Andrea Titolo*

NORMALIZED DIFFERENCE WATER INDEX FOR CULTURAL HERITAGE. A REPRODUCIBLE METHOD FOR MONITORING FLOODED ARCHAEOLOGICAL SITES

Abstract

This paper presents a reproducible and adaptable methodology for monitoring and analyzing archaeological sites that resurface from fluctuating lake waters. This innovative workflow addresses a growing need for a second-phase salvage approach by utilizing spectral indices extracted by free medium-resolution satellite images combined with zonal statistics to analyze the resurfaced area of archaeological sites. The methodology incorporates cloud-computing, open-source GIS, and programming languages tailored for reuse, ensuring reproducibility and ease of use. Thanks to the results available, the different applications in which the workflow was already tested, and the future possibilities, the paper will show how a tool like the one presented here can fill a gap in the current archaeological workflow regarding emerging cultural heritage.

Keywords

Remote sensing, change detection, R, Google Earth Engine, QGIS

1. Introduction

The construction of dams and subsequent creation of artificial reservoirs is one of the significant threats to archaeological sites in Western Asia, together with conflicts and natural catastrophes.[1] For decades, local institutions, governments, and international organizations have sponsored archaeological rescue project prior to the submersion of archaeological sites.[2] Even if a large number of projects took place, recent studies highlighted how the knowledge of potential cultural heritage loss due to dam construction is less than optimal.[3] Most importantly, it has been pointed out that there is a growing need for a post-flooding assessment of cultural heritage and a continuous monitoring of archaeological sites. This need arises from the knowledge that sites, even if submerged, are not necessarily lost for good.[4] Multiple pieces of evidence show how cultural heritage can re-emerge due to cyclical or episodic reservoir water fluctuations, urging for a need for so-called second-phase salvage approaches.[5]

This paper wants to address the topic by describing a reliable, adaptable, and reproducible tool and workflow for monitoring the resurfacing of archaeological sites at a larger scale on a multitemporal basis, bringing examples from two recent case studies that relied on it: the Mosul Dam and the Tishreen Dam. The workflow was created to answer the following research questions:

- How can we identify which sites have been affected by water level fluctuations?
- Can we identify sites more at risk of damage and erosion among those affected by water fluctuations?
- How can we document and monitor these sites in the long term and in the future?
- Can we make this process reproducible and adaptable to multiple case studies?

This workflow was initially tested in an early-stage shape on three different dams in Iraq in order to check whether the issues above could be tackled efficiently or not.[6] The second iteration, presented here, is considered to be more mature, as it draws from the experiences (both in code quality and methodology) of the Mosul Dam[7] and the Tishreen Dam[8] applications, and is now integrated in a still in development tool with a broader scope, such that a new presentation of the computational methods underpinning it seemed necessary. While the two case studies are described in separate papers,[9] the detailed methods behind the workflow will be presented in detail here. A summary of the outcomes that the workflow can deliver will also be described (based on the aforementioned applications), followed by a discussion on the workflow's potential, usefulness, and reproducibility.

---

* Dipartimento di Studi Storici, Università di Torino.

[1] On the conflicts generated by water management, see e.g. Lupu 2001; Gleick 2014; Al-Ansari 2016. For archaeology and dam construction, see e.g. Luke, Meskell, 2019; Marchetti *et alii* 2019; Eidem 2020.

[2] Eidem 1999; Wilkinson 2004.

[3] Marchetti *et alii* 2018; 2019.

[4] See e.g. Eidem 2015; Eidem *et alii* 2019; Eidem 2020; Sconzo, Qasim 2021; Sconzo, Simi, Titolo 2023.

[5] See Sconzo, Simi, Titolo 2023; this volume; Sconzo, Simi this volume.

[6] Titolo 2021.

[7] Sconzo, Simi, Titolo 2023.

[8] Sconzo, Simi, Titolo this volume.

[9] Sconzo, Simi, Titolo 2023; this volume.

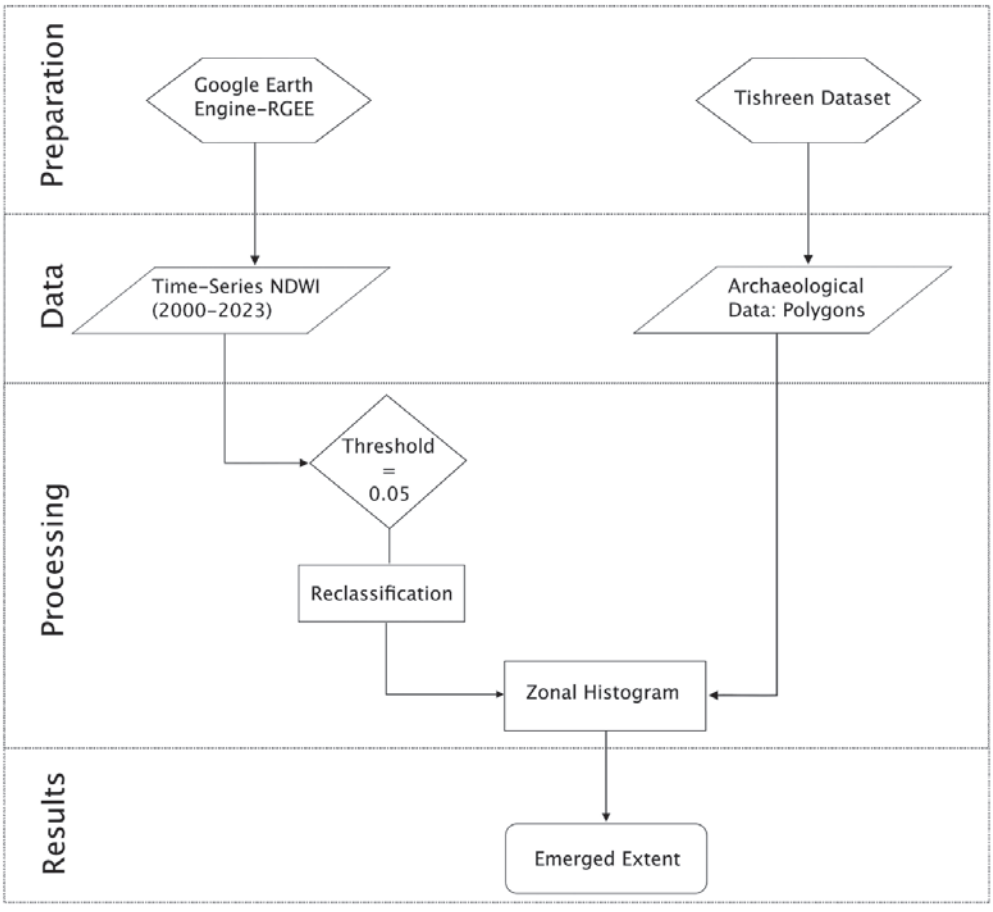

Fig. 1 - Conceptual workflow for extracting resurfaced area information from archaeological sites polygons.

