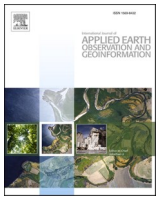
International Journal of Applied Earth Observations and Geoinformation 112 (2022) 102882

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Time-series surface water gap filling based on spatiotemporal neighbourhood similarity

Bingxin Bai [a], Yumin Tan [a,*], Kailei Zhou [a], Gennadii Donchyts [b], Arjen Haag [b], Albrecht H. Weerts [b,c]

[a] Beihang University, School of Transportation Science and Engineering, Beijing 100191, China
[b] Deltares, 2629 HV Delft, the Netherlands
[c] Wageningen University, Hydrology and Quantitative Water Management Group, 6700 AA Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Surface water
Google Earth Engine
Gap filling
Spatiotemporal neighbourhood

ABSTRACT

Optical satellite-derived surface water monitoring is challenging because of the spatial gaps in images caused by clouds, cloud shadows, voids, etc. Here, an efficient method for filling gaps in time-series surface water images is proposed, based on the spatiotemporal characteristics of water. This method utilises the accurately classified historical ternary (gap, water, non-water) or binary (water, non-water) water image time-series and the clear part of the ternary gap water image. Pixels with values of 0 and 1 in the same period water occurrence image are first used to correct the gap water image. The spatial neighbourhood similarity is then calculated as a quality control band for mosaicking the accurately classified historical water images. The final result is generated by replacing the gap pixels with a mosaic image. The proposed method was implemented on the Google Earth Engine, and 93 Landsat 8 top-of-atmosphere (TOA) images were used to verify its validity. Quantitative evaluations were adequate, with a mean accuracy, recall, and precision of 0.98, 0.90, and 0.85, respectively. The proposed method could improve the utilisation of optical remote sensing data and would be applicable to the production of large-area homogeneous surface water time-series and water resource monitoring.

1. Introduction

Detailed long-term maps describing the location and extent of rivers, lakes, reservoirs, and wetlands, recording the time of events such as floods, river migration, lake expansion and retreat, reservoir water storage and discharge, and wetland evolution, can provide insights into the impacts of climate change and water resource management (Pekel et al., 2016; Yamazaki and Trigg, 2016).

There has been much research on the dynamic monitoring of surface water at different scales, including regional (Che et al., 2019; Heimhuber et al., 2016), continental (Mueller et al., 2016), and global scales (Donchyts et al., 2016; Klein et al., 2017; Pekel et al., 2016; Schwatke et al., 2019). At the global scale, commonly used products include the Joint Research Centre (JRC) Global Surface Water (GSW) dataset (Pekel et al., 2016), Global WaterPack (Klein et al., 2017) and Database for Hydrological Time Series of Inland Waters (DAHITI) (Schwatke et al., 2019). The JRC GSW dataset was generated from the entire Landsat archive between 1984 and 2020, and contains maps of the location and temporal distribution of surface water with 30 m spatial resolution. The global WaterPack product is derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) with a spatial resolution of 250 m and a temporal resolution of 1 day. DAHITI provides a variety of hydrological information, of which the monthly status of lakes and reservoirs is based on the Landsat archive and Sentinel-2 from 1984 to 2018.

The above-mentioned surface water data sets are all based on optical images which are vulnerable to clouds. Image synthesis is a common method for generating high-quality products with minimum cloud coverage; however, this method sacrifices time resolution and may therefore miss some short-term hydrodynamic processes (Chen et al., 2013). In addition, few cloud-free images are available in some regions, such as the tropics (Bai et al., 2020), and it is difficult to obtain monthly cloud-free water products by image synthesis. This is one reason why, even though the entire archive of Landsat 5, 7, and 8 has been leveraged, there are still gap pixels in GSW's monthly water products. Therefore, the filling of surface water gaps caused by clouds, cloud shadows or other voids is required to produce spatially continuous high-frequency

* Corresponding author.
E-mail addresses: baibx@buaa.edu.cn (B. Bai), tanym@buaa.edu.cn (Y. Tan), zhoukailei@buaa.edu.cn (K. Zhou), Gennadii.Donchyts@deltares.nl (G. Donchyts), Arjen.Haag@deltares.nl (A. Haag), albrecht.weerts@deltares.nl (A.H. Weerts).

https://doi.org/10.1016/j.jag.2022.102882
Received 16 December 2021; Received in revised form 16 June 2022; Accepted 21 June 2022
Available online 28 June 2022

surface water products.

Multiple gap filling techniques (also called reconstructions in some studies) for surface water monitoring have been developed in recent years. In research on monitoring small reservoir storage from Landsat, Avisse et al. (2017) assumed that the elevation of a frequently immersed pixel may be lower than that of a rarely immersed pixel, and proposed a method to reconstruct missing water pixels based on the updated topography. Based on the assumption that if all uncontaminated pixels with the same water occurrence are classified as water, then all contaminated pixels that have a water occurrence greater than or equal to that value should also be water, Zhao and Gao (2018) proposed a method to correct all types of contaminated pixels in the reservoirs of GSW. In the method proposed by Yao et al. (2019), isobaths (extracted from nearly cloud-free images) were used to recover water areas under contamination through efficient vector-based interpolation. The two studies described above (Yao et al., 2019; Zhao and Gao, 2018) mainly focused on lakes or reservoirs (i.e. single water bodies), and only water status that was not obscured by clouds was considered; therefore, these methods would not work if a single water body was completely covered by clouds. In our previous research (Bai et al., under review), we proposed a Bayes-based gap filling algorithm which considers the surrounding water correlation. Although this algorithm produces fairly

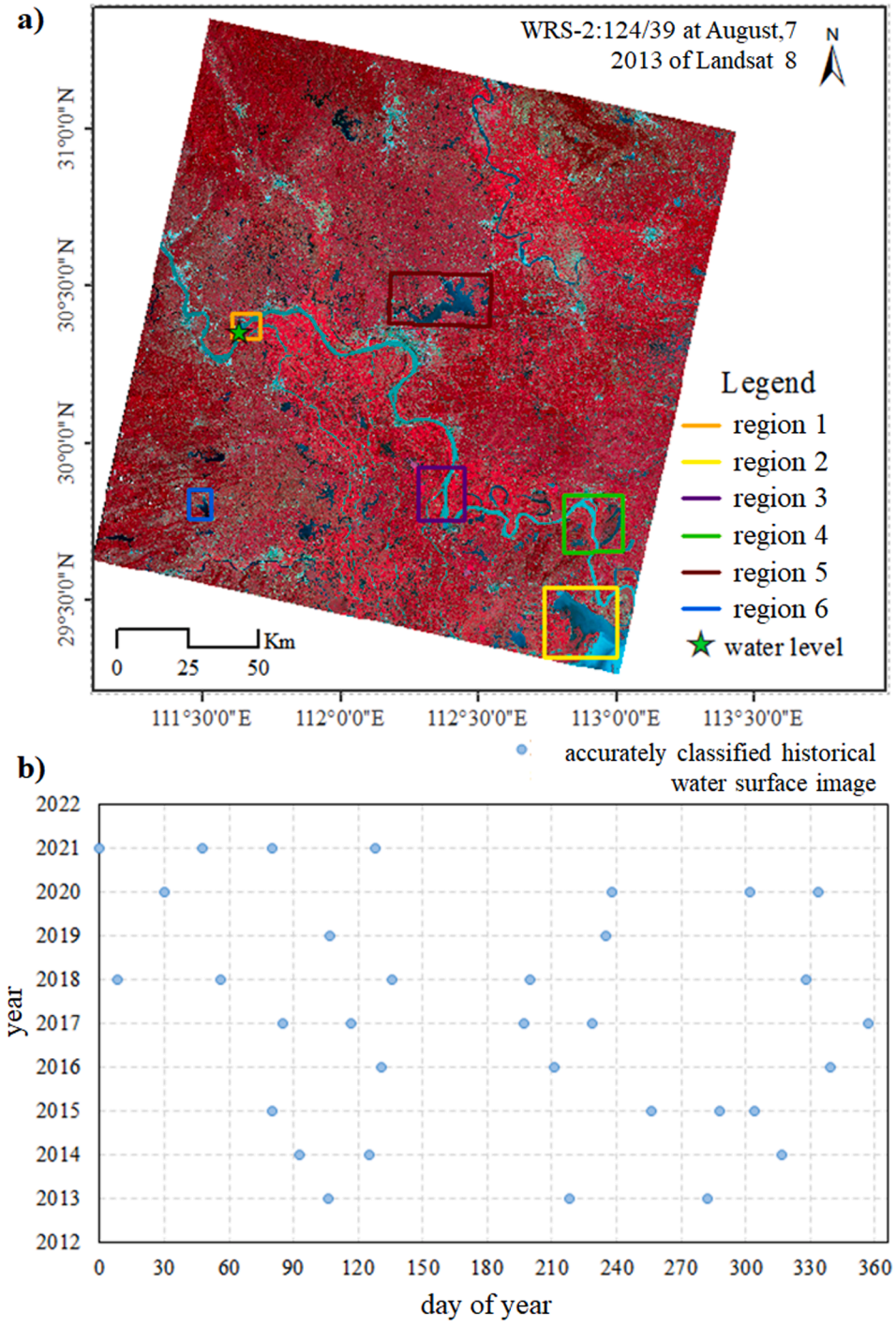


Fig. 1. a) Landsat 8 scene (path/row:124/39) chosen in this paper (RGB rendering: near-infrared, red, green). b) Temporal distribution of the accurately classified historical water image time-series.

good reconstruction results, its application over large areas is limited because of one of its inputs is cloud-free binary (water and non-water) images that are difficult to obtain over large areas.

In this study, an efficient and scalable method is developed to fill the surface water gaps in time-series images caused by clouds, cloud shadows, or terrain shadows at the pixel level based on spatiotemporal neighbourhood similarity. In contrast to the method proposed by Avisse et al. (2017), our method does not depend on terrain data, which are difficult to obtain in many regions. In contrast to the methods proposed by Yao et al. (2019) and Zhao and Gao (2018), our method considers the relationship between connected or disconnected water bodies by considering spatial neighbourhood similarity, which enables the method to be workable even if a reservoir or lake is completely cloud-covered. In addition, in contrast to the method proposed by Bai et al. (under review), the proposed method can use ternary (gap, water, non-water) historical water images as empirical data, which is conducive to execution over a large area. The proposed method was developed on the Google Earth Engine (GEE) (Gorelick et al., 2017), and can be calculated scene-by-scene, which makes it scalable. Furthermore, it is potentially useful for the frequent monitoring of surface water.

