state calculated using Equations (2) and (3). Refined time series water maps
339 were generated by applying the refinement process to each pixel accordingly.

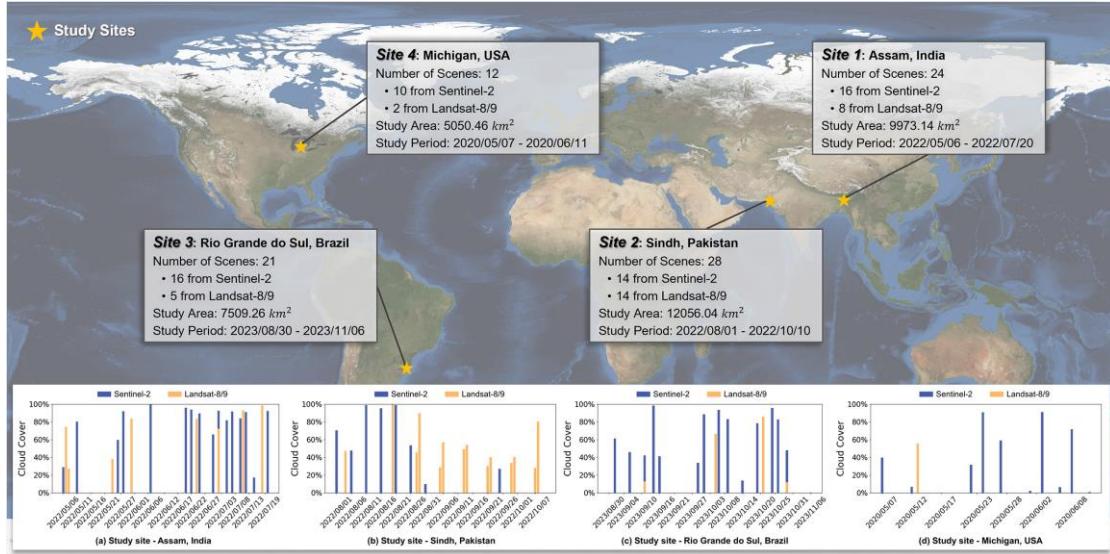
340 Flood inundated areas were identified by excluding pre-flood water from the refined time series
341 water maps. In this study, the maximum water extent prior to a flood was determined from the
342 time series water maps. Water within this extent was labeled as pre-flood water on the flood extent
343 map, while water outside this extent was labeled as floodwater. Consequently, three classes,
344 namely, non-water, pre-flood water, and flood water, were labeled on the final time series flood
345 extent maps.

346 **3. Results and analysis**

347 **3.1 Experimental data**

348 In this study, four global study sites located in four different countries were selected. The
349 geographical locations of these sites, along with the corresponding number of harmonized Sentinel-2
350 and Landsat-8/9 scenes for experiments, are shown in Fig. 5. This figure also offers a comprehensive
351 overview of the cloud percentages derived from cloud and cloud shadow flags of each scene in HLS
352 datasets within the study periods.

353 Among the four study sites, Site 1 is situated in Assam, India, known for having one of the highest
354 frequencies of flood occurrences globally (Rentschler et al., 2022). The selected flood event occurred
355 between May and August 2022, extensively affecting Assam, India and the Sylhet region of
356 Bangladesh. Site 2 is located in Sindh, Pakistan, which experienced a flood event in August and
357 September 2022. This event was notable for its widespread impact, significant consequences, high
358 satellite revisit frequency, and minimal cloud cover during the flooding, making it a focus for flood
359 mapping study (Tulbure et al., 2022). Site 3 is located in Rio Grande do Sul, Brazil, which
360 experienced a flood in October 2023 due to intense rainfall causing river overflow. This event
361 featured heavy cloud cover, low satellite revisit frequency, and limited available observational
362 information, marking its research significance. Site 4 is located in Michigan, USA, where a dam
363 failure flood occurred on May 19, 2020. Unlike the other three events, the dam failure caused
364 simultaneous expansion and shrinking of the water extent, making the reconstruction of cloud-
365 covered water areas challenging.



366

367 **Fig. 5.** Locations of four global study sites and cloud cover percentage distribution of HLS image
 368 time series for experiments.

369 **3.2 Accuracy evaluation of the large foundation model for water extraction**

370 The Prithvi-100M-Sen1Floods11 model ([Jakubik et al., 2023](#)), fine-tuned on the Sen1Floods11
 371 dataset ([Bonafilia et al., 2020](#)), was used for water extraction in HLS images. Among the 446
 372 labeled 512×512 chips in the Sen1Floods11 dataset, encompassing all 14 biomes and spanning 11
 373 flood events, 267 chips were used for model fine-tuning, 89 chips for model validation, and the
 374 remaining 90 chips for model testing. The performance of the pre-trained model and the fine-tuned
 375 model for water extraction was evaluated in terms of five metrics, namely, overall accuracy (OA),
 376 precision, recall, mean intersect over union (mIoU), and F1-score. The accuracy comparison
 377 results are provided in [Table 1](#). The results showed an increase of 1.08% in precision and a
 378 decrease of 0.56% in the recall, which resulted in an overall improvement of the fine-tuned
 379 model in OA with increased training epochs compared to the pre-trained model across test images
 380 in the Sen1Floods11 dataset. Note that the commission errors in water maps are more serious than
 381 the omission errors, as the latter can potentially be recovered during the reconstruction process of
 382 the proposed method, while the former might lead to more severe errors. In this case, the fine-
 383 tuned model is more appropriate than the pre-trained model. The 97.35% OA indicates the
 384 promising capacity of the fine-tuned model for flood extent mapping.

385 **Table 1.** Accuracy comparison of pre-trained and fine-tuned models for water extraction with test
 386 images in the Sen1Floods11 dataset.

	OA	Precision	Recall	mIoU	F1-Score
--	----	-----------	--------	------	----------

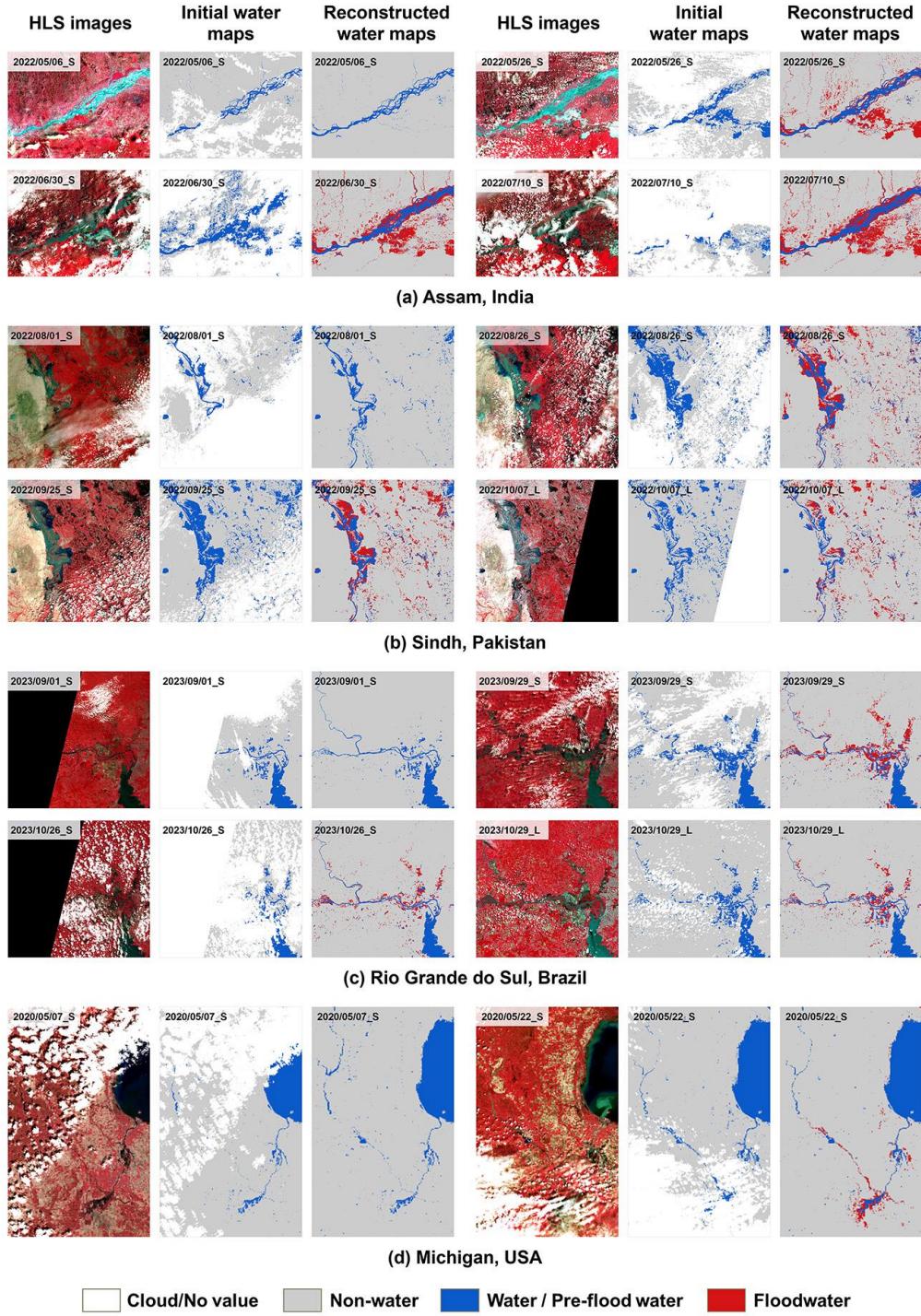
Pre-trained model	97.25%	87.85%	90.54%	0.805	0.892
Fine-tuned model	97.35%	88.93%	89.98%	0.809	0.895

387 **3.3 Comparison of reconstructed flood maps and accuracy assessment**

388 **3.3.1 Time series flood extent mapping with the proposed reconstruction method**

389 Fig. 6 shows typical seamless time series flood mapping results at four study sites generated by the
 390 proposed approach, and the complete flood mapping results are presented in Figs. S1–S4 as
 391 supplementary materials. Pseudocolored HLS images, initial water maps, and reconstructed flood
 392 maps are provided for each study site. From the pseudocolored images and initial water maps, the
 393 cloud cover during the flood period was significant, and valid flood observation data were sparse and
 394 limited. HLS images partly covering study areas were also used for experiments to enhance flood
 395 mapping frequency. As a result, the average satellite observation intervals for the four study sites were
 396 significantly improved from 8.3, 5.0, 6.2, and 4.4 days to 3.1, 2.5, 3.2, and 2.9 days, respectively. In
 397 no-value areas of these spatially incomplete images, the same reconstruction process as that in the
 398 cloud-covered areas was applied. The visual evaluation of the reconstructed flood maps suggested the
 399 effectiveness of the proposed method in seamless flood extent mapping under varying cloud
 400 conditions, i.e., in overall flood inundation trend and reconstruction details under different flood
 401 stages among the study sites.