2. The Workflow

While the main objective of this workflow, namely to distinguish water from other surfaces on satellite images, seems relatively simple, the reality is quite different. Various methods are listed in the scientific literature, and each tends to account for different environmental situations (see below). Multiple approaches exist because the identification relies on how the water absorbs the sunlight at different wavelengths and frequencies of the electromagnetic spectrum, and the surrounding landscape highly influences the results. The idea behind the workflow is rather straightforward: remote sensing offers a method to distinguish between water and non-water surfaces, which is usually aimed at delimiting water surfaces, but for archaeological sites, this is not always enough and will merely show if a site is under the water or not. However, once we have this information, we can revert the question and quantify the resurfaced site area based on available information. In order to obtain these results, several steps are required and described below (see also Fig. 1), namely:

- The choice of appropriate method of delineating water on satellite images.
- The application of a change detection technique, which is divided into:
  - Choice of suitable satellite images and a timeframe for the analysis.
  - Processing of the satellite images using a change detection algorithm.
  - Post-processing to remove noise and false positives/negatives.
- The calculation of the percentage of the resurfaced area and the creation of an output dataset.

2.1 *Spectral Indices*

Several alternatives have been proposed for extracting water surfaces from satellite images.[10] One is density slicing. This single-band method relies on the Near-Infrared (NIR) band to obtain information about

[10] Du *et alii* 2014; Li *et alii* 2013, for a more recent application see Pang *et alii* 2024.

water surfaces.[11] The downsides of density slicing approaches are their sensibility to shadows (either from reliefs, clouds, or urban areas) and the resulting risk of introducing errors. Another method is the Tasselated Cap Transformation, which compresses all the information into three bands (greeness, brightness, wetness), thus enhancing spectral band information. However, this process tends to overestimate the presence of water, and it does not behave reliably when shadows are present.

A more common method involves the use of spectral indices (SI). SI are widely used in a myriad of applications, both in archaeology and not.[12] SI are combinations of two or more bands, and they indicate the relative abundance of the element of interest inside the image. SI have some advantages over the methods mentioned before, as they are usually easier to obtain and tend to lead to more robust results.[13] There are different indices available, but for the workflow, the Normalized Difference Water Index (NDWI)[14] was chosen, as it is generally agreed that it performs better at delineating water masses.[15] In particular, a variation of the original known as mNDWI (Modified NDWI)[16] was employed, because, thanks to the use of the Short-wave Infrared Band instead of the original Near Infrared one, it can perform better than the regular NDWI (see below). The mNDWI is expressed by the following formula (in round brackets are the names of the spectral bands used).

$$mNDWI = \frac{(Green) - (SWIR)}{(Green) + (SWIR)}$$

This formula was chosen because it can avoid registering urban elements as water surfaces (having a similar spectral response, they can often get mixed), and it performs better in areas without dense vegetation. While the formula is sensible to snow,[17] which can sometimes be wrongly registered as water, the study areas in which this workflow has been applied are unlikely to be extensively covered in snow, or at least in a measure that would hamper the identification of water. These choices worked for the Iraqi and Syrian dams on which the workflow was tested (see below). However, the interplay between natural and anthropic elements and their effect on satellite images must always be carefully evaluated before proceeding.

The equation will generate a singleband image with pixel values between -1 and 1. The interpretation of this index is based on an arbitrary conventional threshold of 0, meaning that all pixel values above zero are considered water surfaces, and below zero are considered land or, more correctly, non-water surfaces. While the NDWI is used to identify water bodies, if the extension of an archaeological site is known, then it can be used to check if the surface is underwater or not by simply counting the number of land and water pixels inside the site area. In fact, the difference between an emerged and a submerged archaeological site on satellite images will be whether pixels can be identified as representing water or not.

2.2 *Change Detection Technique*

2.2.1 *Choice of Satellite Images and Timeframe*

At its core, change detection comprises a series of methods of identifying and quantifying changes in the pixels on a raster image (drone, satellite, or aerial). The aim is to detect changes in the same geographical area over time by comparing pixel values of each image and applying change detection algorithms according to the application's need. There are a variety of change detection methods,[18] but they all follow a similar workflow: pre-processing of the chosen satellite images to remove noise and inaccuracies, application of the change detection algorithm, and post-processing.[19]

After deciding how to extract water surfaces (see above), a choice regarding which satellite images to use has to be made. Since the idea is to be as affordable and reliable as possible, the workflow relies on freely available medium-resolution images, i.e., Landsat and Sentinel-2 images. While their medium resolution might not be ideal for smaller-scale analyzes, the advantages in accessibility, costs, processing time, and disk space vastly outweigh the downsides.[20] Moreover, many artificial lakes were constructed years before most commercial satellites came into operation. Thus, Landsat images are almost a necessary choice for a multi-temporal analysis. Images were taken at the Bottom-of-Atmosphere reflectance to minimize the impact of atmospheric reflectance on the resulting spectral index.[21] There is also a need to define a time frame and interval to acquire the satellite images. In the first iteration of the workflow, the analysis was conducted with one image per year from

---

[11] For an archaeological application see Marchetti *et alii* 2019.

[12] see e.g. Agapiou, Hadjimitsis, Alexakis 2012; Calleja *et alii* 2018; Fuldain González, Varón Hernández 2019; Kalayci *et alii* 2019.

[13] Du *et alii* 2014.

[14] McFeeters 1996.

[15] Du *et alii* 2014; Qiao *et alii* 2012; Li *et alii* 2013; Acharya *et alii* 2016.

[16] Xu 2006.

[17] Acharya *et alii* 2016.

[18] For a comprehensive review, see e.g. Lu *et alii* 2004; Cheng *et alii* 2023.

[19] Cheng *et alii* 2023, 3-4.

[20] In particular, regarding issues and costs for archaeologists and remote sensing specialist working in SWA, see Rayne *et alii* 2020 with related bibliography.

[21] Agapiou *et alii* 2011; Hadjimitsis *et alii* 2010.