2. Data and methods

2.1. Data

The Landsat 8 top-of-atmosphere (TOA) image was adopted in this study for consistency with other surface water mapping studies (Pekel et al., 2016; Tulbure and Broich, 2013). 93 images accessed from the GEE with path 124 and row 39 (see Fig. 1a) in the Worldwide Reference System-2 (WRS-2), acquired from April 2013 to May 2021, were selected.

The river shown in the image scene in Fig. 1a is the Jingjiang River in China, with a width of approximately 2000 m and winding shape; in the bottom right of the scene is part of Dongting Lake, which has a characteristic seasonal water shortage. In addition to rivers and lakes, the land use/land cover types in this study area mainly include reservoirs, ponds, wetlands, paddy fields, mountain forest land, and artificial land.

After cloud masking (Section 2.3.1) and initial water detection (Section 2.3.2), 32 images with high classification accuracy checked by visual interpretation were chosen as the historical empirical data (accurately classified historical water image time-series). The acquisition dates of these images are shown in Fig. 1b. The purpose of visual inspection is to ensure that the classified surface water is as accurate as possible, and to avoid the impact of classification errors on the gap filling algorithm.

2.2. Water level data

One water level data point (https://dahiti.dgfi.tum.de/en/13287/water-level-altimetry/) from the fusion of satellite altimetry Jason-2, Jason-3 missions (a 10-day orbital cycle) (Crétaux et al., 2011; Schwatke et al., 2015), which are available from DAHITI (Schwatke et al., 2015), was used to evaluate the gap-filled results; its location is shown in Fig. 1a.

2.3. Methods

2.3.1. Data pre-processing and water detection

The Landsat 8 TOA images of the GEE were orthorectified. On this basis, cloud and cloud shadow masking was performed using the quality assessment (QA) band which is contained in each Landsat 8 image.

K-means-based unsupervised classification was used to obtain surface water. This requires only one parameter, the number of clusters, which is usually determined according to the number of land use/land cover types in the scene. In this study, we set it to five, based on experience. Seven bands or indices from Landsat 8 were taken as inputs of the K-means: red band, green band, near-infrared band, normalized difference vegetation index (NDVI) (Tucker, 1979), normalized difference water index (NDWI) (McFeeters, 1996), modified NDWI (MNDWI) (Xu, 2006), and normalized difference build-up index (NDBI) (Zha et al., 2003); these are commonly used for water detection (Jiang et al., 2012; Pekel et al., 2014).

The water cluster number is different for different images, which makes it difficult to automatically separate the water from the classified image time-series (Wu et al., 2019). Here, we dealt with this by visually selecting pixels that were indisputably water on the Landsat 8 image and then used the average cluster value of these pixels to extract water.

Finally, every Landsat 8 image was classified as a water image with two categories (0: non-water, 1: water) or three categories (0: non-water, 1: water, and 2: gap pixels). In this study, all pixels covered by clouds, cloud shadows, terrain shadows, and other unrecognised pixels were classified as gap pixels. After the above step, there may be some false positive errors (non-water pixels incorrectly detected as water) or false negative errors (water pixels incorrectly detected as non-water) for water, and these incorrectly classified pixels were further processed as gap pixels by manual masking.

2.3.2. Gap filling

In this paper, we propose a gap filling method based on the following two characteristics of surface water. i) The local spatial correlation of water, that is, the surface water extent, is highly correlated with its surroundings (Zhao and Gao, 2018b; Bai et al., under review). Therefore, when some water is covered by clouds, the extent of the water under the clouds could be calculated from the extent of the surrounding water uncovered by clouds. ii) The temporal similarity of water, which includes two aspects. The first is that the water status is often similar at adjacent moments, that is, when a pixel is marked as water at a previous moment and is water at a later moment, then the pixel is more likely to be water at the current moment. The second aspect is that water usually has obvious seasonal variations, so that water status could be estimated by comparison with the same period (e.g., winter or summer) in history; that is, if a pixel has never been water in the same period historically, then the pixel is likely not water, and vice versa.

The first above-mentioned characteristic of water can be evolved into spatial neighbourhood similarity, which refers to the number of clear pixels with the same category between the gap pixel neighbourhood and historical neighbourhood in this study (Fig. 2). For example, in Fig. 2, in the similarity image, the number 5 indicates that there are five clear pixels in the corresponding neighbourhood of the historical image with the same category as the clear pixels in the neighbourhood of the example gap pixel. By comparing the neighbourhood of the example gap pixel with pixels in its corresponding position in individual historical images, an image of spatial neighbourhood similarity at each historical date can be generated. Then, the category of the gap pixel can be assigned to that of the pixel in the historical image with the greatest similarity.

\Temporal similarity is expressed as the same period water occurrence in this paper. One pixel is more likely not water if it has never been classified as water during the same period in history. Similarly, one pixel is more likely to be water if the pixel has always been classified as water in the same period. In the accurately classified historical surface water image time-series, images with a day of year (DOY) difference within 100 days (approximately one quarter) from the gap date were used to generate the same period water occurrence. This process is illustrated in Fig. 3.

Fig. 4 shows the framework of the proposed gap filling method. There were two inputs: the ternary (0: non-water, 1: water, and 2: gap pixels) gap surface water image, and the accurately classified ternary or binary (0: non-water, 1: water) historical surface water image time-series. The output was a gap-filled surface water image with two classes (0: non-water, 1: water).

The detailed gap filling process was as follows. First, a ternary gap

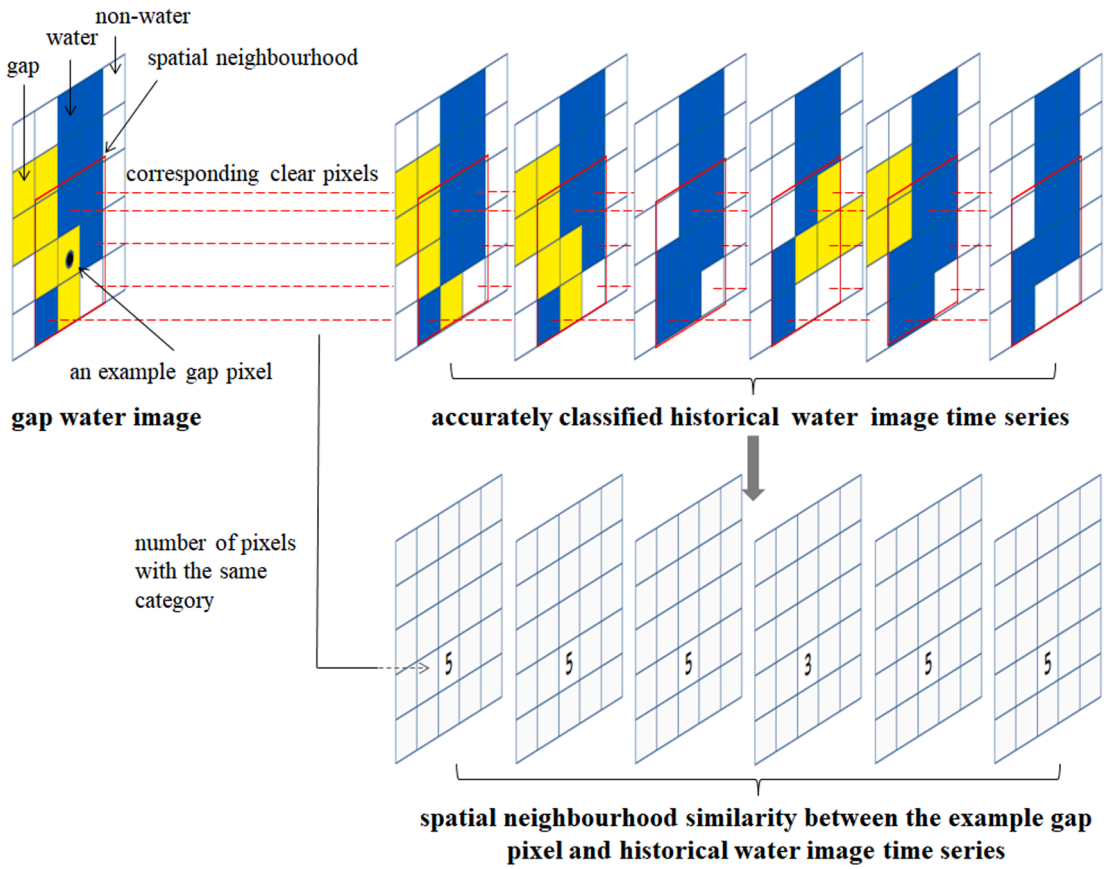


Fig. 2. Schematic of spatial neighbourhood similarity generation.

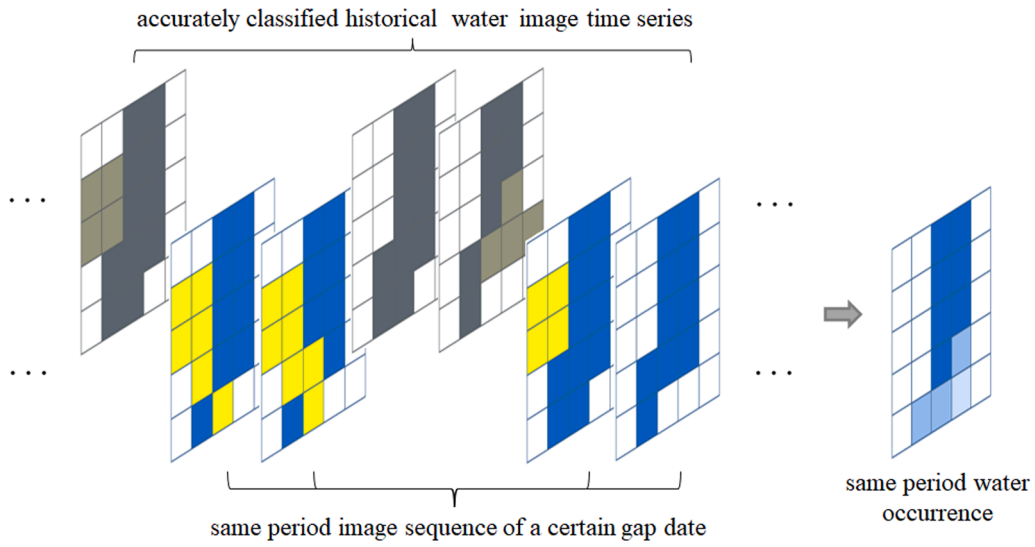


Fig. 3. Schematic of same period water occurrence generation.

water image was preliminarily corrected using the same period water occurrence. Pixels were classified as water if the value of the corresponding position in the same period water occurrence image was equal to 1 (i.e., a pixel in the same period history is always water). Similarly, the pixels were set to non-water, where values were equal to 0 in the same period water occurrence image.