402 In consideration of the connection between flood events and rainfall, comparing the areas of flood
 403 inundated regions with rainfall amounts could help generally evaluate the accuracy of the
 404 reconstructed flood maps. In this study, we collected daily precipitation data, i.e., CPC Global Unified
 405 Gauge-Based Analysis of Daily Precipitation from the NOAA Physical Sciences Laboratory website
 406 (<https://psl.noaa.gov/>), of each site during the study period. In Fig. 7, areas of identified pre-flood
 407 water and floodwater are compared with daily precipitation over the four study sites. Changes in flood
 408 inundated areas were not only related to daily precipitation over the study sites, but were also affected
 409 by topographic relief and the rainfall conditions nearby the study sites. Moreover, biases might exist
 410 in estimated change trends of floodwater areas because of the potential reconstruction error under
 411 varying cloud cover conditions. The comparison results showed a general consistency between the
 412 changes in floodwater areas and rainfall with consideration of the time lag effects, which
 413 demonstrated the effectiveness of the proposed method from a different perspective.



414

415

Fig. 6. Reconstruction results of the proposed method for seamless time series flood extent mapping

416

over four study sites. (a) Assam, India; (b) Sindh, Pakistan; (c) Rio Grande do Sul, Brazil; and (d)

417

Michigan, USA. For each study site, 4 (for the first three sites) or 2 (for the last site) images acquired

418

at different flood stages, and the corresponding water maps before and after reconstruction are

419

selected to be shown as examples. The 1st and 4th columns show the original HLS images, the 2nd

420

and 5th columns show the initial water maps overlaid with cloud/shadow masks, and the 3rd and

421

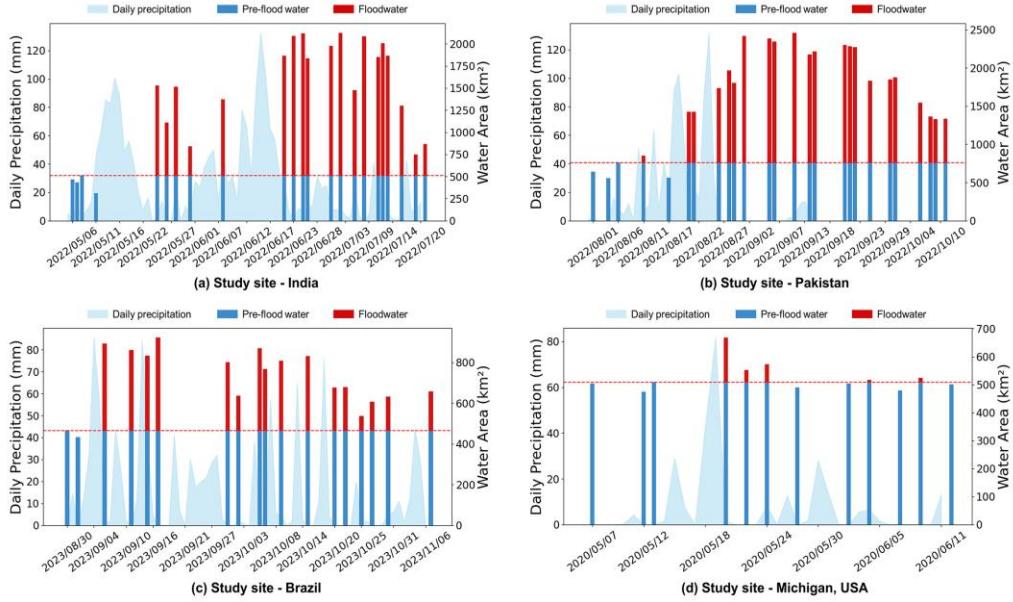
6th columns show the reconstructed water maps with flooded areas highlighted in red. Note that 'L'

422

or 'S' follows the acquisition dates on each image to indicate that the HLS images were obtained

423

from Landsat or Sentinel-2 satellites, respectively.



424

425 **Fig. 7.** Comparison of areas of identified floodwater (red bar) with daily precipitation (light blue
426 background) over four test sites during flooding periods.

427 3.3.2 Reconstruction performance evaluation under varying cloud cover conditions

428 Simulation reconstruction experiments and comparative analysis were conducted to
429 quantitatively evaluate the performance of the proposed method in the reconstruction of cloud-
430 covered areas in water maps. This evaluation encompassed groups of simulation experiments in
431 reconstructing cloud-covered areas and refining time series water maps, and comparisons of the
432 reconstructed results of the proposed method with those of a gap-filling approach (Zhao and Gao,
433 2018). In this study, benchmark datasets with reference water maps were constructed to evaluate
434 different reconstruction methods. We selected locally cloud-free HLS images with a size of
435 approximately 1000×1000 pixels from each of the four study sites and manually annotated the
436 extents of surface water. For the comprehensive evaluation of the performance of different
437 methods in diverse scenarios, HLS images acquired at different flood stages (i.e., before/after
438 flood, during flood) were involved to obtain reference water maps (i.e., ground truth) for
439 comparison. In addition, cloud masks with three levels of cloud coverage (i.e., low: <30%,
440 medium: 30%–60%, high: >60%) were overlaid with reference water maps to simulate cloud-
441 covered water maps for reconstruction. Different methods were employed to reconstruct the
442 simulated cloud-covered water maps together with other water maps in the time series of each
443 study site. The reconstruction results of different methods were compared against reference water
444 maps for quantitative accuracy evaluation, in which five metrics, namely, OA, precision, recall,

445 mIoU, and F1-score, were adopted for accuracy measurement.

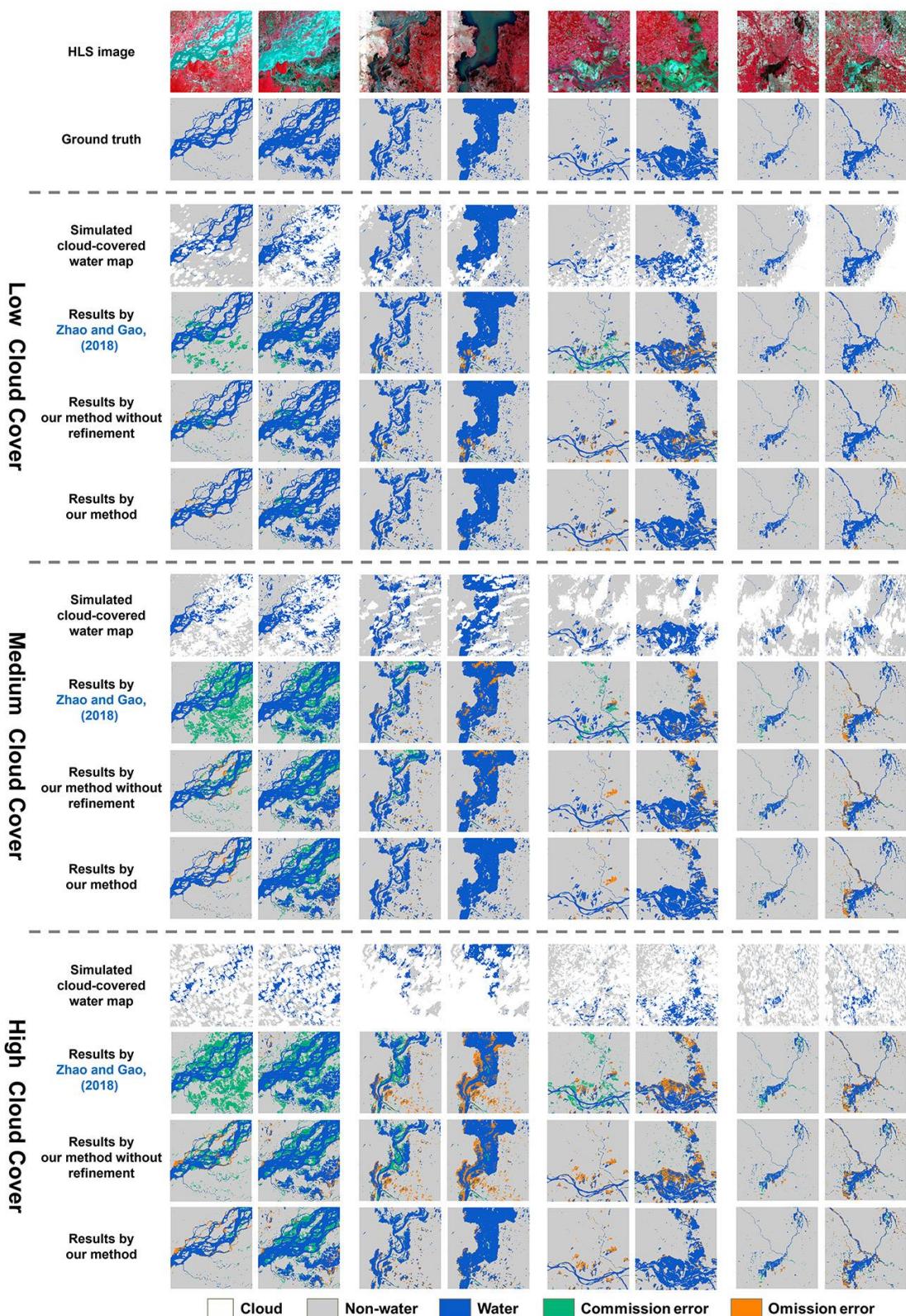


Fig. 8. Comparison of water extent maps reconstructed by the compared gap-filling method (Zhao and Gao, 2018) and the proposed method with and without spatiotemporal modeling-based refinement through simulated data experiments.