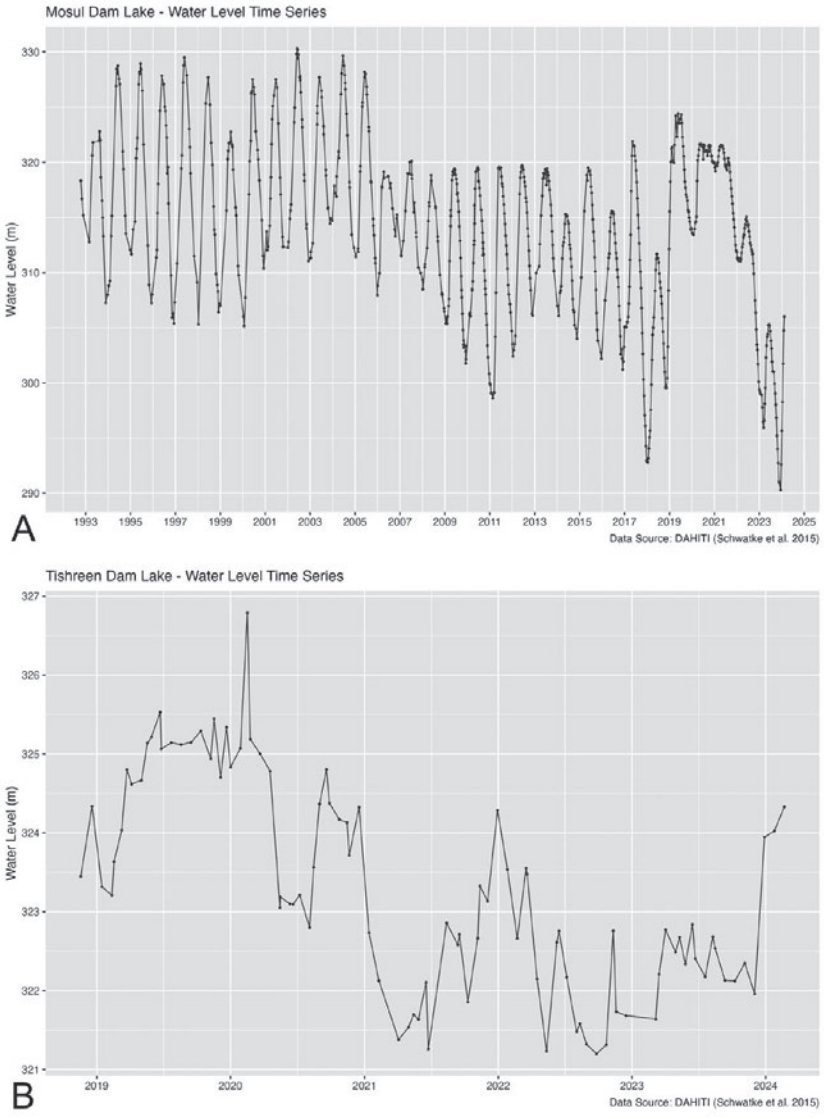

Fig. 2 - Example of information gathered from the DAHITI. A) Mosul Dam water level plot. B) Tishreen Dam water level plot.

the filling of the reservoir to the present day. However, this choice might be limiting and not helpful in highlighting water fluctuations within the same year. For this reason, the workflow relied on quantitative data of reservoir water level variations provided by the Database for Hydrological Time Series of Inland Waters (DAHITI).[22] The DAHITI registers water level and other lake characteristics through remote sensing over several years (Fig. 2). Data on the water level were leveraged in order to link the results of the analysis to a precise value. Thus, the time frame was usually limited to the available coverage of the DAHITI for our study area. Since the DAHITI records water levels throughout the year, but not necessarily for all the months, a decision was made to extract minimum and maximum water levels for each year and to limit our analyses to two instances per year. This means that in the current workflow, we take two images per year, one for the months with minimum water level and one for the maximum water level based on the DAHITI data.

2.2.2 *Processing of Satellite Images*

Usually, the larger the time period, the larger the number of satellite images to utilize. Manual processing of all these images would take a long time; thus, the workflow relies on cloud computing to quickly apply different processes over multiple images. The most commonly used platform for remote sensing

[22] Schwatke *et alii* 2015.

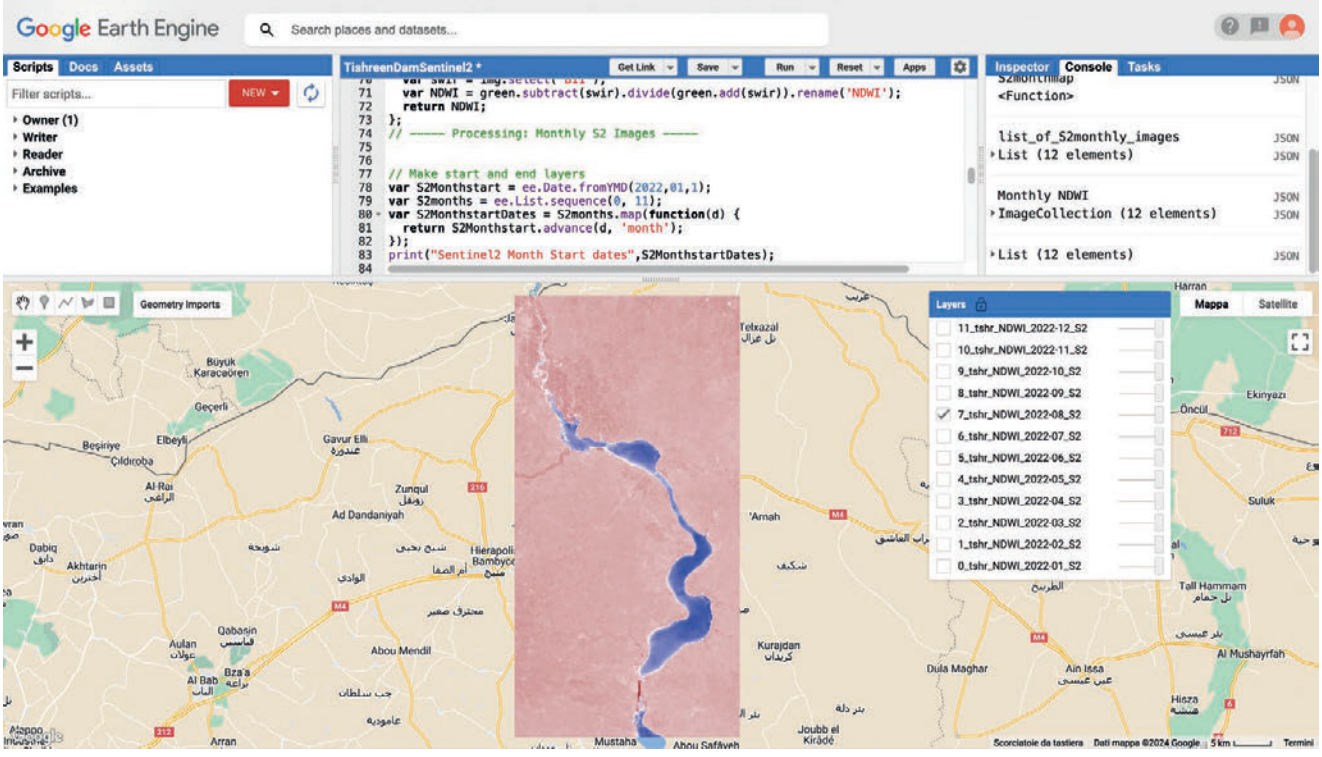

Fig. 3 - Google Earth Engine Code Editor interface and NDWI composites.

cloud-computing is Google Earth Engine (hereafter GEE).[23] The platform, launched in 2017, offers petabytes of geospatial data and allows the quick application of different and complex algorithms. The platform has seen a growing number of easily accessible web apps; however, a basic understanding of Javascript is still required for specialized applications, to write the code necessary for the analyses. Two scripts were written to generate NDWI images for Landsat and Sentinel-2 data.[24] The NDWI images were generated as monthly composites by applying first a cloud filter to avoid cloudy images, a cloud-masking function to mask out clouds and unwanted pixels, and then by averaging the NDWI values for each pixel of each image that fell within the monthly range.[25] Using monthly composites instead of single images has some advantages as it helps, for example, to reduce the impact of clouds or reliefs shadows.[26] The GEE scripts generates images for all 12 months of each year. The images then go through a manual visual inspection to assess if the algorithm outputs valid images or not (Fig. 3).[27] While generated composite images were selected for the months highlighted by the DAHITI, if a composite was not available for the selected months, or its quality was not good enough, the closest month with the highest/lowest water level was selected, and so on. As stated above, the generated images will contain pixels ranging from -1 to 1.

2.2.3 *Post-processing*

As mentioned before, the interpretation of the NDWI relies on an arbitrary threshold, with values above this threshold counted as water surfaces and below as non-water surfaces. This threshold is usually set at 0, but during the reclassification process (see below), it is essential and generally advisable to experiment with different values in order to ensure a better-quality output and avoid errors.[28] Most er-

[23] Gorelick *et alii* 2017.

[24] The scripts are available at: https://github.com/ReLand-Project/ReLandTishreen and https://github.com/ReLandProject/MosulDrownedLandscapes. The same scripts were merged and translated in R (see discussion) for enhanced reproducibility.