Second, the similarity of the spatial neighbourhoods was calculated. A logical AND operation between the gap water image and each historical surface water image was designed to obtain spatial difference binary images. The spatial difference binary image was then masked with an image in which the water occurrence was neither 0 nor 1. After performing a spatial neighbouring summation operation (using a window with a radius of 70 pixels in this study) on the spatial difference binary image time-series, the spatial neighbourhood similarity image time-series can be generated. The category of the gap pixel was considered to be the same as that of the pixel at that historical moment with the highest value in the spatial neighbourhood similarity image time-series.

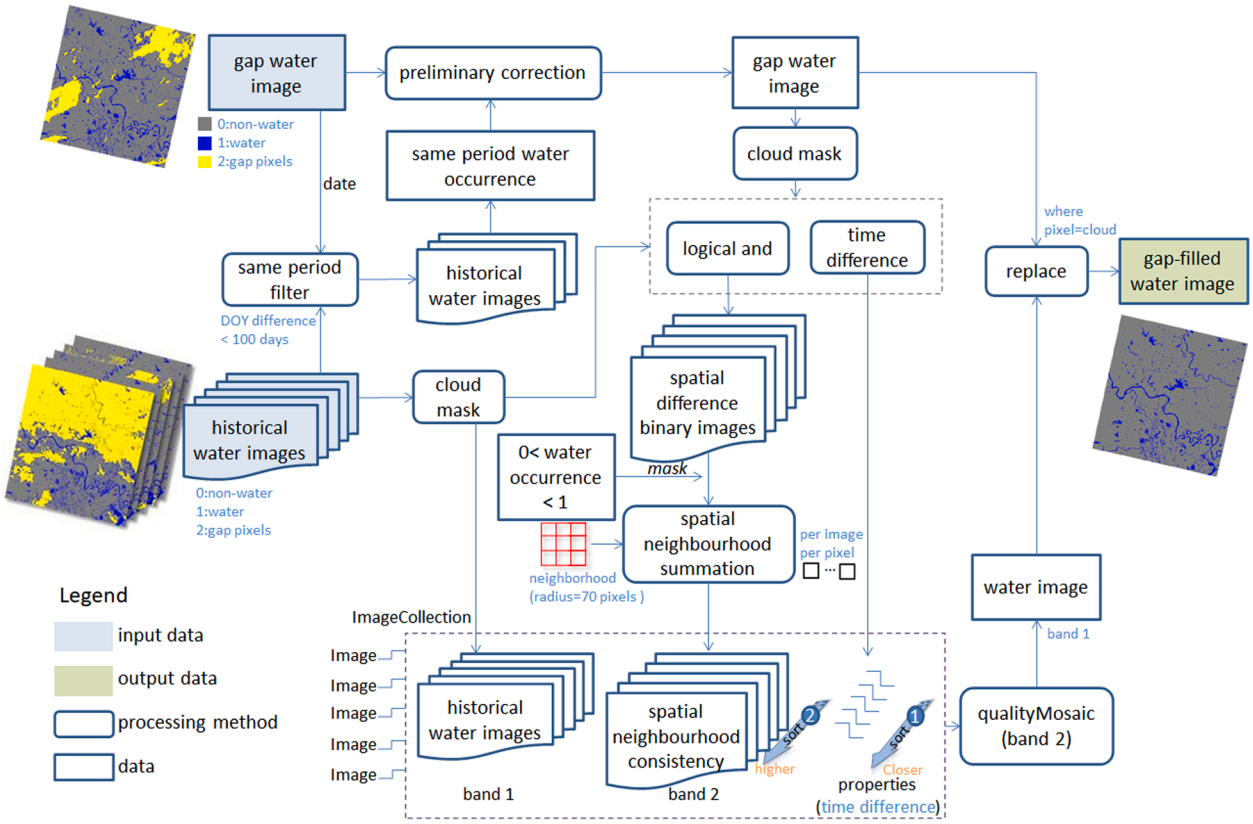


Fig. 4. Graphical workflow of the proposed gap filling method.

The final step was mosaicking. Here, we constructed an ImageCollection object in GEE, where each accurately classified historical surface water image was taken as one band, while the spatial neighbourhood similarity image was another band. The time difference (unit: days) between the historical water image and the gap water image was set as an attribute. Then, the ImageCollection was sorted according to the time difference attribute. After that, the *qualityMosaic* method in GEE was used to mosaic the accurately classified historical water image band according to the spatial neighbourhood similarity band. Hence, each pixel in the mosaic image was a historical pixel on the date which with the highest spatial neighbourhood similarity. When the similarity was too low (usually owing to greater cloud coverage), the similarity calculation was considered unreliable, and other strategies (see Fig. 5) were adopted to fill the gaps. In this study, when the similarity was less than 30 (the value should increase as the radius of the spatial neighbourhood increases, and vice versa), the pixel at the closest date was used instead. Here, if the closest date was greater than 64 days (multiple of the 16-days revisit cycle of Landsat, around 2 months), the same period water occurrence was used to correct the mosaicked image by setting the pixels to non-water class if the occurrence was less than 0.6, or setting to water otherwise. The pixels in the mosaicked image were then used to replace the pixels that remained in the gap class after the initial correction.

The above processing was performed on each image that contained gaps to obtain the gap-filled water image time-series.

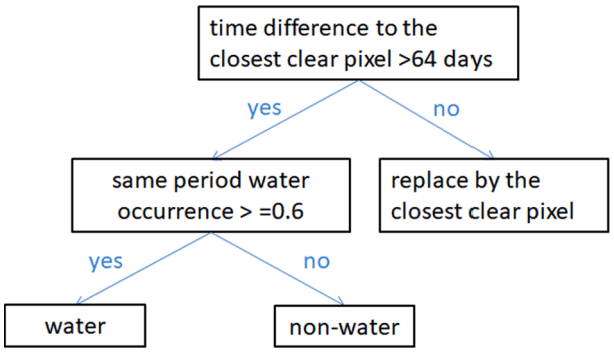


Fig. 5. Strategies adopted when the spatial neighbourhood similarity calculation was unreliable.

2.3.3. Validation

The proposed method was evaluated in two ways. One was to compare the water area time-series with the water levels, which has been used in many studies (Yao et al., 2019; Schwatke et al., 2019).

The second was to construct the gap filling task by manually adding clouds to the cloud-free water image and then comparing the gap-filled water image with the original cloud-free water image. Here, 12 cloud-free water images from accurately classified historical surface water image time-series were selected as the original images. To make the evaluation more comprehensive and objective, gap filling tasks were constructed in three scenarios: i) randomly adding different clouds to each cloud-free water image; ii) adding the same cloud to each cloud-free water image; and iii) adding different clouds to one cloud-free water image. 5000 sample pixels were randomly generated in the cloud areas to calculate the precision, accuracy, and recall indicators (Olson and Delen, 2008; Stillinger et al., 2019). Accuracy indicates how many pixels are correctly predicted for all water and non-water samples, recall indicates how many water pixels in the samples are predicted correctly, and precision refers to how many pixels in the predicted water are actually water.

3. Results

3.1. Gap filling with different strategies

Figs. 6–8 show three situations where different strategies were used to fill gaps in region 2 (marked with a yellow rectangle in Fig. 1). In each example situation, the filling result of the intermediate date was used as the illustration.

Fig. 6 shows an example of gap filling based on spatial neighbourhood similarity. Fig. 6d-f show three temporally continuous surface water images in an accurately classified historical surface water image time-series. There were gap pixels in the surface water image on 6 May 2014, which were caused by the inconsistent spectrum of water in the optical image (Fig. 6b). The gap pixels are not very large and can be filled based on spatial neighbourhood similarity. The gap-filled result is shown in Fig. 6g. By comparing it with the original optical image, we can see that the gap-filled surface water extent has high accuracy. In addition, the surface water extents on the three dates were quite different, and the gap-filled result maintained this change. This would be difficult to achieve using methods such as image synthesis or linear interpolation.

When the spatial neighbourhood similarity was unreliable, as in the gap filling task of 14 September 2015, a different strategy was employed, as shown in Fig. 7. Because the surface water image was substantially covered by clouds, the area inside the orange ellipse in Fig. 7e could not be correctly filled based on the spatial neighbourhood similarity. In this situation, the filling strategy was to use the closest image in the accurately classified historical surface water image time-series for filling. Here, the water image on 16 October 2015 was the closest to the gap image, with a time difference of 32 days; hence, it was used for filling. The water in the ellipse in the gap-filled water image (Fig. 7g) was the same as that on 16 October 2015 (Fig. 7f). The uncertainty of this filling strategy is relatively high, especially for areas with frequent water extent changes and large time differences.