Table 2. Accuracy assessment of reconstructed water maps under varying cloud cover conditions before and during flood events.

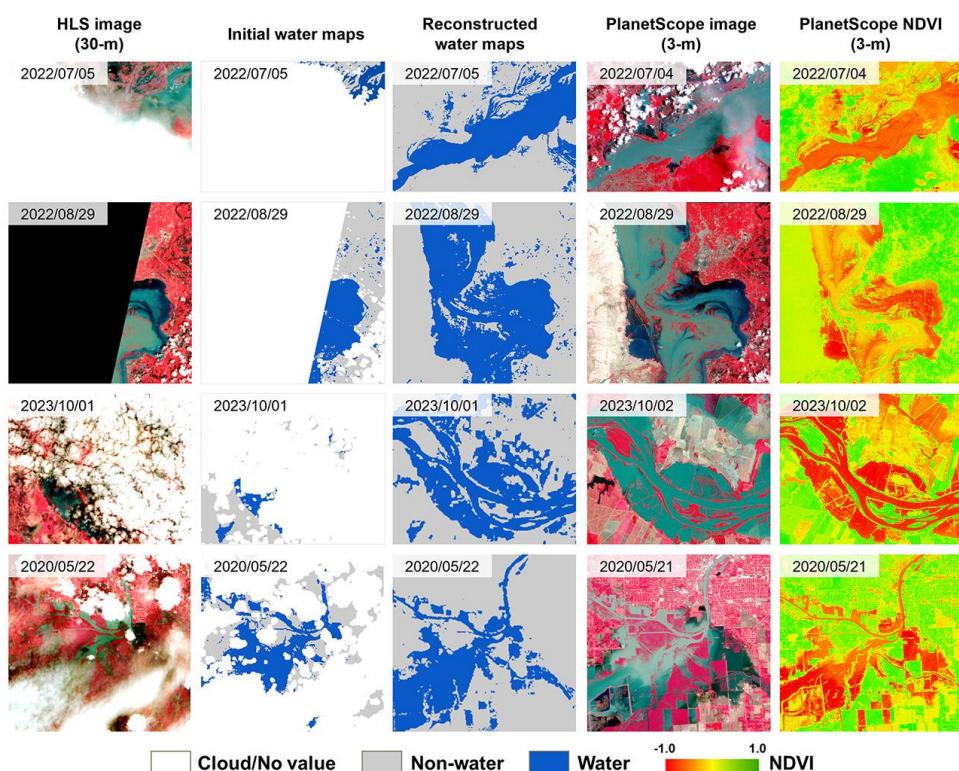
Phases	Cloud Cover	Method	OA	Precision	Recall	mIoU	F1-Score
During Floods	Low	Zhao and Gao (2018)	96.74%	96.37%	92.39%	0.891	0.942
		Ours without refinement	<u>97.29%</u>	<u>96.79%</u>	<u>93.22%</u>	<u>0.904</u>	<u>0.949</u>
		Ours	98.50%	97.49%	96.50%	0.942	0.970
	Medium	Zhao and Gao (2018)	92.67%	90.73%	86.19%	0.779	0.874
		Ours without refinement	<u>94.56%</u>	<u>92.56%</u>	<u>87.83%</u>	<u>0.816</u>	<u>0.898</u>
		Ours	96.14%	93.57%	91.96%	0.864	0.924
	High	Zhao and Gao (2018)	88.63%	87.93%	75.85%	0.663	0.797
		Ours without refinement	<u>91.08%</u>	<u>89.61%</u>	<u>78.76%</u>	<u>0.714</u>	<u>0.833</u>
		Ours	94.75%	91.84%	88.57%	0.822	0.898
	Mean Acc.	Zhao and Gao (2018)	92.68%	91.67%	84.81%	0.777	0.871
		Ours without refinement	<u>94.31%</u>	<u>92.99%</u>	<u>86.60%</u>	<u>0.811</u>	<u>0.893</u>
		Ours	96.46%	94.30%	92.34%	0.876	0.931
Before/After Floods	Low	Zhao and Gao (2018)	96.21%	81.04%	94.41%	0.776	0.867
		Ours without refinement	<u>98.35%</u>	<u>95.39%</u>	90.86%	<u>0.868</u>	<u>0.928</u>
		Ours	98.79%	97.71%	<u>90.93%</u>	0.890	0.940
	Medium	Zhao and Gao (2018)	90.62%	68.56%	94.10%	0.650	0.776
		Ours without refinement	<u>96.39%</u>	<u>87.42%</u>	88.72%	<u>0.782</u>	<u>0.877</u>
		Ours	98.07%	94.43%	<u>89.18%</u>	0.845	0.915
	High	Zhao and Gao (2018)	86.80%	58.38%	85.58%	0.518	0.673
		Ours without refinement	<u>93.71%</u>	<u>81.45%</u>	76.37%	<u>0.635</u>	<u>0.775</u>
		Ours	96.87%	92.10%	<u>80.62%</u>	0.747	0.853
	Mean Acc.	Zhao and Gao (2018)	91.21%	69.33%	91.36%	0.648	0.772
		Ours without refinement	<u>96.15%</u>	<u>88.09%</u>	85.32%	<u>0.762</u>	<u>0.860</u>
		Ours	97.91%	94.75%	<u>86.91%</u>	0.828	0.903

453 Fig. 8 shows the reconstruction results of the simulation experiments, and Table 2 provides the
454 detailed accuracy evaluation results of the compared gap-filling method and the proposed method
455 with and without the refinement process described in Section 2.3. The reconstruction results
456 showed that our methods generally outperformed the compared gap-filling method in terms of
457 OA, mIoU, and F1-score under different conditions, with much less water misidentification errors,
458 as shown in Fig. 8, and such superiorities strengthened during flood. While the compared gap-
459 filling method generally achieved higher recall compared with our methods during non-flood
460 periods, it sacrificed precision. Our method employed a stable ratio-based strategy to characterize
461 water extent, making it less sensitive to the impact of varying cloud cover conditions.
462 Furthermore, the locally asymptotic window reconstruction strategy employed in the proposed
463 method considered the heterogeneity of local areas during flood, contributing to better
464 reconstruction results.

465 Specifically, during flood and non-flood periods, the OA of the proposed method was 96.46%
466 and 97.91%, respectively, showing absolute improvements of 3.78% and 6.70%, respectively,
467 over the compared gap-filling method. The F1-score of the proposed method was 0.931 and 0.903,
468 also showcasing obvious absolute increases of 0.060 and 0.131, respectively, compared with the
469 gap-filling method. Under three conditions with low, medium, and high cloud coverage, the
470 average F1-score of the proposed method was 0.955, 0.920, and 0.875, respectively. Although the
471 reconstruction accuracy decreased as the cloud cover percentages increased, our method still
472 demonstrated significant improvements in F1-score compared with the gap-filling method during
473 flood and non-flood periods, with absolute increases of 0.051, 0.095, and 0.140 under low,
474 medium, and high cloud coverage conditions, respectively. The above comparisons highlighted the
475 effectiveness of the proposed method under different flood stages and varying cloud cover
476 conditions. Furthermore, compared with the results of our method without refinement under low,
477 medium, and high cloud coverage, our method with refinement showed absolute improvements of
478 0.016, 0.033, and 0.071 in terms of F1-score, respectively. This finding indicated that the time
479 series refinement via spatiotemporal modeling played a more crucial role in the reconstruction of
480 cloud-covered areas as cloud cover percentages increased, leading to greater accuracy
481 improvements.

482 **3.3.3 Comparison of reconstruction results with high-resolution PlanetScope images**

483 Commercial PlanetScope satellite images with a high resolution of 3 m were introduced to
 484 further validate the reconstructed flood maps with real flood situations. PlanetScope satellites
 485 capture images of the entire Earth surface almost daily. In contrast with the 30 m-resolution HLS
 486 images, 3 m PlanetScope images allow for the identification of water with finer details, including
 487 smaller water bodies. We collected the PlanetScope images acquired at the same or nearest
 488 neighboring date as the HLS images over the four study sites. Due to the differences in imaging
 489 times of different satellites, cloud cover conditions and flooding extents in the HLS and
 490 PlanetScope images can vary substantially. Therefore, PlanetScope images can only be used for
 491 visual validation. Fig. 8 offers a visual comparison of the HLS images and the initial and
 492 reconstructed water maps with PlanetScope images and their derived Normalized Difference
 493 Vegetation Index (NDVI) maps. NDVI has been considered effective in distinguishing surface
 494 water from land pixels (Vermote and Saleous, 2007; Zhu and Woodcock, 2012), and is thus used
 495 to highlight the surface water in PlanetScope images. The comparison showed that the proposed
 496 method could achieve relatively accurate water map reconstruction even for images with limited
 497 cloud-free areas.