[25] Monthly composites are pixel operations on multiple satellite images acquired during a selected time frame. All the images within the timeframe are combined using specific mathematical operations (in our case the median) called *reducers* in GEE, to generate a single (usually clearer) image. Of course, usually the more images available, the better the results.

[26] Rayne *et alii* 2020; Sagar *et alii* 2017.

[27] It might happen, for example, that not enough images are present for the selected time frame, and that missing pixels removed from the cloud-masking function are not filled with other data from the same period because of the lack of images. In this case, a different month was selected, to avoid potentially missing information that could alter the final result.

[28] Acharya *et alii* 2016; Fisher, Danaher 2013; Ji, Zhang, Wylie 2009.

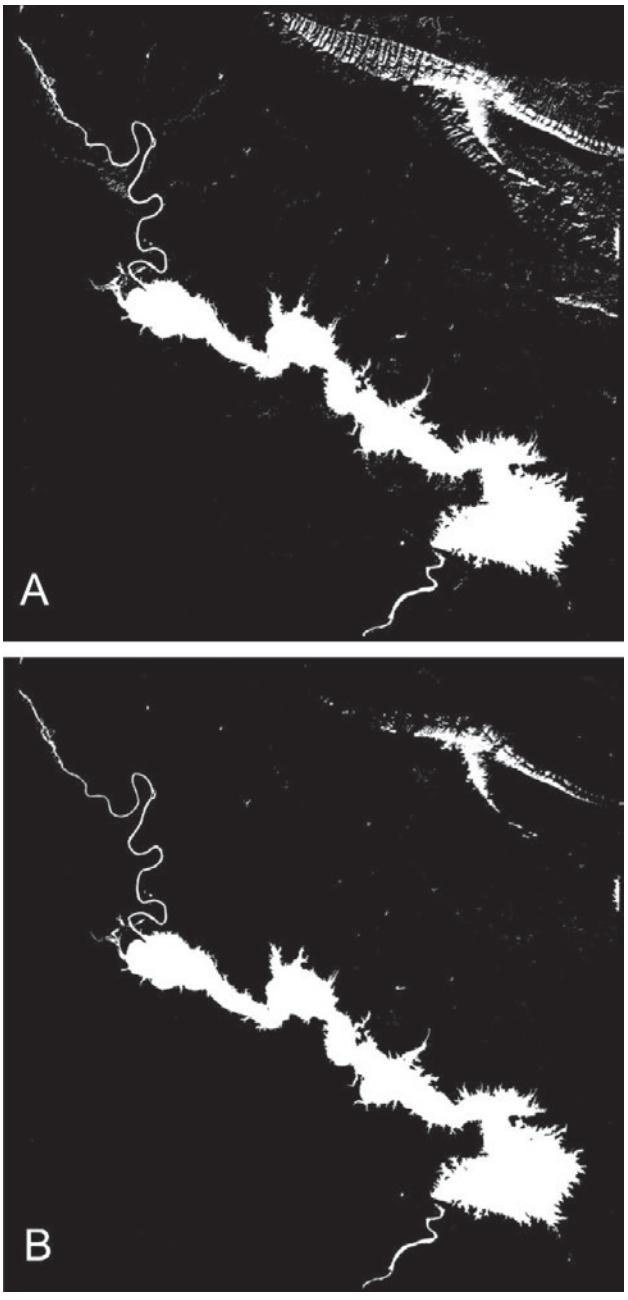

Fig. 4 - Example of difference results on applying a 0 threshold (A) and a 0.05 threshold (B) to the reclassification process.

rors, especially with the formula employed here, are caused by shadows and snow, which typically have a similar spectral signature to water in the SWIR band. An example is the images generated in one of the two major case studies to which this workflow was applied, i.e., the Mosul Dam. Here, tweaking the threshold to 0.05 instead of plain 0 helped minimize errors, bringing sensible improvement visible also at the naked eye (Fig. 4). For the Tishreen Dam, on the other hand, a threshold of 0 seemed accurate enough. As briefly noted by these two examples, while a threshold set at 0 can still lead to useful results, it is likely that different areas will require different tweaks and an accurate inspection before proceeding with the reclassification.

The threshold is essential for the next step of the post-processing. The generated NDWI images contain different pixel values ranging from -1 to 1; however, to calculate the temporal change within each site area, it is necessary to have unique values to quantify (i.e., to calculate the percentage of each single value). This involves reclassifying images to a set of values, in our case two: non-water and water surfaces. The workflow leverages the open-source programming language R to do that.[29] An R script was written to reclassify the generated composites based on the selected threshold quickly.[30] As mentioned before, to interpret the NDWI, values above the threshold are considered water and values below are considered non-water surfaces. Thus, citing the example of the Mosul Dam above, in the reclassification script, all pixel values below 0.05 are reclassified as 0 or non-water surface, and all those above or equal to 0.05 are reclassified as 1 or water surface.

However, a visual inspection of the resulting reclassification output is not enough, and there is a need for a qualitative measurement of the reclassified results. This is crucial because it will inform on how good the distinction between water and non-water features is and how effectively the workflow is identifying emerged site areas. The process for obtaining this qualitative measure is called accuracy assessment and is evaluated using a confusion or error matrix.[31] The accuracy assessment also aids with another common issue of reclassification of medium-resolution images, i.e., mixed pixels. Inaccuracies may occur when more than one class (in this case, land/water) is present within a single pixel, thus possibly resulting in an incorrect pixel attribution or a failure of change detection. This is usually addressed by manually inspecting several sample pixels against a higher resolution image from the same time and area of the

[29] https://www.r-project.org/ (last accessed 2024-03-07).

[30] This is done leveraging the {raster} package (https://cran.r-project.org/web/packages/raster/index.html, last accessed 2024-03-07).

[31] The error matrix is a way of comparing the expected results (the one from the reclassification) with actual pixel information from either ground truth or higher-resolution reference images (Congalton 1991; Congalton, Green 2019). In the error matrix, there are three main pieces of information present: Overall Accuracy (which measures how accurate is the overall reclassification of the entire image), Producer's Accuracy (measures the "omission errors", i.e., false negatives, answering the question "how likely is it that a pixel known to be water has been mapped as such in the reclassified image), and User's Accuracy (measures the"commission errors", i.e., false positives, answering the question of how likely it is that a pixel counted as water in the reclassified image is also water in the reference image).

Satellite Images	Overall Accuracy	Producer's Accuracy	User's Accuracy
Landsat 5	98.01%	98.63%	94.74%
Landsat 7	98.59%	100%	95.22%
Landsat 8	98.36%	99.24%	95.08%
Sentinel-2	98.36%	99.24%	94.99%

Tab. 1 - Confusion Matrix for the Mosul Dam Dataset.