When the spatial neighbourhood similarity calculation was unreliable and the time difference between the gap image and the closest image was greater than 64 days, the same period water occurrence was used to determine the category of the gap pixels. Fig. 8 shows an example. Images from 31 January and 26 August 2020 are two temporally continuous surface water images in accurately classified historical water image time-series. The image from 22 May 2020 was almost completely covered by clouds, so the spatial neighbourhood similarity would be unreliable, and the temporally closest image was from 26 August 2020, with a time difference of 97 days. In this case, the same period water occurrence was used to fill the image. Similarly, in such a

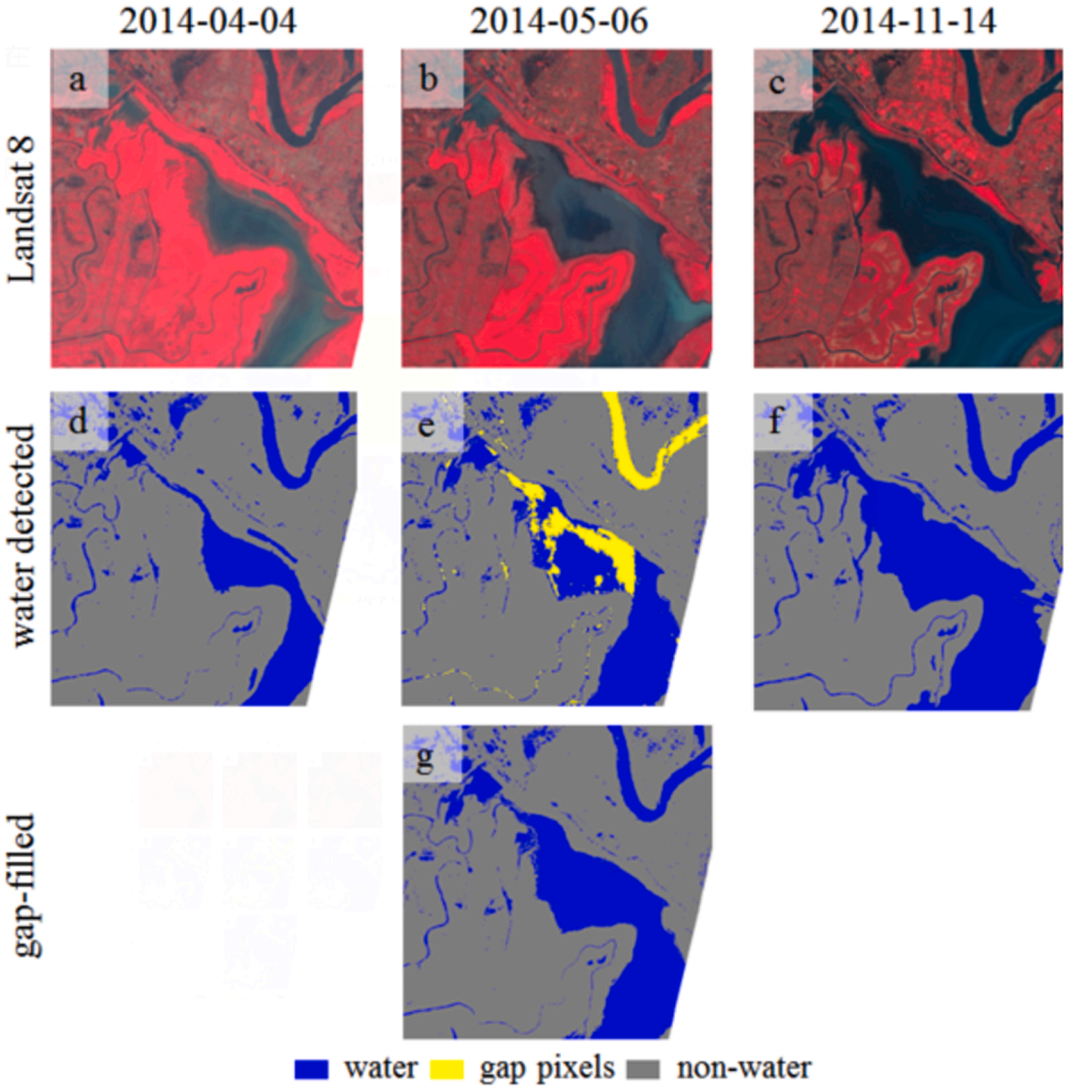


Fig. 6. An example of gap filling using spatial neighbourhood similarity. A-c) Landsat 8 images (RGB rendering: near-infrared, red, green). D-f) Water detected based on (a-c). g) Gap filling result on 6 May 2014. **(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)**

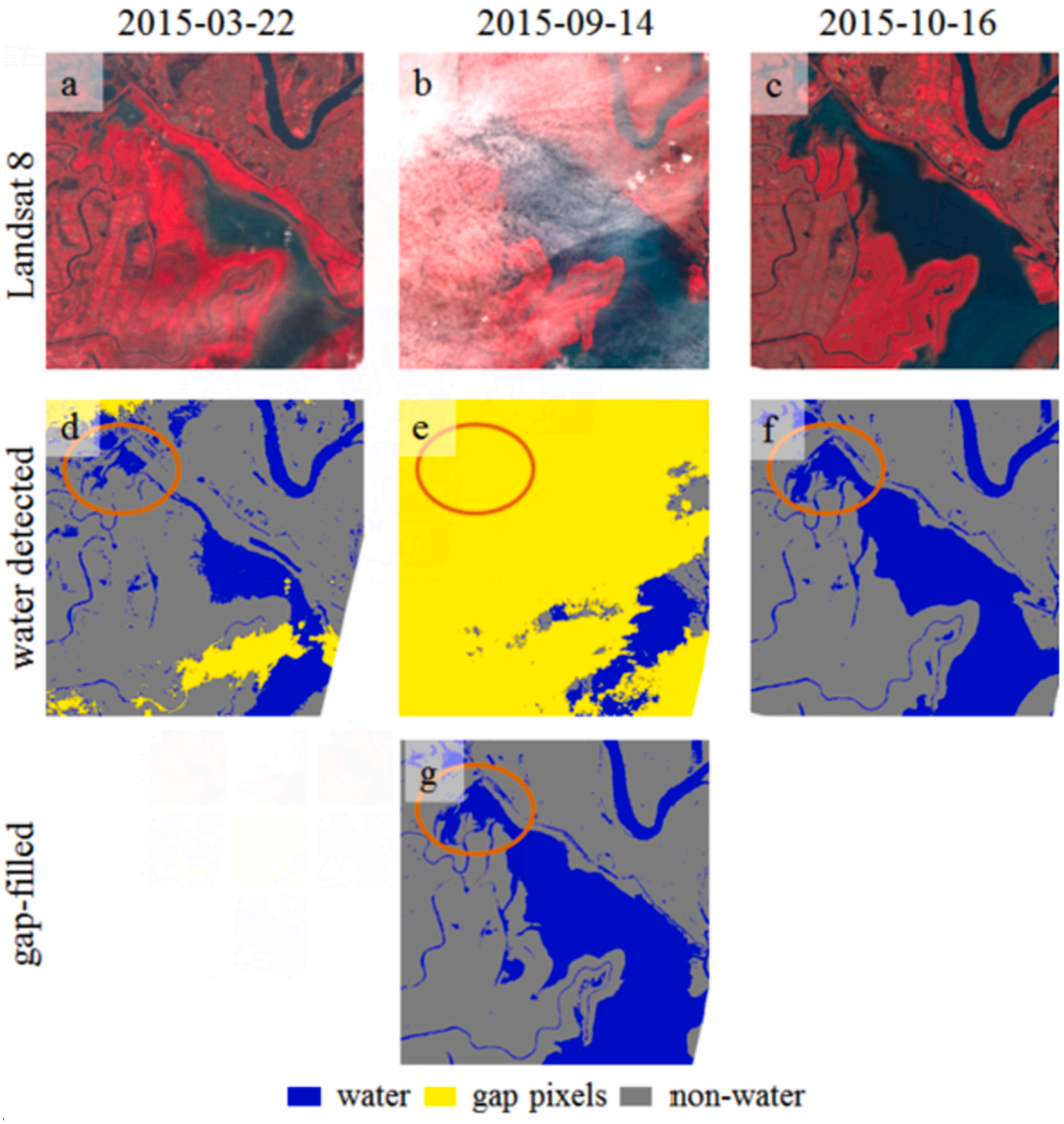


Fig. 7. An example of gap filling using the closest clear pixels. A-c) Landsat 8 images (RGB rendering: near-infrared, red, green). D-f) water detected based on (a-c). g) gap filling result on 14 September 2015. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

case, the uncertainty of gap filling will also be large. As shown by the orange ellipse in Fig. 8b and g, there is a false positive error for water. This example can also be used to illustrate the influence of "closest" parameter on the filling result. If it is set too large, such as 97 days, then the water image on 26 August 2020 will be used as a replacement for the gap pixels. Clearly, in areas with frequent water changes, the uncertainty of gap filling based on this strategy would increases.

3.2. Comparison with JRC GSW monthly water data

The GSW dataset shows high accuracy with less than 5% omission and less than 1% false water detection (Pekel et al., 2016). There are three categories in the GSW monthly water image: water, not water, and no observations (e.g., pixels contaminated by clouds or scan line corrector [SLC] failure).

The water area time-series of the gap-filled surface water images and the GSW monthly water products (example regions 2–4 marked on Fig. 1), are shown in Fig. 9-11. Note that in this study, we only used Landsat 8, whereas GSW monthly water data in the corresponding months were generated from the combination of Landsat 7 and 8. The water area time-series after gap filling showed more obvious time variations than those of the GSW products.

From the water occurrence image, it can be seen that the water extent in region 2 shows frequent changes. From the time-series in Fig. 9a, both the GSW and gap-filled water area show seasonal variations, but the latter shows more obvious variations.

The fill percentage (i.e., the proportion of pixels whose category changes before and after filling) after removing areas where water occurrence equal to 0 or 1 was calculated, and is shown in Fig. 9b. If the clear part of the gap water image is accurately classified, the fill percentage indicates the uncertainty of the water surface to a certain extent. The larger the value, the more corrected pixels there are in the gap image, and the greater the uncertainty of the water image classification, and vice versa.

Fig. 9c shows the gap-filled result for 23 July 2019. At this date, the Landsat 8 image was partially contaminated by clouds, and there were some stripe contaminations in the GSW water image in the corresponding month. It can be seen that the shape of the water after gap filling is essentially the same as that in the GSW product, but there are no stripes.