498 **Fig. 9.** Comparison of reconstructed flood extent maps with high-resolution PlanetScope images.
 499

500 **4. Discussion**

501 **4.1 Benefits of cloud reconstruction for flood mapping**

502 To further demonstrate the significance and benefits of cloud reconstruction for flood mapping,
503 we compared the differences between flood maps generated with and without cloud
504 reconstruction. As illustrated in Fig. 10, the comparisons involved example HLS images taken
505 during floods, the maximum water extents in the GSW dataset, and maximum water extents and
506 flood duration maps composited with initial and reconstructed time series water maps. The initial
507 water maps might be spatially incomplete owing to cloud cover. In this paper, the maximum water
508 extent denotes the maximum-extent flood inundated areas, and the flood duration map represents
509 the frequency (i.e., the duration of floods) at which inundation was observed. The results showed
510 that the reconstructed maximum water extents provided timely flood information compared with
511 the reference extents in the GSW dataset and more spatially complete flood inundated areas
512 compared with those composited with initial water maps. Additionally, the reconstructed seamless
513 water maps enhanced the spatial completeness of the composited flood duration maps in
514 comparison with the flood duration maps composited with initial water maps. Similar conclusions
515 on the benefits of cloud reconstruction for time series water mapping were discussed in recent
516 studies (Bai et al., 2022; Huang et al., 2023). The maximum water extent and flood duration maps
517 are crucial for flood monitoring and management, such as in flood impact assessments (Bofana et
518 al., 2022). Although uncertainties may arise in the reconstruction, our method generally provides
519 an effective way to generate accurate and complete flood maps and thus benefits flood mapping
520 under cloudy and rainy scenarios.

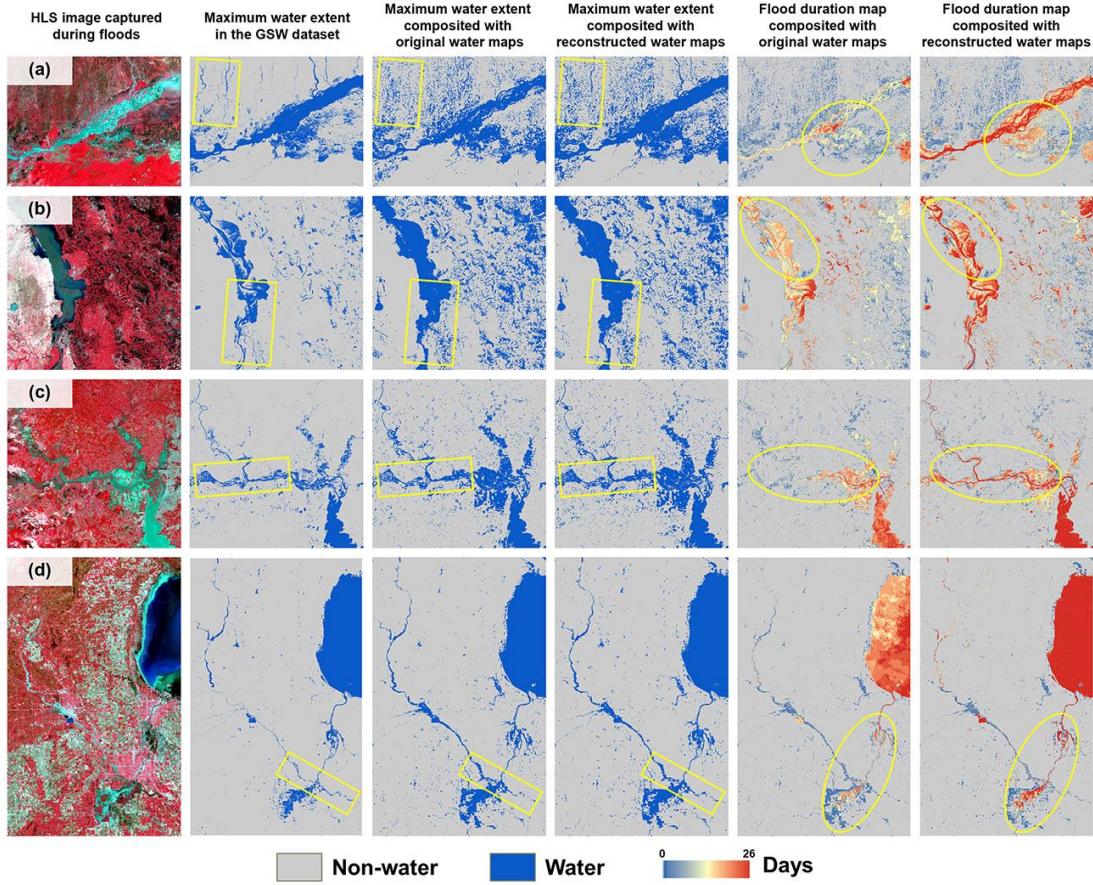
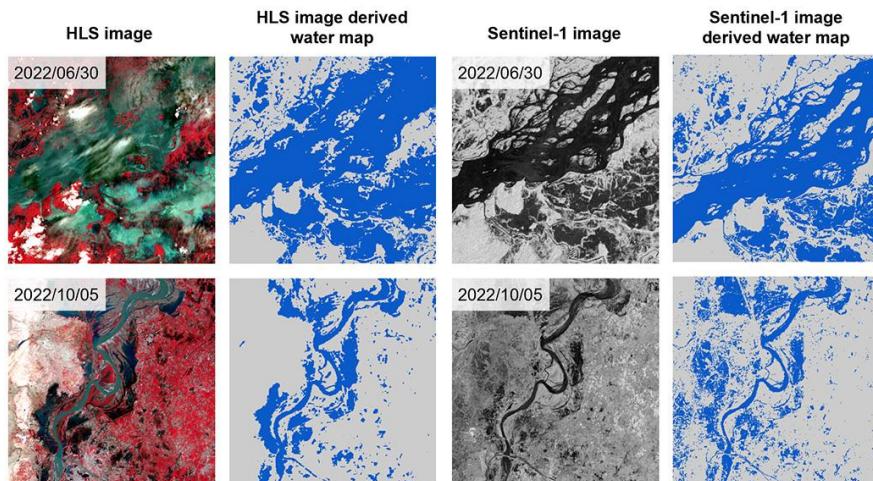


Fig. 10. Comparison of maximum water extent and flood duration maps composed with initial and reconstructed time series water maps. HLS images captured in (a) Assam, India on 2022/07/15; (b) Sindh, Pakistan on 2022/09/10; (c) Rio Grande do Sul, Brazil on 2023/09/11; (d) Michigan, USA on 2020/05/20. Yellow polygons in the figure panels highlight the major differences among the compared maps.

4.2 Incorporating SAR images for enhanced flooding area reconstruction

All-weather SAR can penetrate clouds to conduct flood monitoring in all weather conditions; thus, SAR data are promising for delineating surface water under cloudy conditions. Theoretically, incorporating SAR data can enhance the temporal frequency of flood monitoring and provide spatially complete water maps. To validate the importance of SAR data in enhancing cloud-covered flooding area reconstruction, we collected Sentinel-1 SAR images within each study site from Google Earth Engine (GEE). Note that each Sentinel-1 image available on GEE was preprocessed with the Sentinel-1 toolbox, and the steps including noise removal, radiometric calibration, and terrain correction. Sentinel-1-derived water maps were obtained on the basis of the Sentinel-1 dual-polarized water index (Huang et al., 2023), which were then resampled and stacked with time series water maps derived from HLS images in a time order. The newly stacked

538 time series maps were finally processed through time series refinement to reconstruct simulated
 539 cloud-covered water maps, as described in [Section 3.3.2](#). We compared the accuracy evaluation
 540 results with and without incorporating Sentinel-1 images to assess the impact of SAR data on
 541 enhancing flood mapping. The experimental results showed that incorporating SAR data could
 542 enhance the reconstruction accuracy with a slight increase of 0.12% in F1-score. However, as
 543 shown in [Fig. 11](#), owing to the differences in imaging mechanisms between SAR and optical
 544 images, inconsistencies existed in the water extraction results derived from SAR and optical
 545 images ([Markert et al., 2018](#)), even when both were acquired on the same day. Thus, while
 546 incorporating SAR images is beneficial to enhance the flooding area reconstruction, especially
 547 under heavy cloud cover conditions, the harmonization of water maps derived from SAR and
 548 optical images warrants further exploration to strengthen this benefit.



549
 550 **Fig. 11.** Comparison of flood mapping results generated from optical (i.e., Landsat and Sentinel-2)
 551 and SAR (i.e., Sentinel-1) images acquired on the same day.

552 4.3 Efficiency

553 Overall, the proposed method demonstrates high efficiency in processing time. Specifically,
 554 using one of the study areas in Assam, India, as an example, there are 24 HLS images in the time
 555 series, each with a size of 3849×2880 pixels and 76.10% mean cloud cover. Tested on a desktop
 556 computer with an Intel i5-12490F CPU in a Python environment, our method took 9,234 seconds
 557 to complete the processing steps and generate the seamless time series flood maps. Note that the
 558 method by [Zhao and Gao \(2018\)](#) applied a global threshold for cloud reconstruction, allowing for
 559 a comprehensive comparison with our methods that apply global and local thresholds,
 560 respectively, in both efficiency and accuracy. Additionally, only the time cost for the cloud

561 reconstruction step is considered among the methods. The efficiency assessment results show that
562 the proposed method took 6,703 seconds with local reconstruction and 641 seconds with global
563 reconstruction, as being compared to 482 seconds by the method of [Zhao and Gao \(2018\)](#). While
564 our method, using a local sliding window strategy for local reconstruction, took longer than the
565 method by [Zhao and Gao \(2018\)](#), which applied global reconstruction. Our method yields a
566 significant improvement in mean overall accuracy, rising from 82.65% to 92.76% in Assam, India,
567 as detailed in Section 3.3.2. Nevertheless, the efficiency of our method could be further optimized
568 through implementation enhancements and parallel processing.

569 **4.4 Limitations**

570 Cloud cover significantly impairs the capabilities of optical satellites in monitoring floods.
571 Although the reconstruction method proposed in this paper is effective in mitigating the influence
572 of clouds to support continuous time series mapping of flood extent, limitations remained within
573 its multiple steps.

574 The fine-tuned large foundation model exhibits high accuracy in water extraction, but it still
575 introduces errors, such as misclassifying cloud shadows as surface water and omitting the surface
576 water obstructed by vegetation, thus leaving room for improvements. Furthermore, the QA band of
577 the HLS products tends to overestimate actual cloud cover, complicating the reconstruction
578 process given the limited number of valid pixels. Therefore, developing an enhanced large
579 foundation model with robust cloud masking capabilities is worthy of further exploration.