Satellite Images	Overall Accuracy	Producer's Accuracy	User's Accuracy
Landsat 7	98.74%	98.57%	97.18%
Landsat 8	98.31%	98.48%	95.59%
Sentinel-2	99.16%	98.57%	98.57%

Tab. 2 - Confusion Matrix for the Tishreen Dam dataset.

reclassified one, registering classification errors, or using ground validation of reclassified areas. Generally, it is quite problematic to carry out an accuracy assessment for time-series change detection images, the main reason being that usually there are no higher resolution images for most of the analyzed periods and no field data.[32] This, in turn, means that the accuracy assessment needs to use the same data that were employed for generating the NDWI, which is not always optimal. However, when a limited number of classes is used (in this case, as explained, just two), it has been proven that these images can still yield optimal results.[33] During this workflow, validation has been carried out for two images per sensor, one with a lower water level, and one with a higher water level. All the images used for validation were true-color composites, pansharpened for Landsat 8, or at a higher resolution for Sentinel-2.[34] Sampling is carried out at the per-pixel level using a stratified random sampling[35] with two classes: water and non-water pixels. This method allows for allocating enough samples for each class depending on their area proportion, thus accounting for the lake area variation between images. Validation is conducted in QGIS using the open-source Semi-Automatic Classification Plugin.[36] When mixed pixels were detected, the reference class was assigned to either water or non-water, depending on the most prevalent class inside the sample. Overall, the NDWI showed promising results for both the case studies of Tishreen and the Mosul Dam and for all the satellites, confirming that the reclassification and the threshold choice were appropriate for the two regions.[37] In particular, with a minimum overall accuracy of 98% (see Tab. 1 and Tab. 2 for details on both dams) the reclassification can be considered good quality enough. With these results, we can be quite confident that the reclassification works as expected, but again, each new application should engage in some sort of accuracy assessment to find the optimum threshold[38].

After the image reclassification and the accuracy assessment of the results, the next step is to calculate the percentage of emerged area for each site under investigation. The workflow leverages the Zonal Histogram Algorithm, which counts each unique value from a raster layer contained within polygon features, which is why there is a need for a set of unique values (the reclassification output) (Fig. 5). The zonal histogram is a QGIS algorithm and was used in all the case studies mentioned above.[39] The output of a zonal histogram is a series of shapefiles (the same number as the satellite images used) with the count of each unique pixel in their attribute table. After applying the Zonal Histogram, there is a need to convert the counts for each monthly image in percentages, merge the results, and obtain a single shapefile with the amount of resurfaced extent for each site. The whole procedure is carried out again using the R programming lan-

[32] Congalton, Green 2019; Olofsson *et alii* 2014.

[33] Özelkan 2020; Sagar *et alii* 2017.

[34] The resolution of the reclassified Sentinel-2 NDWI composites is 20m (because of the use of the SWIR band), while the true color composites are 10m.

[35] Congalton, Green 2019.

[36] Congedo 2021.

[37] Generally, when there are only two classes in the reclassified image and the separation between the lake and its surroundings is clear, good results are expected (Fisher, Danaher 2013).

[38] The lower User's accuracy compared to the Overall and Producer's accuracy is usually tied to the medium-resolution nature of the satellite images and the difficulty of dealing with mixed pixels. If this value is lower, it might also indicate a need to tweak the threshold further.

[39] The algorithm documentation is available at: https://docs.qgis.org/latest/en/docs/user_manual/processing_algs/qgis/rasteranalysis.html#zonal-histogram (last accessed 2024-03-07).

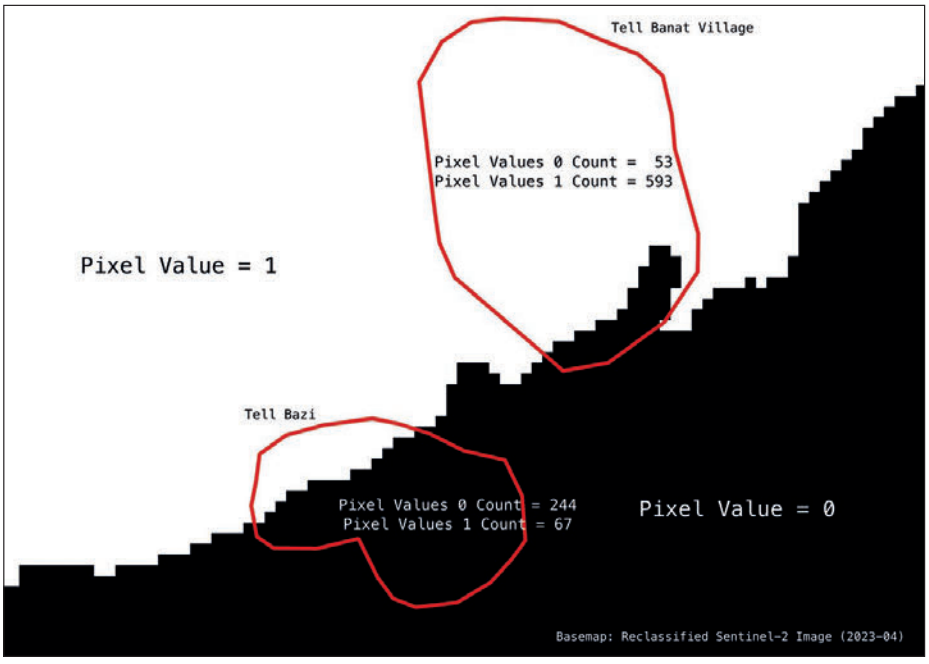

Fig. 5 - Conceptual model of the Zonal Histogram Algorithm.

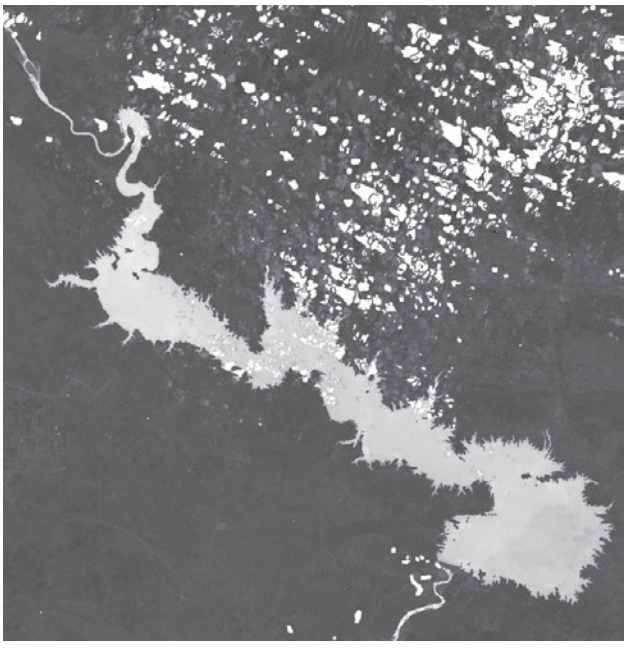

Fig. 6 - Example of missing information due to clouds and lack of enough images for the median reducer to fill missing pixels.

guage, and specifically the {qgisprocess} package,[40] which allows to run QGIS algorithms inside R. A script (also available in the project repository, see below) was written that apply the algorithm, converts the results to percentages and merge the results, calculating also other fields useful for a classification of the results, which will be discussed below.

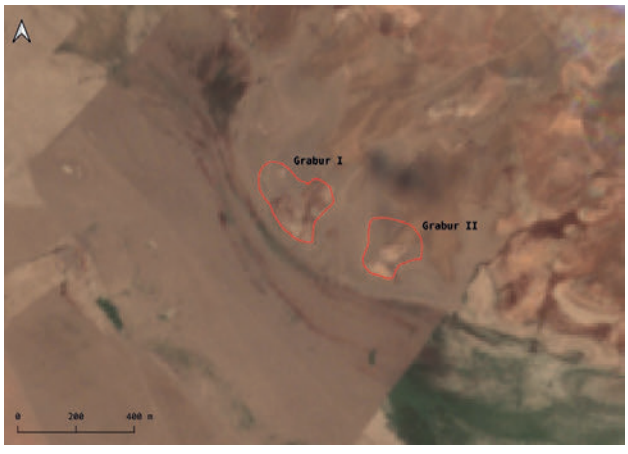

Fig. 7 - Example of area that might be classified as non-water surface but not necessarily accessible on foot.