Region 3 is a section of the Jingjiang River with sandbars in the middle. The water extent in this area has also changed over time. However, compared with lakes, the seasonal variation in this area was not as obvious. Fig. 10a shows that the gap-filled water area time-series

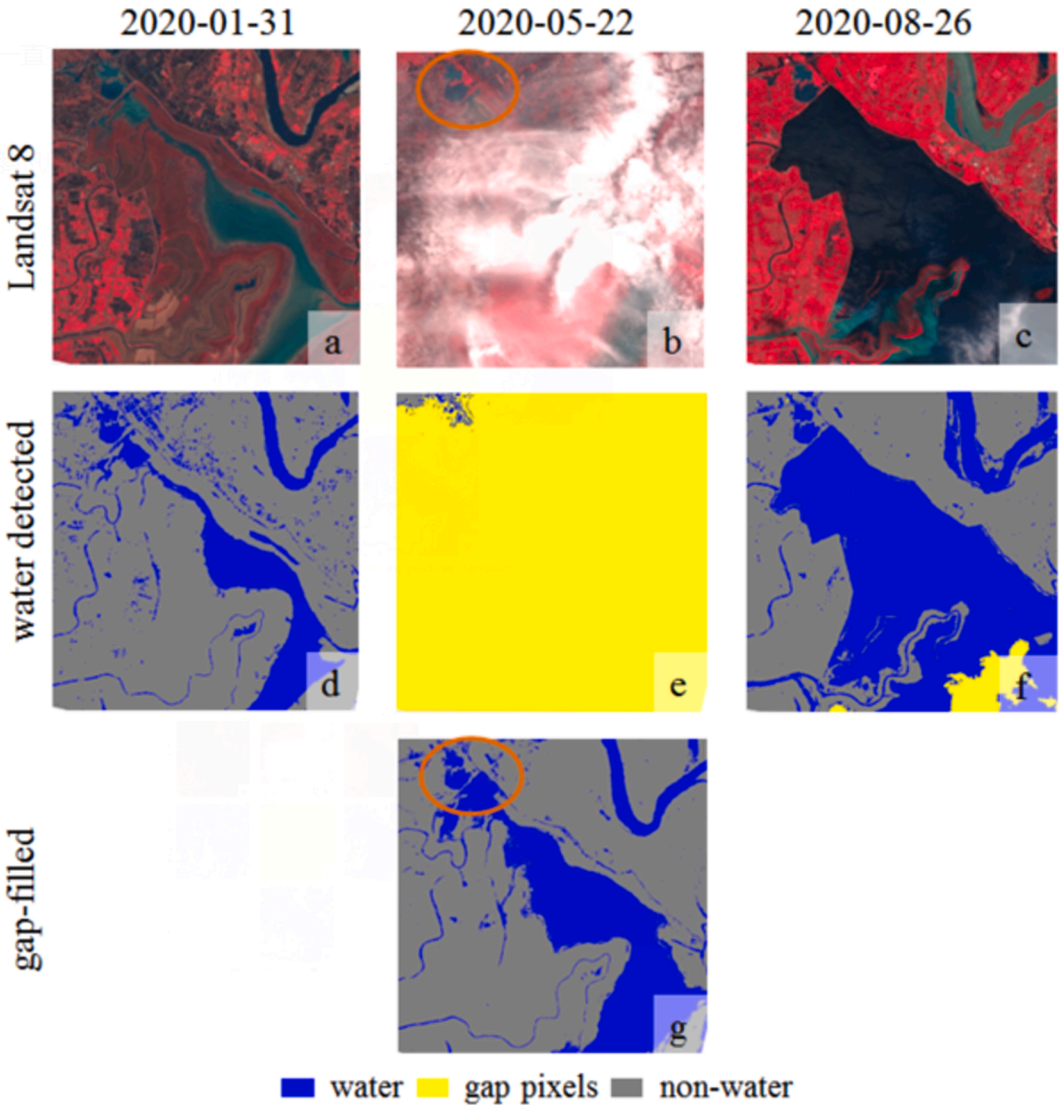


Fig. 8. An example of gap filling using the same period water occurrence. A-c) Landsat 8 images (RGB rendering: near-infrared, red, green). D-f) Water detected from (a-c). g) Gap filling result on 22 May 2020.

has more obvious seasonal characteristics than that of the GSW.

An example (9 July 2014) of the gap filling result is shown in Fig. 10c. Parts of the water bodies are covered by clouds. Because there was no cloud-free Landsat image in this month, only a small part of the water body was identified in the GSW product. The water image based on our newly proposed method showed enhanced spatial details of the river.

Region 4 is also a section of the river. Some parts of the region are often inundated with water. The seasonal and inter-annual fluctuations show good consistency between the gap-filled and GSW water area time-series (Fig. 11), and show an upward trend.

An example of gap filling on 23 February 2018 is shown in Fig. 11c. There are thin clouds in the Landsat 8 image, so we can visualise the real extent of water. It can be seen that, although most pixels are masked into clouds, the water body is reconstructed accurately.

For the other two example regions (5 and 6, marked on Fig. 1), the gap-filled results also show good accuracy (see Figures S1 and S2 in the Supplementary Materials).

3.3. Consistency with water level data

The surface water area time-series before and after gap filling around (orange rectangle in Fig. 1) the water level point (marked by green star in Fig. 1), and water level data are shown in Fig. 12. It can be seen that the water area time-series before gap filling has almost no obvious time variations, whereas after gap filling, it shows obvious seasonal variability and maintains a strong correlation with the water level.

Fig. 13 shows the scatterplot of the water area (before and after gap filling) time-series and water levels obtained by linear interpolation. The coefficient of determination R^2 in a polynomial fitting is 0.62 after gap filling. In Fig. 13a, three points (27 September and 13 October 2014, and 18 October 2016) deviate significantly. Because of the date mismatch between the water image and the water level, and the linear interpolation operation on the water level data, the three points correspond to the two water level dates in Fig. 12. It can be seen that the relevant water levels have a large deviation in the sequence, which may be due to the relatively high water level (there was a flood in this basin in the summer of 2016), or it may have been generated during data processing (the satellite water level data contains some uncertainty, such as system noise). After removing the three points in Fig. 13a, the R^2 improved to 0.79.

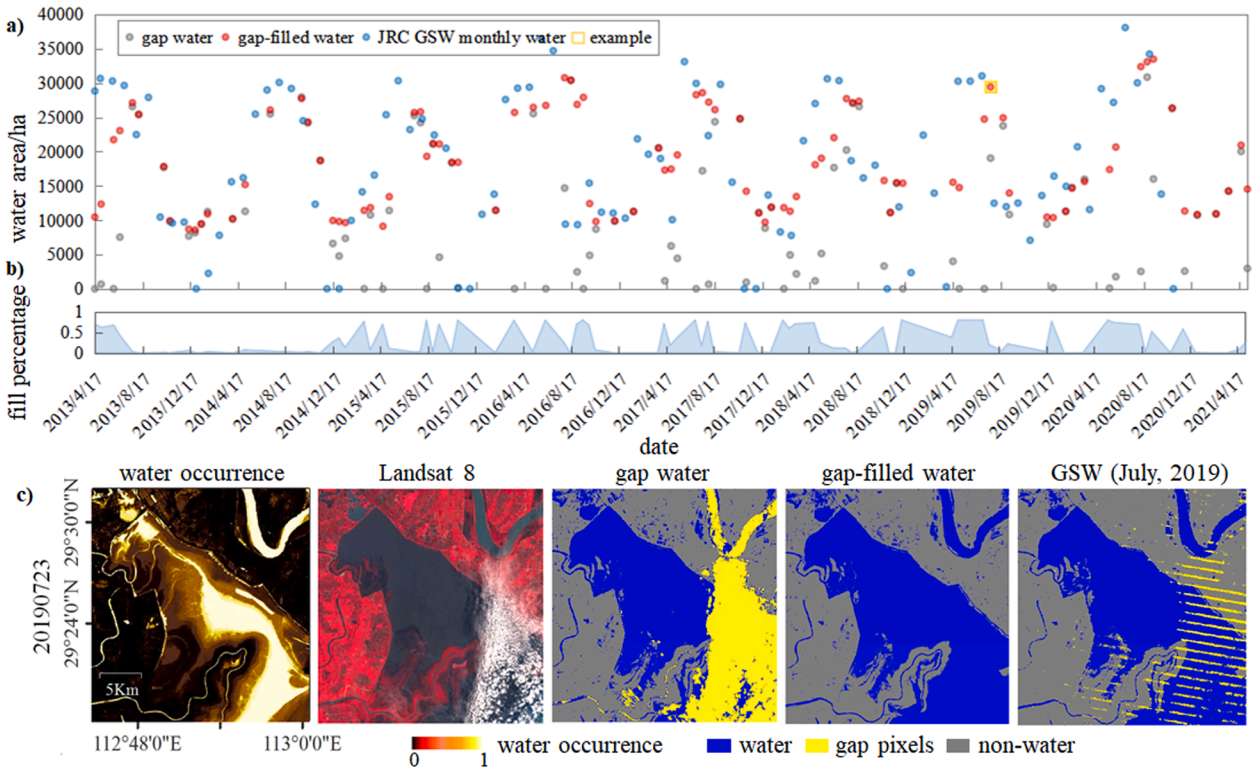


Fig. 9. a) Water area time-series of the yellow rectangle (region 2) in Fig. 1.b) Fill percentage after removing areas where water occurrence was equal to 0 or 1.c) Gap-filled result of the example time indicated in (a).

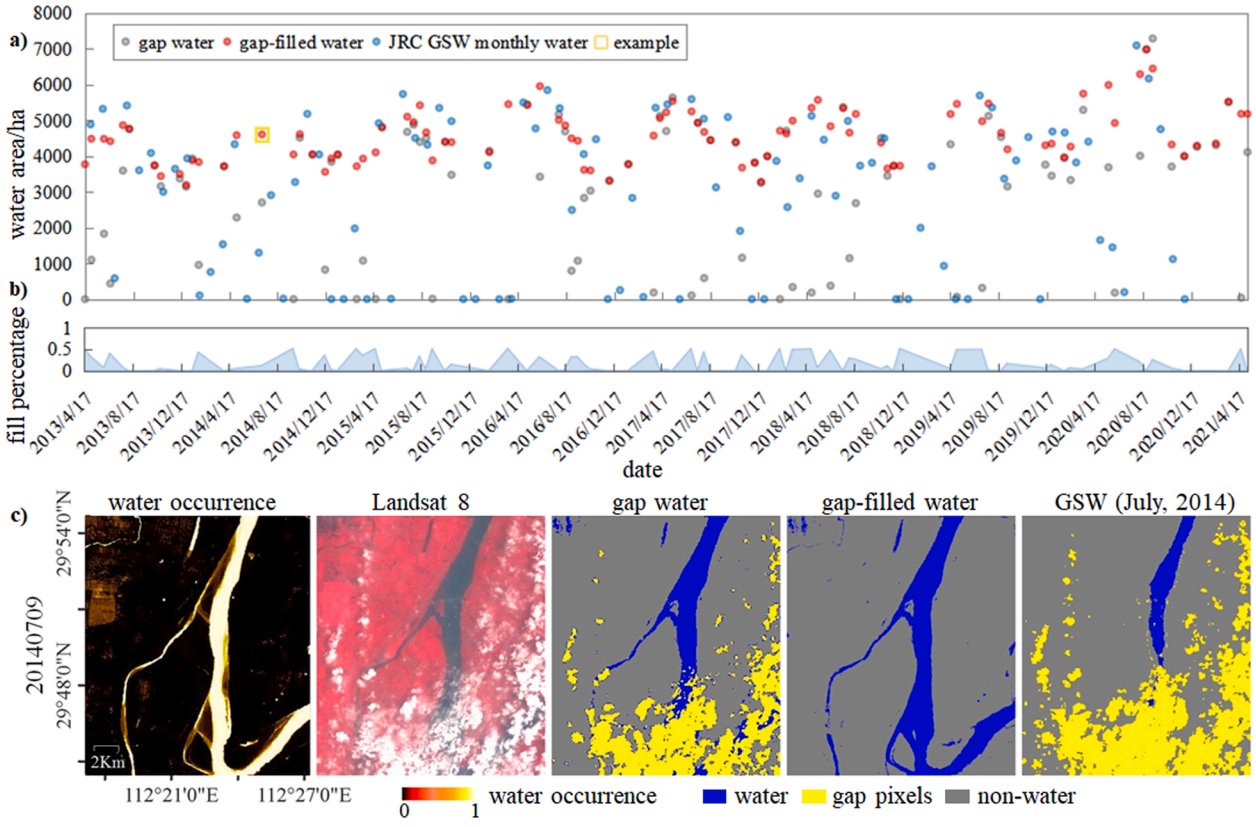


Fig. 10. As in Fig. 9, but for the purple rectangle (region 3) in Fig. 1.