580 The performance of the proposed method decreases under conditions of persistent and heavy
581 cloud cover. Our analysis indicates that for water map reconstruction, the method should be
582 applied when the cloud cover is less than 96% to maintain a typical F1-score accuracy above 0.8.
583 When cloud cover exceeds 96%, the results become considerably uncertain given insufficient
584 cloud-free observational data. Therefore, a cloud cover percentage below 96% is essential for
585 effective water map reconstruction of cloud-covered areas.

586 Reconstructing cloud-covered areas in water maps during flood periods is more challenging
587 compared with that during non-flood periods. The reason is that cloud cover is often more severe
588 during flood periods than during non-flood periods, leading to a reduction in available temporally
589 auxiliary information, thereby making the improvement through time series refinement less

590 apparent. In addition, with the relatively low frequency of flooding in surface water dynamics,
591 cloud-covered flooding area reconstruction is limited by the utilized water occurrence data, which
592 might introduce omissions of floodwater in the reconstructed water maps, especially for extreme
593 flooding events. While the time series refinement of water maps via spatiotemporal modeling is
594 beneficial to alleviate this issue, such a limitation, sourced from water occurrence data, may hinder
595 the potential of the proposed method in capturing extreme flood events intensified by climate
596 change.

597 **5. Conclusions**

598 This study investigated seamless flood extent mapping using HLS image time series. We
599 developed a robust method to reconstruct cloud-covered areas in time series water maps and
600 evaluated its effectiveness for spatially continuous flood extent mapping under various flood
601 scenarios with differing cloud cover conditions. The method was proven effective in enhancing
602 time series flood monitoring and outperformed the compared gap-filling method in reconstructing
603 cloud-covered flooding areas. Flood extent mapping experiments at four global study sites
604 suggested an improvement in the average reconstruction accuracy, as measured by F1-score, from
605 0.871 to 0.931 during flood periods and from 0.772 to 0.903 during non-flood periods. The
606 superiority of the proposed method became increasingly prominent as cloud cover increased, with
607 an improvement in F1-score of up to 0.140 under high cloud coverage conditions.

608 The experimental results also indicated that reconstructing cloud-covered areas in time series
609 water maps benefited the composition of maximum flood extent maps and flood duration maps.
610 The reconstructed maps exhibited better spatial completeness and consistency compared with
611 those composed without involving reconstruction. In addition, we consider the incorporation of
612 SAR images as a promising way to enhance reconstruction results under persistent and dense
613 cloud cover conditions, even though uncertainties and inconsistencies occurred between water
614 maps derived from optical and SAR images owing to differences in their physical imaging
615 mechanisms. The future combined use of optical and SAR images for flood monitoring should
616 address this issue to ensure consistent flood extent mapping.

617 Overall, the proposed method provides an effective approach for flood monitoring under cloudy
618 and rainy scenarios, thus supporting emergency response and disaster management. Future studies

619 could explore the harmonization of water maps derived from multimodal and multisensor data.
620 The delineation of water and cloud is crucial for effective flood monitoring, and the accuracy of
621 water maps and cloud masks promises further enhancement with the introduction of a multi-
622 temporal foundation model for image time series. Hydrodynamic modeling and additional
623 auxiliary data, such as hydrodynamic simulation and accurate high-resolution DEM data, when
624 available, can be utilized to refine reconstruction results further, especially when the proposed
625 method is applied to mapping extreme flooding events.

626 **Acknowledgments**

627 This research has received funding from Global STEM Professorship (P0039329) and Hong Kong
628 Polytechnic University (P0046482 and P0038446). This research was partially supported by the
629 National Natural Science Foundation of China (No. 42101357). The authors also acknowledge
630 Planet Labs for providing high-resolution PlanetScope images that supported this study. Last but
631 not least, the authors are grateful to the editors and reviewers for their constructive comments and
632 suggestions which helped improve the manuscript.

633 **References**

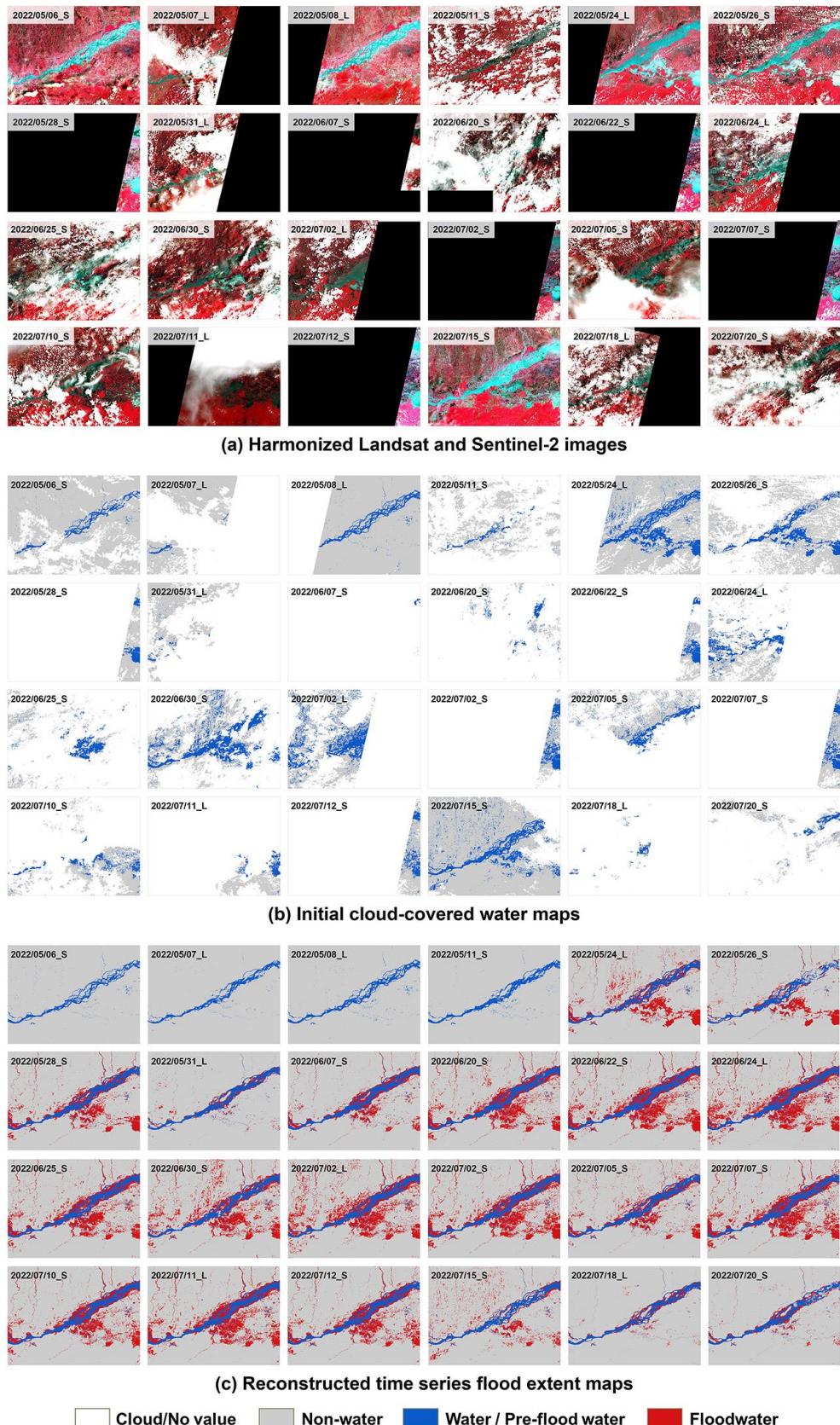
- 634 Bai, B., Tan, Y., Donchyts, G., Haag, A., Xu, B., Chen, G., Weerts, A.H., 2023. Naive Bayes
635 classification-based surface water gap-filling from partially contaminated optical remote sensing
636 image. *J. Hydrol.* 616. <https://doi.org/10.1016/j.jhydrol.2022.128791>
- 637 Bai, B., Tan, Y., Zhou, K., Donchyts, G., Haag, A., Weerts, A.H., 2022. Time-series surface water gap
638 filling based on spatiotemporal neighbourhood similarity. *Int. J. Appl. Earth Obs. Geoinf.* 112,
639 102882. <https://doi.org/10.1016/j.jag.2022.102882>
- 640 Benoudjit, A., Guida, R., 2019. A novel fully automated mapping of the flood extent on sar images
641 using a supervised classifier. *Remote Sens.* 11, 779. <https://doi.org/10.3390/rs11070779>
- 642 Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2022. Deep learning methods for flood
643 mapping: a review of existing applications and future research directions. *Hydrol. Earth Syst. Sci.*
644 26, 4345–4378. <https://doi.org/10.5194/hess-26-4345-2022>
- 645 Bofana, J., Zhang, M., Wu, B., Zeng, H., Nabil, M., Zhang, N., Elnashar, A., Tian, F., da Silva, J.M.,
646 Botão, A., Atumane, A., Mushore, T.D., Yan, N., 2022. How long did crops survive from floods
647 caused by Cyclone Idai in Mozambique detected with multi-satellite data. *Remote Sens. Environ.*
648 269, 112808. <https://doi.org/10.1016/j.rse.2021.112808>
- 649 Bonafilia, D., Tellman, B., Anderson, T., Issenberg, E., 2020. Sen1Floods11: A georeferenced dataset to
650 train and test deep learning flood algorithms for sentinel-1, in: IEEE Computer Society Conference
651 on Computer Vision and Pattern Recognition Workshops. IEEE, pp. 835–845.
652 <https://doi.org/10.1109/CVPRW50498.2020.00113>