2.3 *Technical Limitations*

Change detection is a complex challenge that depends on many factors according to the chosen method. There are, thus, some technical limitations that even threshold tweaking and accuracy assessment cannot overcome. Firstly, a solid archaeological dataset is highly advised. In the Mosul Dam case study, this was already prepared thanks to a recent work of literature review, remote and ground surveys,[41] and was further improved for the case study in question.[42]

[40] Dunnington *et alii* 2023.
[41] Sconzo, Simi 2020.
[42] Sconzo, Simi, Titolo 2023.

Name	2020-08	2021-05	AlwaysEm	AlwaysSub	Affected	NeverEm	NeverSub
Duluk	100	100	1	0	0	0	0
Hammam Saghir N	41	100	0	0	1	0	0
Tell Ahmar	53	91	0	0	1	0	0
Tell El-Kebir	0	0	0	1	0	1	0

Tab. 3 - Partial example of attribute table of the output shapefile with quantitative columns.

The same care for a dataset creation was applied for the Tishreen dam case study.[43] The main reason as to why an accurate dataset is needed is twofold. First, the major complication to the dataset comes in the case of overlapping and contradictory information in published data. This issue can affect the effective reconstruction of the site areas, and in return make the resurfaces calculation less effective. A second issue comes from the types of sites recorded during the past archaeological surveys and it is also linked to how this workflow operate. All the processes highlighted above assume that a site extends horizontally on a (relatively) limited and identifiable area. This poses a challenge when dealing with types of sites for which the extension cannot be practically reconstructed. A primary example of this is the burial caves in the Tishreen Dam, which, while recorded and analyzed, are likely not correctly represented in the results.[44] This is not a limitation of the workflow per se, as the latter operates correctly, but something to keep in mind when applying it.

Due to being based on multispectral images, the workflow is also sensible to clouds. This means that in regions where clouds are heavily present, they can impact the availability of satellite images and, thus, the applicability of the workflow to the entire selected period. Another element that was identified during the work on the Tishreen Dam is that the workflow can be exposed to errors induced by null (or NA) values. NA values are missing data present in the satellite images due to registration errors, other factors, or sometimes the removal of clouds from the cloud masking process (Fig. 6). While NA values are not a substantial issue if recognized early, they can impact the quantitative calculation if that is not the case.[45] Currently, a manual inspection is in place to remove any NA value before this calculation are made, but in the future, this will be automated.[46]

Lastly, the tool discriminates between water and non-water surfaces, but it does not return any information about the surroundings of a site or its accessibility. The NDWI correctly identifies a terrain as non-water surface. However, that terrain might still be inaccessible due to river deposits, for example (Fig. 7). This is not an error in the reclassification or index extraction. However, sometimes, this factor needs to be accounted for if field surveys are planned based on the results. For this reason, inspecting recent true-color satellite images before or during field surveys is always advisable.

3. Results

This section will briefly explain what potential results the workflow can produce. However, an entire case study is not included below, as each application with the respective results has been presented separately.[47] Therefore, only a general presentation of all the possible outcomes from applying the workflow will follow. Primarily, the workflow will produce quantitative results, which can be connected to qualitative information to enhance the understanding or expectations regarding, for example, sites survival rates, as discussed below.

First and foremost, the workflow outputs two shapefiles (one for the maximum water levels and one for the minimum water levels) with an identical structure. An attribute table contains the emerged percentage of each site for each observation, i.e. as many columns as satellite images utilized. Appended to this table are also five more qualitative columns with records in a binary format (0 or 1) (Tab. 3): two of these are relative to the whole study period and have identical values in both shapefiles (NeverEm, NeverSub). These two columns record whether a site was never exposed or never touched by the waters. The other three columns are relative to the water level period (i.e., minimum or maximum water level) and record whether a site has been affected by the water fluctuations (Affected), if it has always been outside or under the water in the respective water level period (AlwaysEm, AlwaysSub).

[43] Sconzo Simi, Titolo, this volume.

[44] See the discussion in Sconzo, Simi, Titolo, this volume.

[45] This is because the NA values, if not removed or accounted for, can distort the percentage calculation.

[46] See also Sconzo, Simi, Titolo, this volume, note 56 on another limitation specific for the Tishreen Dam, that is not however directly related to the methods presented here, but to the lack of DAHITI data for the years of analysis.

[47] See Sconzo, Simi, Titolo 2023 for the Mosul Dam case study and Sconzo Simi, this volume, for the Tishreen Dam Case study.

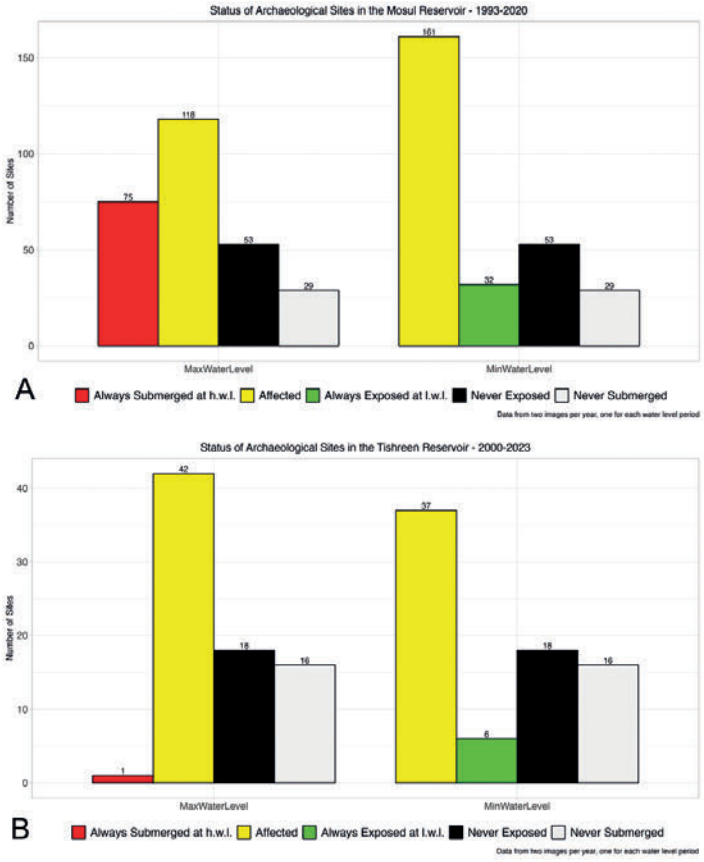

Fig. 8 - Example of quantitative results: emergence pattern categories for the Mosul Dam (A) and the Tishreen Dam (B).