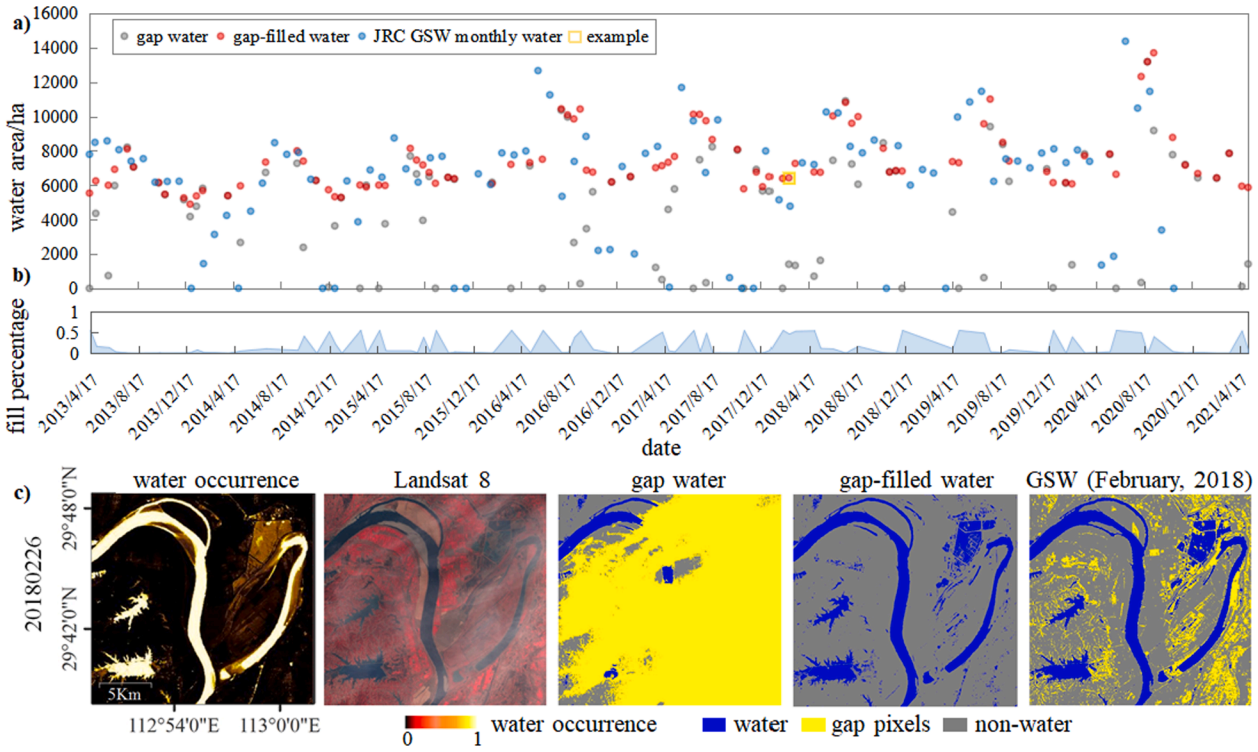


Fig. 11. As in Fig. 9, but for the green rectangle (region 4) in Fig. 1.

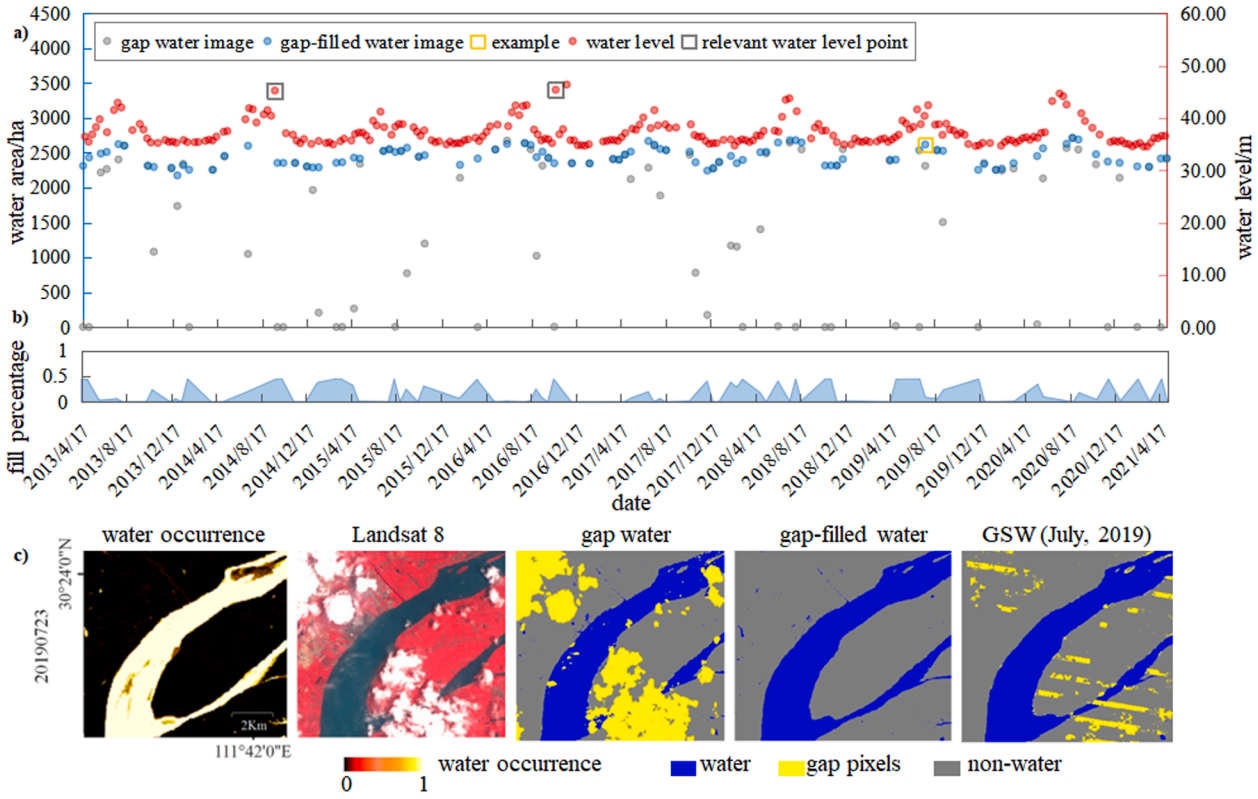


Fig. 12. a) Water area and water level time-series of the orange rectangle (region 1) in Fig. 1. b) Fill percentage after removing areas where water occurrence was equal to 0 or 1. c) Gap-filled result of the example date indicated in (a).

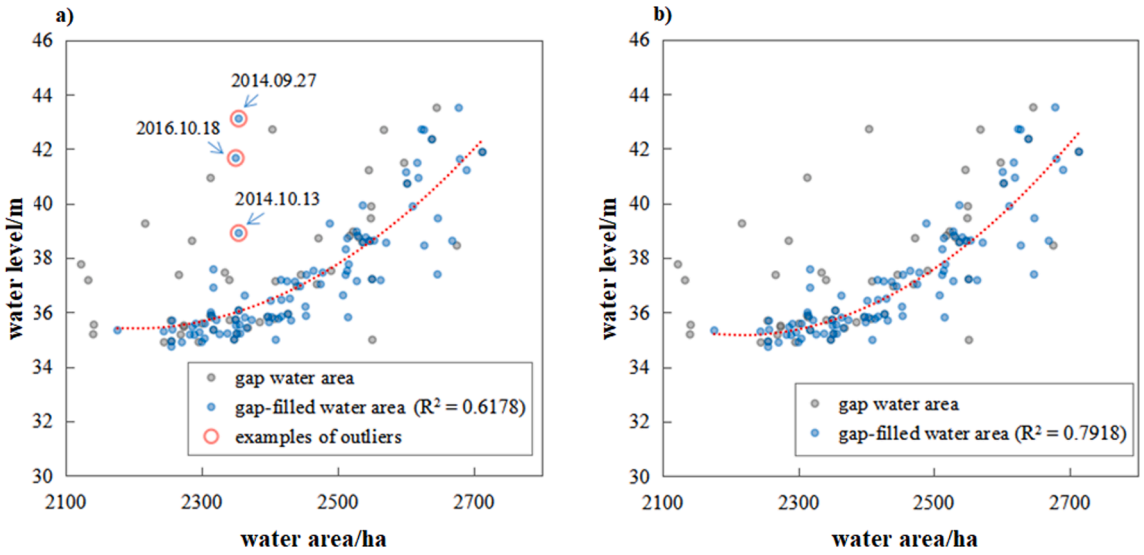


Fig. 13. Scatterplot of the water area and linearly interpolated water level. a) Original. b) Outliers removed.