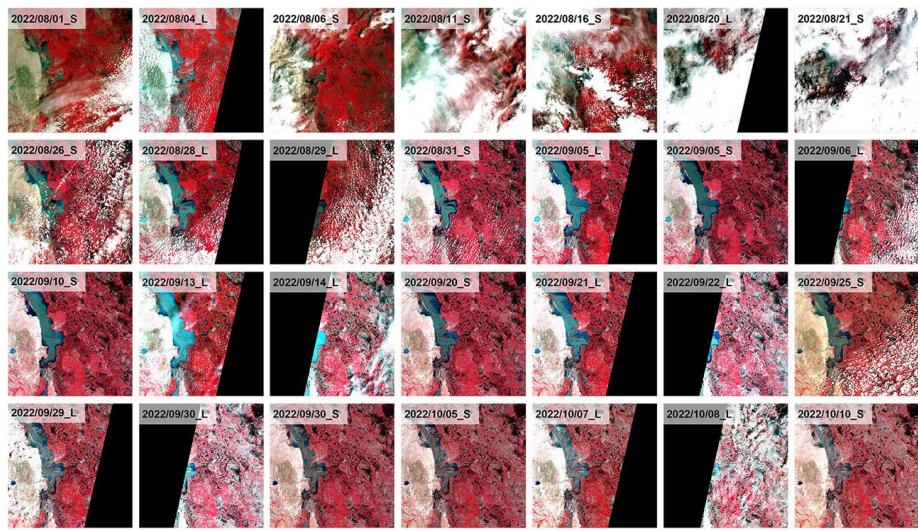
- 653 Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S. V., Justice, C.,
654 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.*
655 219, 145–161. <https://doi.org/10.1016/j.rse.2018.09.002>
- 656 Deng, Y., Jiang, W., Tang, Z., Li, J., Lv, J., Chen, Z., Jia, K., 2017. Spatio-temporal change of lake
657 water extent in Wuhan urban agglomeration based on Landsat images from 1987 to 2015. *Remote*
658 *Sens.* 9, 270. <https://doi.org/10.3390/rs9030270>
- 659 DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J.W., Lang, M.W., 2020. Rapid and robust
660 monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote*
661 *Sens. Environ.* 240, 111664. <https://doi.org/10.1016/j.rse.2020.111664>
- 662 Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M.,
663 Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2021. An Image is Worth 16x16
664 Words: Transformers for Image Recognition at Scale. *ICLR 2021 - 9th Int. Conf. Learn. Represent.*
- 665 Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014. Automated Water Extraction Index: A new
666 technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* 140, 23–35.
667 <https://doi.org/10.1016/j.rse.2013.08.029>
- 668 Goffi, A., Stroppiana, D., Brivio, P.A., Bordogna, G., Boschetti, M., 2020. Towards an automated
669 approach to map flooded areas from Sentinel-2 MSI data and soft integration of water spectral
670 features. *Int. J. Appl. Earth Obs. Geoinf.* 84, 101951. <https://doi.org/10.1016/j.jag.2019.101951>
- 671 He, K., Chen, X., Xie, S., Li, Y., Dollar, P., Girshick, R., 2022. Masked Autoencoders Are Scalable
672 Vision Learners, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision*
673 and Pattern Recognition.
- 674 pp. 15979–15988. <https://doi.org/10.1109/CVPR52688.2022.01553>
- 675 He, Y., Wang, J., Zhang, Y., Liao, C., 2024. An efficient urban flood mapping framework towards
676 disaster response driven by weakly supervised semantic segmentation with decoupled training
677 samples. *ISPRS J. Photogramm. Remote Sens.* 207, 338–358.
<https://doi.org/10.1016/j.isprsjprs.2023.12.009>
- 678 Huang, M., Jin, S., 2020. Rapid flood mapping and evaluation with a supervised classifier and change
679 detection in Shouguang using Sentinel-1 SAR and Sentinel-2 optical data. *Remote Sens.* 12, 2073.
680 <https://doi.org/10.3390/rs12132073>
- 681 Huang, W., Wang, F., Duan, W., Tang, L., Qin, J., Meng, X., 2023. Image similarity-based gap filling
682 method can effectively enrich surface water mapping information. *ISPRS J. Photogramm. Remote*
683 *Sens.* 202, 528–544. <https://doi.org/10.1016/j.isprsjprs.2023.07.011>
- 684 IPCC, 2023. *Climate Change 2021 – The Physical Science Basis*. Cambridge University Press.
685 <https://doi.org/10.1017/9781009157896>
- 686 Jakubik, J., Chu, L., Fraccaro, P., Gomes, C., Nyirjesy, G., Bangalore, R., Lambhate, D., Das, K.,
687 Oliveira Borges, D., Kimura, D., Simumba, N., Szwarcman, D., Muszynski, M., Weldemariam, K.,
688 Zadrozny, B., Ganti, R., Costa, C., Edwards Blair & Watson, C., Mukkavilli, K., Schmude Johannes
689 & Hamann, H., Robert, P., Roy, S., Phillips, C., Ankur, K., Ramasubramanian, M., Gurung, I.,
690 Leong, W.J., Avery, R., Ramachandran, R., Maskey, M., Olofossen, P., Fancher, E., Lee, T., Murphy,
691 K., Duffy, D., Little, M., Alemohammad, H., Cecil, M., Li, S., Khallaghi, S., Godwin, D., Ahmadi,
692 M., Kordi, F., Saux, B., Pastick, N., Doucette, P., Fleckenstein, R., Luanga, D., Corvin, A., Granger,
693 E., 2023. Prithvi-100M. <https://github.com/NASA-IMPACT/hls-foundation-os>.
694 <https://doi.org/10.57967/hf/0952>

- 695 Kasetkasem, T., Phuhinkong, P., Rakwatin, P., Chanwimaluang, T., Kumazawa, I., 2014. A flood
696 mapping algorithm from cloud contaminated MODIS time-series data using a Markov random field
697 model. *Int. Geosci. Remote Sens. Symp.* 2507–2510.
698 <https://doi.org/10.1109/IGARSS.2014.6946982>
- 699 Konapala, G., Kumar, S. V., Khalique Ahmad, S., 2021. Exploring Sentinel-1 and Sentinel-2 diversity
700 for flood inundation mapping using deep learning. *ISPRS J. Photogramm. Remote Sens.* 180, 163–
701 173. <https://doi.org/10.1016/j.isprsjprs.2021.08.016>
- 702 Li, J., Narayanan, R.M., 2003. A Shape-Based Approach to Change Detection of Lakes Using Time
703 Series Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 41, 2466–2477.
704 <https://doi.org/10.1109/TGRS.2003.817267>
- 705 Li, M., Lu, S., Du, C., Wang, Y., Fang, C., Li, X., Tang, H., Baig, M.H.A., Ikhumhen, H.O., 2022.
706 Time-series surface water reconstruction method (TSWR) based on spatial distance relationship of
707 multi-stage water boundaries. *Int. J. Digit. Earth* 15, 2335–2354.
708 <https://doi.org/10.1080/17538947.2022.2159553>
- 709 Li, S., Sun, D., Goldberg, M.D., Kalluri, S., Sjoberg, B., Lindsey, D., Hoffman, J.P., DeWeese, M.,
710 Connelly, B., McKee, P., Lander, K., 2022. A downscaling model for derivation of 3-D flood
711 products from VIIRS imagery and SRTM/DEM. *ISPRS J. Photogramm. Remote Sens.* 192, 279–
712 298. <https://doi.org/10.1016/j.isprsjprs.2022.08.025>
- 713 Li, S.Z., 2009. Markov random field modeling in image analysis. Springer Science & Business Media.
- 714 Li, X., Ling, F., Cai, X., Ge, Y., Li, Xiaodong, Yin, Z., Shang, C., Jia, X., Du, Y., 2021. Mapping water
715 bodies under cloud cover using remotely sensed optical images and a spatiotemporal dependence
716 model. *Int. J. Appl. Earth Obs. Geoinf.* 103, 102470. <https://doi.org/10.1016/j.jag.2021.102470>
- 717 Li, Z., Shen, H., Weng, Q., Zhang, Y., Dou, P., Zhang, L., 2022. Cloud and cloud shadow detection for
718 optical satellite imagery: Features, algorithms, validation, and prospects. *ISPRS J. Photogramm.*
719 *Remote Sens.* 188, 89–108. <https://doi.org/10.1016/j.isprsjprs.2022.03.020>
- 720 Li, Z., Weng, Q., Zhou, Y., Dou, P., Ding, X., 2024. Learning spectral-indices-fused deep models for
721 time-series land use and land cover mapping in cloud-prone areas: The case of Pearl River Delta.
722 *Remote Sens. Environ.* 308, 114190. <https://doi.org/10.1016/j.rse.2024.114190>
- 723 Liang, J., Liu, D., 2020. A local thresholding approach to flood water delineation using Sentinel-1 SAR
724 imagery. *ISPRS J. Photogramm. Remote Sens.* 159, 53–62.
725 <https://doi.org/10.1016/j.isprsjprs.2019.10.017>
- 726 Liu, F., Chen, D., Guan, Z., Zhou, X., Zhu, J., Ye, Q., Fu, L., Zhou, J., 2023. RemoteCLIP: A Vision
727 Language Foundation Model for Remote Sensing. *IEEE Trans. Geosci. Remote Sens.*
728 <https://doi.org/10.1109/TGRS.2024.3390838>
- 729 Markert, K.N., Chishtie, F., Anderson, E.R., Saah, D., Griffin, R.E., 2018. On the merging of optical
730 and SAR satellite imagery for surface water mapping applications. *Results Phys.* 9, 275–277.
731 <https://doi.org/10.1016/j.rinp.2018.02.054>
- 732 McCormack, T., Campanyà, J., Naughton, O., 2022. A methodology for mapping annual flood extent
733 using multi-temporal Sentinel-1 imagery. *Remote Sens. Environ.* 282, 113273.
734 <https://doi.org/10.1016/j.rse.2022.113273>
- 735 McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation
736 of open water features. *Int. J. Remote Sens.* 17, 1425–1432.