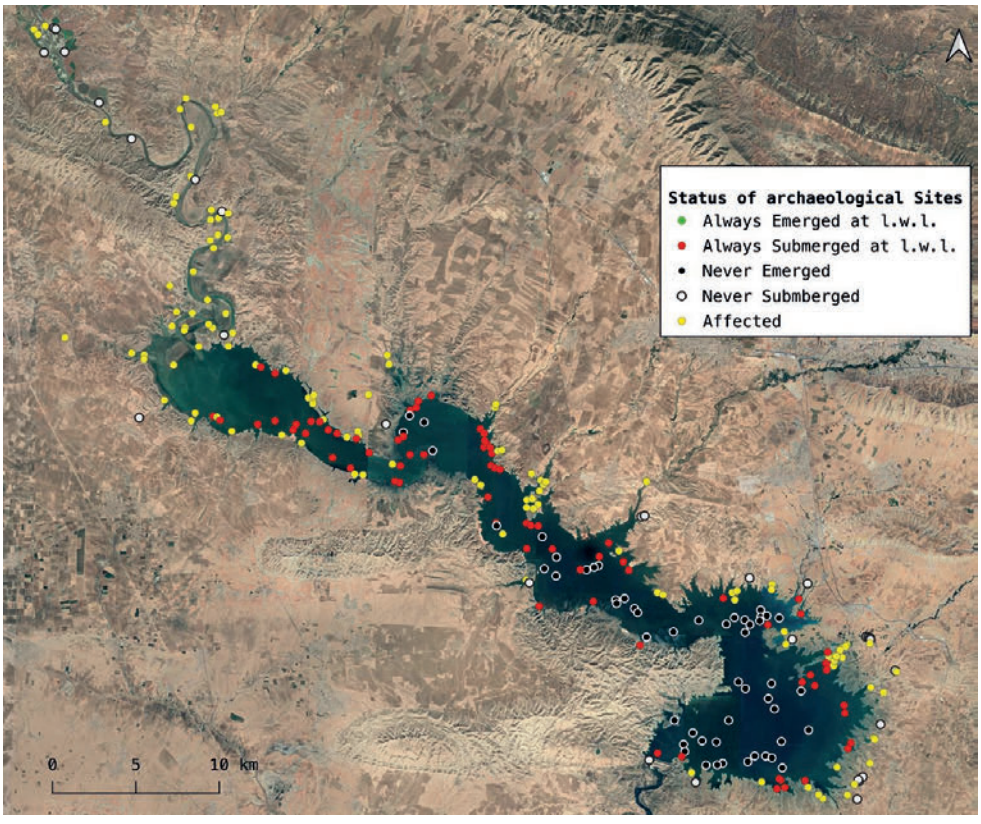

Fig. 9 - Example of quantitative results coupled with spatial data: spatial distribution map with emergence pattern categories for archaeological sites in the Mosul Dam.

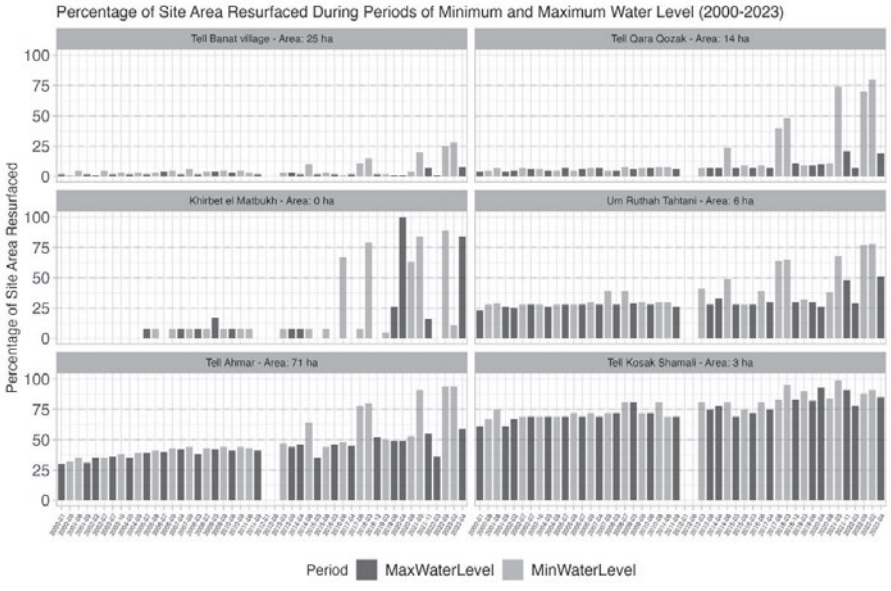

Fig. 10 - Example of single sites analysis results: emergence patterns for selected sites in the Tishreen Dam.

One of the immediate data that can be extracted from this shapefile are quantitative information obtained by counting the number of sites that in the period under examination have never been exposed, those that have never been touched by the ebbing waters, and those that are instead still affected by the water fluctuation processes (Fig. 8). Moreover, by dividing the inspection years in minimum and maximum water level periods, the workflow is able to assess how many sites are outside the water when the lake is at its lower levels, and those which are generally easily submerged when the water is higher.

The creation of a shapefile with this information also means that the spatial distribution of the site categories mentioned before can be observed to further understand behaviors and patterns (Fig. 9). At this point, one could also factor in qualitative data, such as the site type, to inspect whether a site's morphology can influence its emersion patterns. Moreover, understanding site morphology can also help to understand possible damages and survival rates further down the line.

Thanks to the time-series images, the workflow can also quantify emerged archaeological features per year, according to water level periods. This temporal quantification, especially when coupled with water level data (e.g., from DAHITI), can help understand emersion patterns across different years. It is possible to highlight years in which significant reductions in water level occurred and associate these numbers with the number of resurfaced sites. This inspection is useful not only to assess past behaviors but also to plan safeguarding activities. In fact, similar reductions in water level are expected to bring a similar number of sites outside the water. This element is particularly evident when the number of exposed sites and years with similar water levels is associated with the spatial location of the same archaeological features. This association shows how it is possible to leverage the history of the lake to understand emersion patterns and prepare salvage operations in the future. Another outcome of the time-series analysis is also the identification of any sudden changes in the lake's surface and their impact on the cultural heritage. While typically the changes in the lake level are gradual, sudden changes might happen due to water management choices or other external events not always linked to the natural behavior of the reservoir.[48] Knowing how and which sites are more prone to resurface also in this occasion in another valuable information for salvage activities.

While general data are definitely useful, the workflow also allows for the inspection of single sites. General trends might hide local behaviors that are only visible when inspecting each site on its own. We can, in fact, group sites based on their resurface history (Fig. 10). Knowing that, for example, a group of sites will resurface only when there is a substantial reduction in the water level is another valuable addition to the planning of salvage activities. This knowledge can, in fact, help in the prioritization process and lead to more informed undertakings.

Lastly, on a smaller scale, the workflow is also useful for analyzing single sites without the need to examine a larger region or a group of different sites. Inspecting a single site can help understand

[48] For a review, see Sconzo, Simi, this volume.