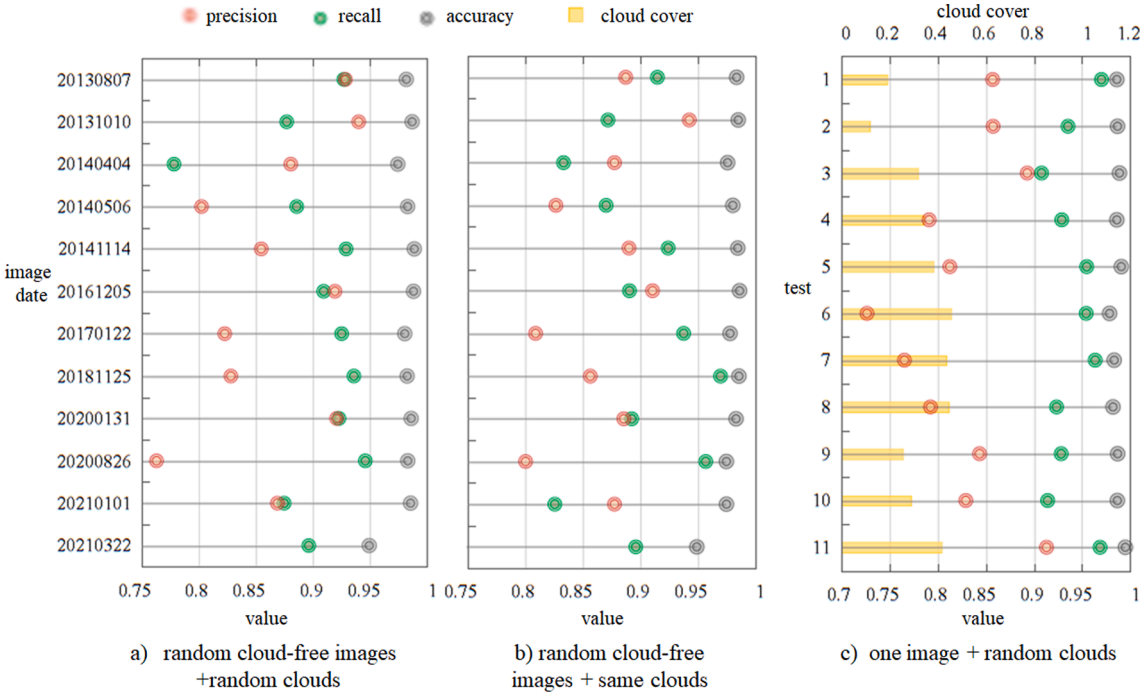


Fig. 14. Evaluation results of the gap filling task constructed using different strategies. a) Different cloud-free images with different clouds added. b) Different cloud-free images with same clouds added. c) Same image with different clouds added.

3.4. Accuracy assessment

The accuracy, recall, and precision of the gap-filled results from the gap water images constructed using different strategies are shown in Fig. 14. It can be seen that the accuracy of the gap-filled results was high in various situations. The average accuracy of gap filling using strategies that randomly add different clouds to different cloud-free images was 0.98, while the average recall and precision were 0.90 and 0.85, respectively. When evaluating the gap filling results of the third strategy, the cloud proportion was also included to show that the gap filling accuracy had no definite relationship with this parameter. This is because the gap filling accuracy is related to many factors, such as the location of the cloud, density of the accurately classified historical water image time-series, and accuracy of the water classification.

At the same time, we should also understand that the samples here are unbalanced, and the evaluation results are relative owing to the strategy of sample selection (Roberts et al., 2017; Wadoux et al., 2021).

We also calculated the root mean square error (RMSE) and between the gap-filled water image and the original cloud-free water image, as shown in Fig. 15. The average RMSE and R^2 of the gap filling task constructed by the strategy of randomly adding different clouds to different images was 0.14 and 0.73, respectively. The smalle the RMSE, the better the gap-filled effect, while the lager theR^2, the better the gap-filled result.

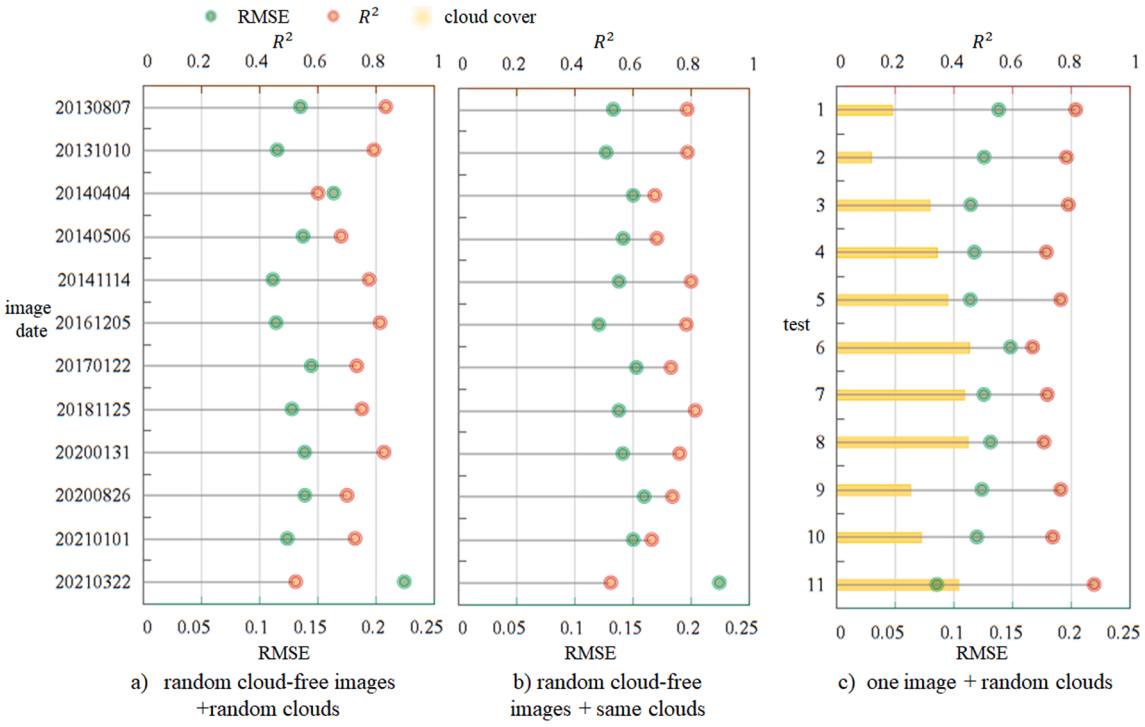


Fig. 15. RMSE and R^2 of the gap filling task constructed using different strategies. a) Different cloud-free images with different clouds added. b) Different cloud-free images with same clouds added. c) Same image with different clouds added.

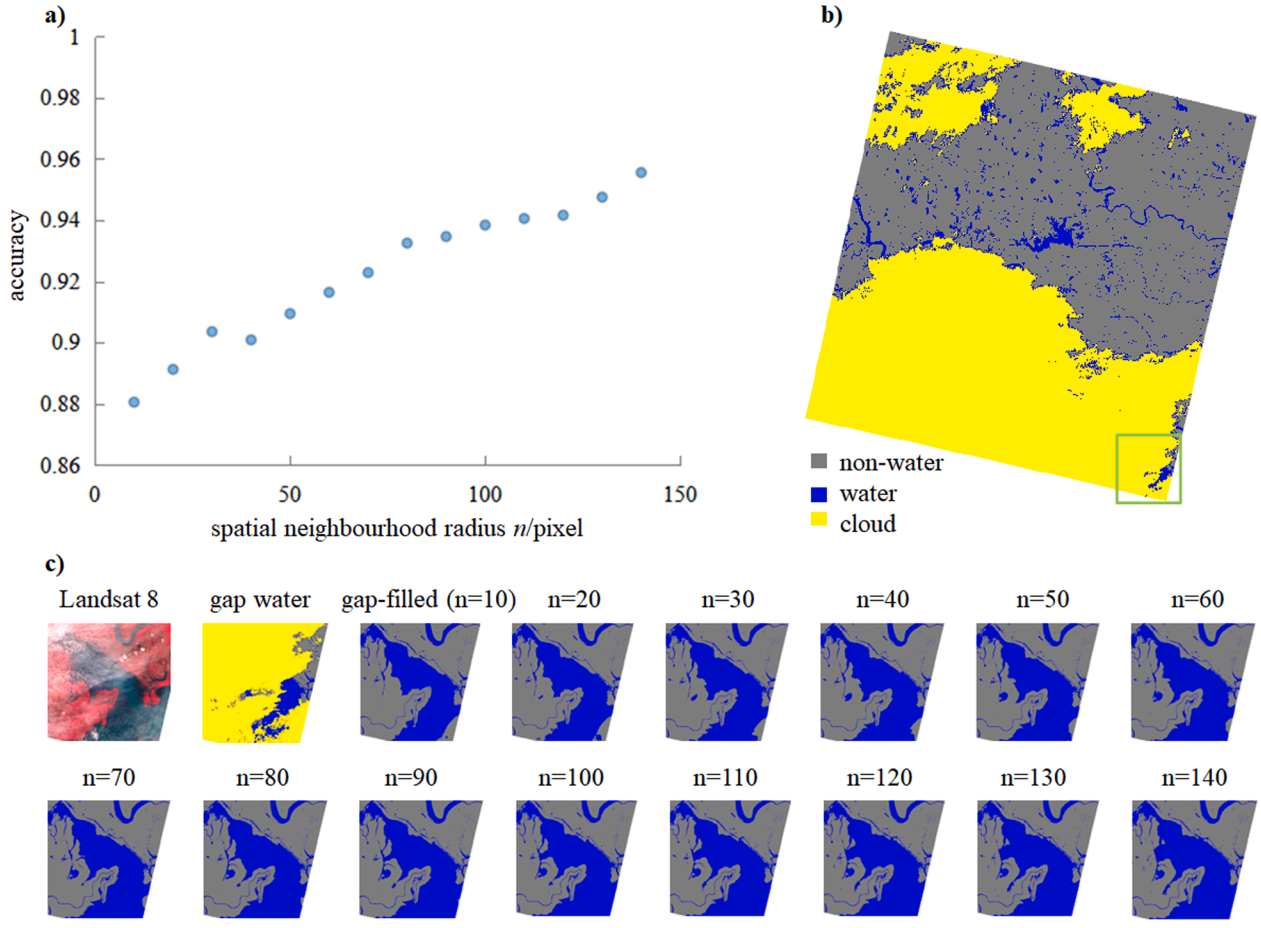


Fig. 16. a) Relationship between the spatial neighbourhood radius and the gap filling accuracy. b) Ternary surface water image from 14 September 2015. c) Gap filling results of the green rectangle in (b) under different spatial neighbourhood radii.