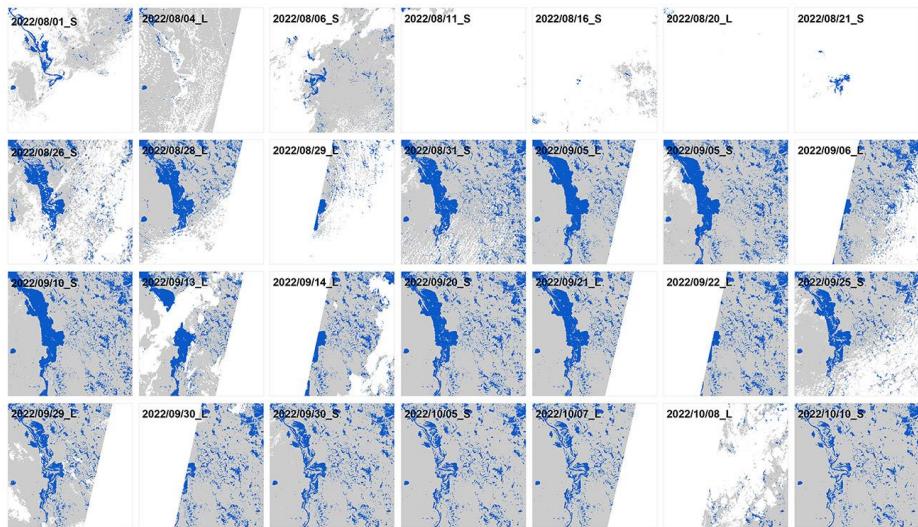
- 737 <https://doi.org/10.1080/01431169608948714>
- 738 Moussouris, J., 1974. Gibbs and Markov random systems with constraints. *J. Stat. Phys.* 10, 11–33.
- 739 <https://doi.org/10.1007/BF01011714>
- 740 Mullen, C., Penny, G., Müller, M.F., 2021. A simple cloud-filling approach for remote sensing water
741 cover assessments. *Hydrol. Earth Syst. Sci.* 25, 2373–2386. <https://doi.org/10.5194/hess-25-2373-2021>
- 742 Notti, D., Giordan, D., Caló, F., Pepe, A., Zucca, F., Galve, J.P., 2018. Potential and limitations of open
743 satellite data for flood mapping. *Remote Sens.* 10, 1673. <https://doi.org/10.3390/rs10111673>
- 744 Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface
745 water and its long-term changes. *Nature* 540, 418–422. <https://doi.org/10.1038/nature20584>
- 746 Psomiadis, E., Diakakis, M., Soulis, K.X., 2020. Combining SAR and optical earth observation with
747 hydraulic simulation for flood mapping and impact assessment. *Remote Sens.* 12, 1–29.
748 <https://doi.org/10.3390/rs12233980>
- 749 Rentschler, J., Salhab, M., Jafino, B.A., 2022. Flood exposure and poverty in 188 countries. *Nat.*
750 *Commun.* 13. <https://doi.org/10.1038/s41467-022-30727-4>
- 751 Shastry, A., Carter, E., Coltin, B., Sleeter, R., McMichael, S., Eggleston, J., 2023. Mapping floods from
752 remote sensing data and quantifying the effects of surface obstruction by clouds and vegetation.
753 *Remote Sens. Environ.* 291, 113556. <https://doi.org/10.1016/j.rse.2023.113556>
- 754 Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L., 2015. Missing Information
755 Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geosci. Remote Sens. Mag.* 3,
756 61–85. <https://doi.org/10.1109/MGRS.2015.2441912>
- 757 Sun, F., Zhao, Y., Gong, P., Ma, R., Dai, Y., 2014. Monitoring dynamic changes of global land cover
758 types: fluctuations of major lakes in China every 8 days during 2000–2010. *Chinese Sci. Bull.* 59,
759 171–189. <https://doi.org/10.1007/s11434-013-0045-0>
- 760 Sun, X., Li, L., Zhang, B., Chen, D., Gao, L., 2015. Soft urban water cover extraction using mixed
761 training samples and support vector machines. *Int. J. Remote Sens.* 36, 3331–3344.
- 762 Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A.,
763 Slayback, D.A., 2021. Satellite imaging reveals increased proportion of population exposed to
764 floods. *Nature* 596, 80–86. <https://doi.org/10.1038/s41586-021-03695-w>
- 765 Tulbure, M.G., Broich, M., Perin, V., Gaines, M., Ju, J., Stehman, S. V., Pavelsky, T., Masek, J.G., Yin,
766 S., Mai, J., Betbeder-Matibet, L., 2022. Can we detect more ephemeral floods with higher density
767 harmonized Landsat Sentinel 2 data compared to Landsat 8 alone? *ISPRS J. Photogramm. Remote*
768 *Sens.* 185, 232–246. <https://doi.org/10.1016/j.isprsjprs.2022.01.021>
- 769 Valman, S.J., Boyd, D.S., Carboneau, P.E., Johnson, M.F., Dugdale, S.J., 2024. An AI approach to
770 operationalise global daily PlanetScope satellite imagery for river water masking. *Remote Sens.*
771 *Environ.* 301, 113932. <https://doi.org/10.1016/j.rse.2023.113932>
- 772 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.,
773 2017. Attention is all you need, in: *Advances in Neural Information Processing Systems*. pp. 5999–
774 6009.
- 775 Vermote, E., Saleous, N., 2007. LEDAPS surface reflectance product description. *Coll. Park Univ.* 1–
776 21.
- 777

- 778 Wang, C., Jia, M., Chen, N., Wang, W., 2018. Long-term surface water dynamics analysis based on
779 landsat imagery and the Google Earth Engine Platform: A case study in the middle Yangtze River
780 Basin. *Remote Sens.* 10, 1635. <https://doi.org/10.3390/rs10101635>
- 781 Wang, D., Zhang, Q., Xu, Y., Zhang, J., Du, B., Tao, D., Zhang, L., 2023. Advancing Plain Vision
782 Transformer Toward Remote Sensing Foundation Model. *IEEE Trans. Geosci. Remote Sens.* 61, 1–
783 15. <https://doi.org/10.1109/TGRS.2022.3222818>
- 784 Wang, Y., Li, Z., Zeng, C., Xia, G.S., Shen, H., 2020. An Urban Water Extraction Method Combining
785 Deep Learning and Google Earth Engine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*
786 <https://doi.org/10.1109/JSTARS.2020.2971783>
- 787 Xu, H., 2005. A Study on Information Extraction of Water Body with the Modified Normalized
788 Difference Water Index (MNDWI). *Natl. Remote Sens. Bull.* 0, 589–595.
789 <https://doi.org/10.11834/jrs.20050586>
- 790 Yang, X., Qin, Q., Yésou, H., Ledauphin, T., Koehl, M., Grussenmeyer, P., Zhu, Z., 2020. Monthly
791 estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data. *Remote*
792 *Sens. Environ.* 244, 111803. <https://doi.org/10.1016/j.rse.2020.111803>
- 793 Zeng, Z., Gan, Y., Kettner, A.J., Yang, Q., Zeng, C., Brakenridge, G.R., Hong, Y., 2020. Towards high
794 resolution flood monitoring: An integrated methodology using passive microwave brightness
795 temperatures and Sentinel synthetic aperture radar imagery. *J. Hydrol.* 582, 124377.
796 <https://doi.org/10.1016/j.jhydrol.2019.124377>
- 797 Zhang, Q., Yuan, Q., Li, Z., Sun, F., Zhang, L., 2021. Combined deep prior with low-rank tensor SVD
798 for thick cloud removal in multitemporal images. *ISPRS J. Photogramm. Remote Sens.* 177, 161–
799 173. <https://doi.org/10.1016/j.isprsjprs.2021.04.021>
- 800 Zhao, G., Gao, H., 2018. Automatic Correction of Contaminated Images for Assessment of Reservoir
801 Surface Area Dynamics. *Geophys. Res. Lett.* 45, 6092–6099.
802 <https://doi.org/10.1029/2018GL078343>
- 803 Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: Cloud,
804 cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sens.*
805 *Environ.* 159, 269–277. <https://doi.org/10.1016/j.rse.2014.12.014>
- 806 Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery.
807 *Remote Sens. Environ.* 118, 83–94. <https://doi.org/10.1016/j.rse.2011.10.028>
- 808

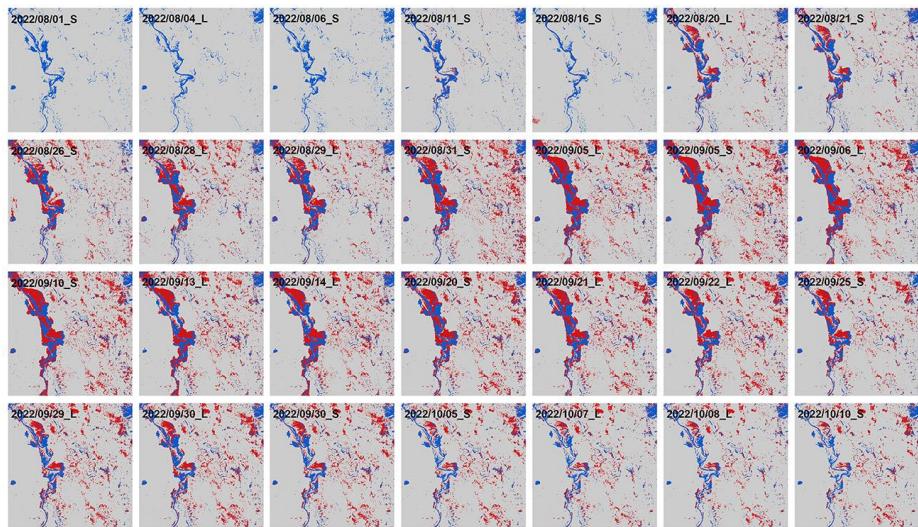




(a) Harmonized Landsat and Sentinel-2 images



(b) Initial cloud-covered water maps

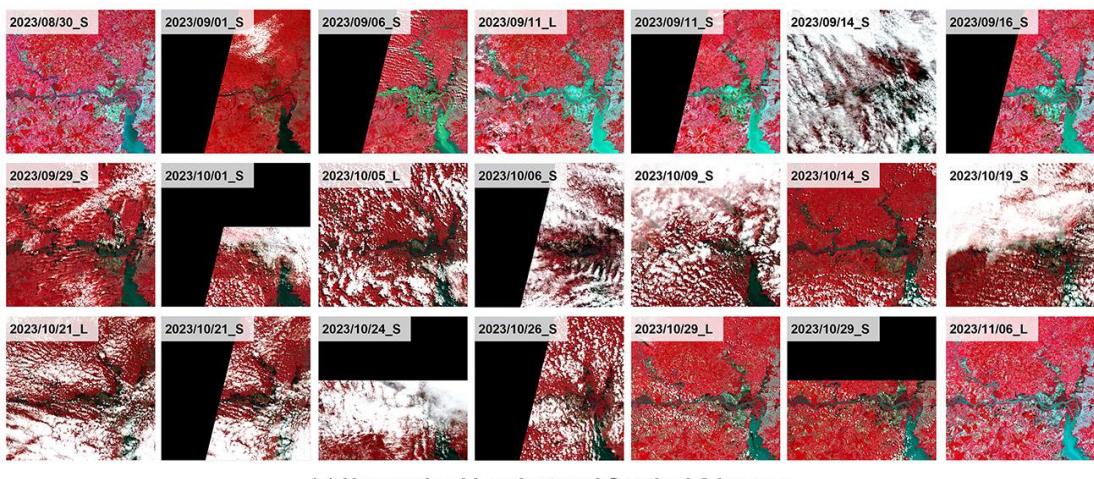


(c) Reconstructed time series flood extent maps

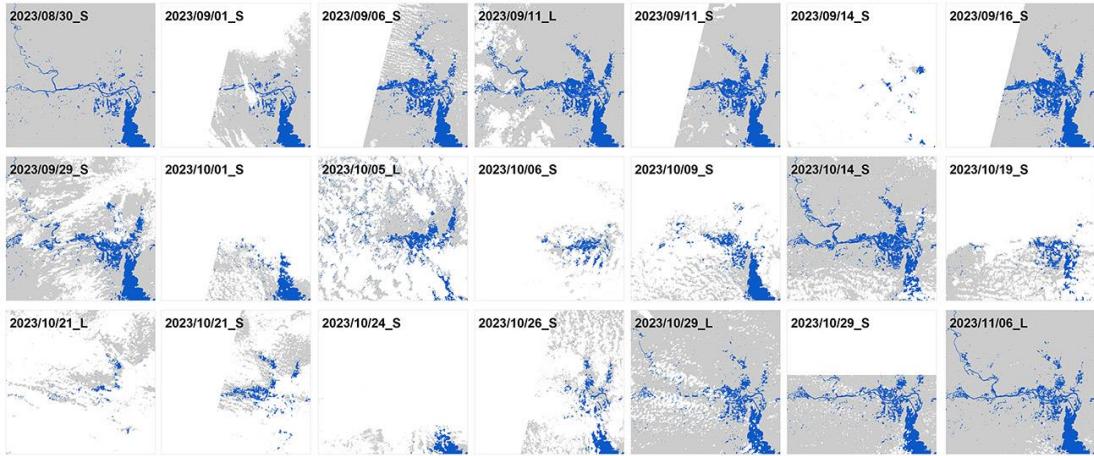
813

814

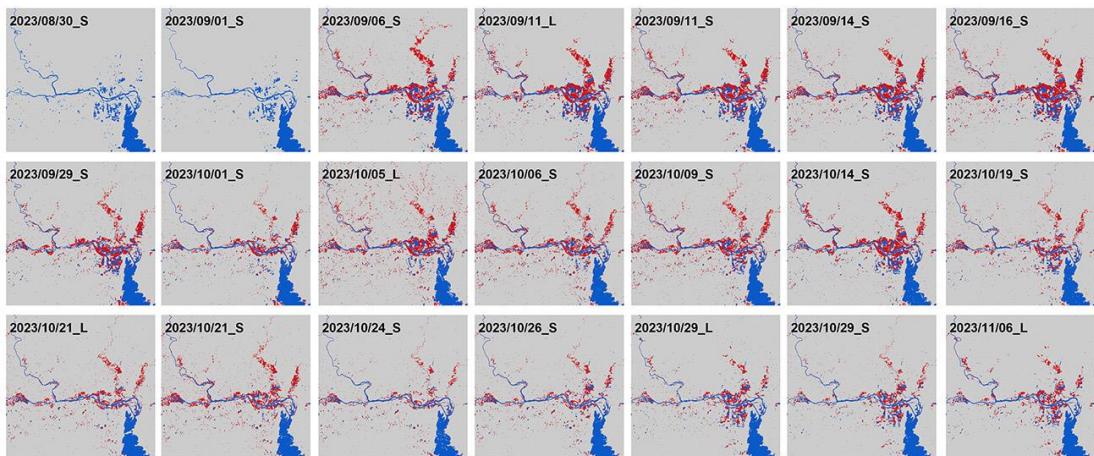
815 **Fig. S2.** Reconstruction results of the proposed method for seamless time series flood extent mapping over Sindh, Pakistan in 2022 flood event.



(a) Harmonized Landsat and Sentinel-2 images



(b) Initial cloud-covered water maps

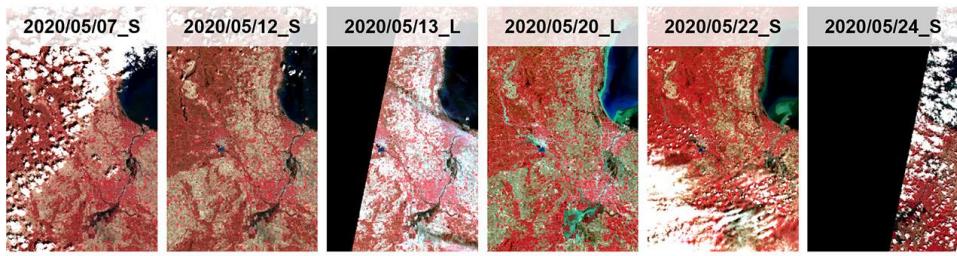


(c) Reconstructed time series flood extent maps

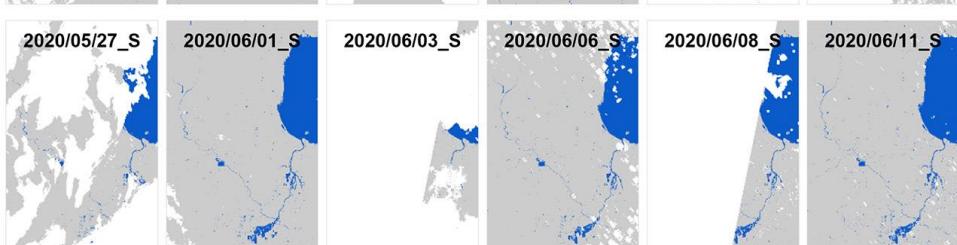
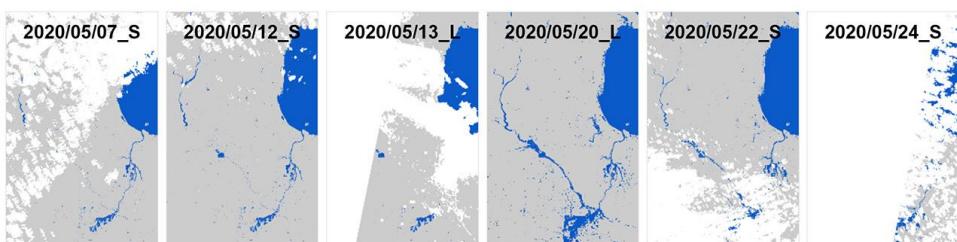
 Cloud/No value Non-water Water / Pre-flood water Floodwater

816

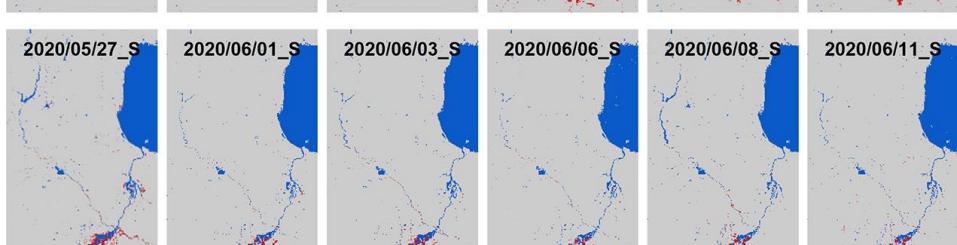
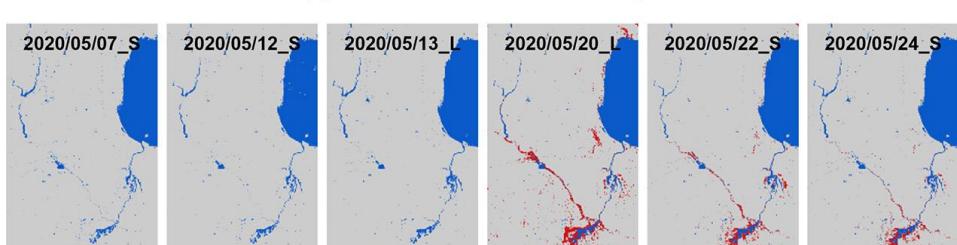
817 **Fig. S3.** Reconstruction results of the proposed method for seamless time series flood extent
818 mapping over Rio Grande do Sul, Brazil in 2023 flood event.



(a) Harmonized Landsat and Sentinel-2 images



(b) Initial cloud-covered water maps



(c) Reconstructed time series flood extent maps

Cloud/No value Non-water Water / Pre-flood water Floodwater

819

820

821

Fig. S4. Reconstruction results of the proposed method for seamless time series flood extent mapping over Michigan, USA in 2020 flood event.