4. Discussion

4.1. The influence of spatial neighbourhood radius on gap filling.

The radius of the spatial neighbourhood has a complex effect on the gap filling accuracy. Generally, the larger the gaps, the larger the spatial neighbourhood radius needed to search for clear pixels for the similarity calculation. For example, we plotted the relationship between the spatial neighbourhood radius and the accuracy of the gap-filled water image on 14 September 2015, as shown in Fig. 16. Because the image was covered by thin clouds, we could visually identify the extent of water; therefore, we manually selected some sample blocks under the cloud, and then randomly generated 2000 sample points for accuracy evaluation. Fig. 16a showed that the filling accuracy increased with an increase in the spatial neighbourhood radius. However, it should also be noted that setting a very large radius reduces the calculation speed and would introduce errors (see Fig. S3 in the Supplementary Material).

When the clouds are small and the water has a narrow shape, a smaller neighbourhood radius could have a high gap filling accuracy, as shown in Fig, S4 in the Supplementary Material.

In general, the selection of an appropriate spatial neighbourhood radius depends on the size of the clouds and water body. When filling time-series images, considering the various situations that may occur, a larger neighbourhood radius should be used. For example, a radius of 70 pixels of Landsat 8 at 30 m spatial resolution was used in this study, and the filling result was acceptable.

4.2. Potential applications

The method proposed in this paper is also suitable for filling stripe gaps that exist in surface water extracted based on Landsat 7 Enhanced Thematic Mapper Plus (ETM+) SLC-off images, and gaps caused by coverage of ice, snow, etc. This provides the possibility of generating spatiotemporally continuous surface water.

4.3. Limitations and opportunities

The proposed method requires an accurate contaminated pixel mask and accurate water classification. In addition, if the cloud coverage is too high and the historical water images before or after the gap date are too far away, the gap filling error would increase.

The proposed method can correct some classification errors (e.g., cloud shadows classified as water) in the gap water image when they occur in areas where the water occurrence equals 0 or 1. However, when errors occur outside these areas, the errors are retained, which affects the filling result by affecting the calculation of spatial neighbourhood similarity.

The requirement of the proposed method on the number of accurately classified historical surface water image time series is to be able to synthesize water images without gaps after contaminated pixel masking, and the more the better. It means that the gap filling accuracy could be improved by adding data from other sensors, such as Landsat 7, and Sentinel-1 and -2.

Recently, Mullen et al. (2021) proposed a cloud-filling approach using a statistical relationship between the inundation status of unmasked pixels and their inundation frequency (IF) value to infer the status of all pixels of the image based on their own IF value. This method has promising applications. In our future research, performance evaluation of different methods will be included.

5. Conclusion

The time-series surface water gap filling method proposed in this study is based on spatiotemporal neighbourhood similarity. After gap filling, the variations of surface water area time-series were consistent with water level data from satellite altimeters (the coefficient of determination R^2 was 0.79), and presented more obvious seasonal characteristics. Quantitative evaluation showed that the average accuracy, recall, and precision were 0.98, 0.90, and 0.85, respectively.

When carrying out surface water monitoring tasks, implementing the proposed method could increase the utilisation of optical remote sensing images and improve the temporal resolution of surface water time-series, making it easier to reveal seasonal and annual variations in surface water. This method is implemented on the Google Earth Engine, which make it easy to conduct time-series calculations and is suitable for large-area applications. It could also be used to fill gaps in existing surface water datasets, such as GSW.

6. Data and code availability:

GEE link of results:
https://code.earthengine.google.com/593bec4b2052f79f95e41c8b9a8f7143.

CRediT authorship contribution statement

Bingxin Bai: Conceptualization, Methodology, Data curation, Writing – original draft, Visualization, Investigation, Software. **Yumin Tan:** Conceptualization, Methodology, Supervision. **Kailei Zhou:** Data curation. **Gennadii Donchyts:** Software. **Arjen Haag:** Software. **Albrecht H. Weerts:** Supervision.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by China's National Key Research and Development Program [grant number 2019YFE0126400] and the China Scholarship Council.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2022.102882.

References

Avisse, N., Tilmant, A., François Müller, M., Zhang, H., 2017. Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas. Hydrol. Earth Syst. Sci. 21, 6445–6459. https://doi.org/10.5194/hess-21-6445-2017.

Bai, B., Tan, Y., Donchyts, G., Haag, A., Albrecht, H.W., 2020. A simple spatio-temporal data fusion method based on linear regression coefficient compensation. Remote Sens. 12, 3900.

Bai, B., Tan, Y., Donchyts, G., Haag, A., Xu, B., Albrecht, H.W., in preparation. Naive Bayes classification-based surface water gap-filling from partially contaminated optical remote sensing image.

Che, X., Feng, M., Sexton, J., Channan, S., Sun, Q., Ying, Q., Liu, J., Wang, Y., 2019. Landsat-based estimation of seasonal water cover and change in arid and semi-arid Central Asia (2000–2015). Remote Sens. 11, 1–15. https://doi.org/10.3390/rs11111323.

Chen, Y., Huang, C., Ticehurst, C., Merrin, L., Thew, P., 2013. An evaluation of MODIS daily and 8-day composite products for floodplain and Wetland Inundation Mapping. Wetlands 33, 823–835. https://doi.org/10.1007/s13157-013-0439-4.

Crétaux, J.F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.C., Nino, F., Abarca Del Rio, R., Cazenave, A., Maisongrande, P., 2011. SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data. Adv. Sp. Res. 47, 1497–1507. https://doi.org/10.1016/j.asr.2011.01.004.

Donchyts, G., Baart, F., Winsemius, H., Gorelick, N., Kwadijk, J., Van De Giesen, N., 2016. Earth's surface water change over the past 30 years. Nat. Clim. Chang. 6, 810–813. https://doi.org/10.1038/nclimate3111.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Heimhuber, V., Tulbure, M.G., Broich, M., 2016. Modeling 25 years of spatio-temporal surface water and inundation dynamics on large river basin scale using time series of Earth observation data. Hydrol. Earth Syst. Sci. 20, 2227–2250. https://doi.org/10.5194/hess-20-2227-2016.

Jiang, Z., Qi, J., Su, S., Zhang, Z., Jiaping, W., 2012. Water body delineation using index composition and HIS transformation. Int. J. Remote Sens. 33, 3402–3421. https://doi.org/10.1080/01431161.2011.614967.

Klein, I., Gessner, U., Dietz, A.J., Kuenzer, C., 2017. Global WaterPack-A 250 m resolution dataset revealing the daily dynamics of global inland water bodies. Remote Sens. Environ. 198, 345–362. https://doi.org/10.1016/j.rse.2017.06.045.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. https://doi.org/10.1080/01431169608948714.

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., Ip, A., 2016. Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. Remote Sens. Environ. 174, 341–352. https://doi.org/10.1016/j.rse.2015.11.003.

Mullen, C., Penny, G., Müller, M.F., 2021. A simple cloud-filling approach for remote sensing water cover assessments. Hydrol. Earth Syst. Sci. 25, 2373–2386.

Olson, D.L., Delen, D., 2008. Advanced data mining techniques. https://doi.org/10.1007/978-3-540-76917-0.

Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. https://doi.org/10.1038/nature20584.

Pekel, J.F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E., Defourny, P., 2014. A near real-time water surface detection method based on HSV transformation of MODIS multi-Spectral time series data. Remote Sens. Environ. 140, 704–716. https://doi.org/10.1016/j.rse.2013.10.008.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-arroita, G., Hauenstein, S., Lahoz-monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography. 40, 913–929. https://doi.org/10.1111/ecog.02881.

Schwatke, C., Dettmering, D., Bosch, W., Seitz, F., 2015. DAHITI- An innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrol. Earth Syst. Sci. 19, 4345–4364. https://doi.org/10.5194/hess-19-4345-2015.

Schwatke, C., Scherer, D., Dettmering, D., 2019. Automated extraction of consistent time-variable water surfaces of lakes and reservoirs based on Landsat and Sentinel-2. Remote Sens. 11, 1–35. https://doi.org/10.3390/rs11091010.

Stillinger, T., Roberts, D.A., Collar, N.M., Dozier, J., 2019. Cloud Masking for Landsat 8 and MODIS terra over snow-covered terrain: Error analysis and spectral similarity between snow and cloud. Water Resour. Res. 55, 6169–6184. https://doi.org/10.1029/2019WR024932.

Tucker, C.J., 1979. Red and photographic infrared lnear combinations for monitoring vegetation. Remote Sens. Environ. 150, 127–150.

Tulbure, M.G., Broich, M., 2013. Spatiotemporal dynamic of surface water bodies using Landsat time-series data from 1999 to 2011. ISPRS J. Photogramm. Remote Sens. 79, 44–52. https://doi.org/10.1016/j.isprsjprs.2013.01.010.

Wadoux, A.M.J., Heuvelink, G.B.M., Bruin, S.D., Brus, D.J., 2021. Spatial cross-validation is not the right way to evaluate map accuracy. Ecol. Modell. 457, 109692 https://doi.org/10.1016/j.ecolmodel.2021.109692.

Wu, Q., Lane, C.R., Li, X., Zhao, K., Zhou, Y., Clinton, N., Devries, B., Golden, H.E., Lang, M.W., 2019. Integrating LiDAR data and multi-temporal aerial imagery to map wetland inundation dynamics using Google Earth Engine. Remote Sens. Environ. 228, 1–13. https://doi.org/10.1016/j.rse.2019.04.015.

Xu, H., 2006. Modification of Normalised Difference Water Index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens. 27, 3025–3033. https://doi.org/10.1080/01431160600589179.

Yamazaki, D., Trigg, M.A., 2016. The dynamics of Earth surface water. Nature 540, 348.

Yao, F., Wang, J., Wang, C., Crétaux, J.F., 2019. Constructing long-term high-frequency time series of global lake and reservoir areas using Landsat imagery. Remote Sens. Environ. 232, 111210 https://doi.org/10.1016/j.rse.2019.111210.

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int. J. Remote Sens. 24, 583–594. https://doi.org/10.1080/01431160304987.

Zhao, G., Gao, H., 2018. Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics. Geophys. Res. Lett. 45, 6092–6099. https://doi.org/10.1029/2018GL078343.