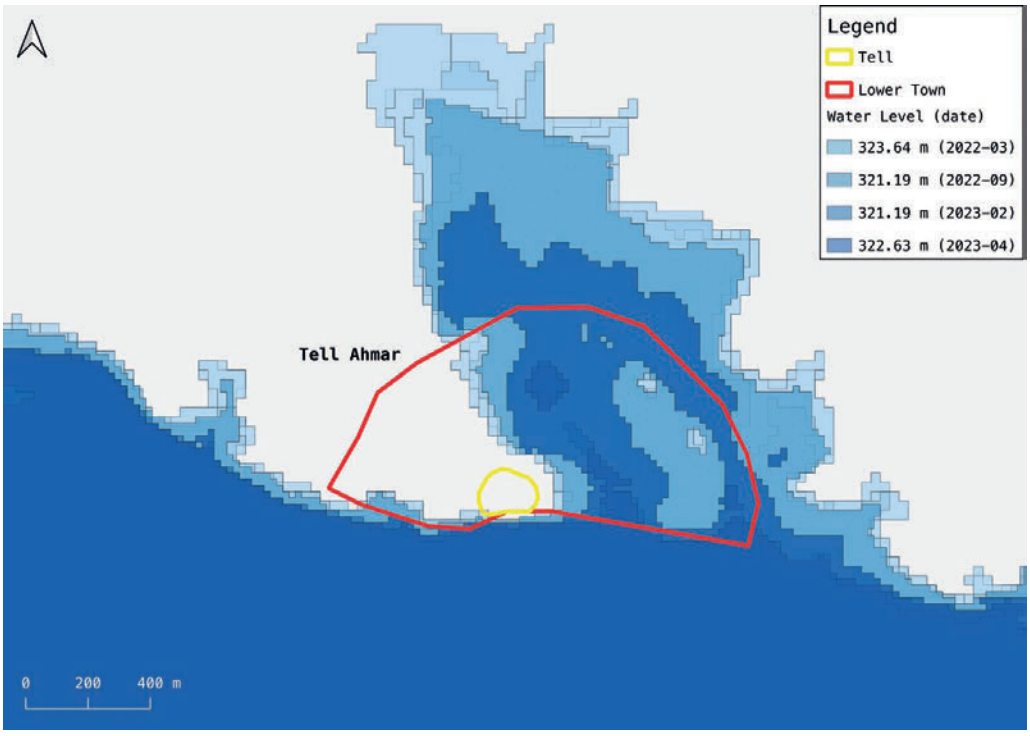

Fig. 11 - Example of single sites results and spatial visualisation: different stages of submersion in the lower town of Tell Ahmar (Tishreen Dam). Water extent extracted from reclassifed NDWI composites.

if there have been different phases of emersion and submersion and if the site was affected by known water fluctuations or similar events. This can also be coupled with a remote sensing analysis to understand the spatial location of the impacted area (Fig. 11). All this information can be gathered to understand, e.g., possible damages to a specific site or to plan field visits.

4. Discussion

It is possible to observe how the tool is flexible and adaptable to non-identical situations and needs. As a matter of fact, the tool provides many different outputs that can be adapted to the needs of each application. At the general level, the tool can distinguish between sites that will never resurface even at the lowest lake reduction, sites that were never touched by the lake, and those still affected by the reservoir's fluctuating dynamics. These data already provide important information as it can lead, thanks to spatial mapping, to more informed ground reconnaissance and rescue projects by prioritizing sites endangered by water fluctuations. Adding to this, having a time-series view of the number of sites resurfaced at each water level period allows to further plan the aforementioned undertakings, as generally speaking, years with similar water levels brings a similar number of sites out of the water. Thus, future periods of water reduction might lead to similar results. The two case studies in which the tool was applied also showed how the results can highlight the complex interplay of natural water fluctuations, water management, and political and strategic factors impacting the resurfacing of archaeological sites.[49]

However, the tool also works at different scales, which can be helpful when a more fine-grained approach is needed. By highlighting a site's (or a group of sites') resurfacing history, the tool can inform a more detailed evaluation of a small-scale rescue effort or a better understanding of potential erosion damages.

The tool is also a first of its kind as a companion to a survey project. In fact, it has been successfully applied at different stages before and during the fieldwork activities of the ReLand Project in Iraqi Kurdistan. Most importantly, this application allowed for a ground validation of the results, which can now be considered accurate beyond the error matrix results. Moreover, thanks to the relative ease of use, the tool helped choose a prioritization strategy that fit the time constraints of an archaeological survey project. In fact, depending on the analyses run every

[49] Sconzo, Simi, Titolo, 2023; this volume.

five days (the time of each Sentinel-2 acquisition), the project could direct its effort to either endangered sites or others that have not been known to resurface before.

There are, of course, technical limitations that must be kept in mind. As mentioned before, a solid archaeological dataset is best suited as an input to this kind of application, as it will necessarily lead to better results. The same goes for the remote sensing data (unprocessed and processed), which must be checked for any missing data or inaccuracies. While the two case studies relied on two observations per year, this is not mandatory, and the workflow can easily be adapted to more complex monitoring processes, for example, using one image per month or week. In fact, a weekly/five-day monitoring was put in place when on the ground during the ReLand fieldwork.

4.1 *Reproducibility and Open-Source*

This tool is not necessarily the easiest to adopt among the scientific community due to the technical knowledge that is still required (see below). However, means have been put in place to enhance reproducibility and understanding behind the technical process, as open sourcing and ease of use have been an integral part of the project since the beginning. First, the workflow is designed to use freely and globally available datasets, ensuring that costs and availability will not be a concern when applying the tool elsewhere. Second, almost all the analyses have been carried out using the R programming language, a powerful scripting language that has become quite popular among archaeologists.[50] The code for Google Earth Engine is instead written in Javascript, but also reproducible in R for scholars more familiar with it (see below). The GEE code is also easily shareable, as it usually requires just one link to share and run the analysis.[51] The use of plain text scripts has been proven to be a suitable method of ensuring reproducibility and transparency in the analysis, especially regarding the Open Science applications.[52] However, just providing the code is not always enough, especially if the application requires a good knowledge of a scripting language. While not all barriers can be torn down, commenting the code, enabling the selection of analysis variables for users, and providing guides and directions is a valuable practice to guide other scholars in re-using the scripts.[53] The code itself is written in a way that users only need to tweak some few initial parameters (variables), while all the processing takes place after these tweaks. This approach is applied where most of the decision-making is necessary, i.e., during the image choice and processing phase. For example, user can select beforehand the area to be analyzed, the satellite images to use, the temporal range, whether to produce NDWI or True Color composites (useful for accuracy assessment), the Google Drive folder name to save the images, their names, and their reference system.[54] This stage is carried out using the GEE scripts, but the same scripts were also converted (and tested) in R, thanks to the {rgee} package,[55] which act as an interface between R and GEE. This ensures further reproducibility and allows for the entire analysis to be carried out in R, for those unfamiliar with JavaScript or GEE.[56] When decision-making is not strictly required (i.e., in the image post-processing and shapefiles creation phase), user-defined variables are not set, but the code is extensively documented to better guide towards its use, with an explanation of functions and operations carried out.[57] The use of GEE, R and QGIS also means that one can choose the more familiar tool to do most analyses and that the entire workflow can potentially be platform-agnostic.[58]

The code is made available on Github, using a GitHub organization profile for the ReLand project where future code, data, and analysis will be hosted.[59] Of course, future datasets, when associated with a published paper, can also be hosted on platforms like Zenodo, for improved sharing between the academic community and long-term archiving through unique IDs.

---

[50] Marwick 2018.

[51] Rayne *et alii* 2020.

[52] Marwick 2017.

[53] See Orengo, Petrie 2017; Rayne *et alii* 2020 for good examples of well-documented code available with the publication.

[54] These parameters are common to the GEE and R scripts, while unique to R are options to display images in an interactive map (only for visualization purposes), and to limit the images to a certain amount (useful for testing purposes).

[55] Aybar *et alii* 2020.

[56] The only downside of the combination of R/rgee is the lack of interactivity, and the slower support for some GEE features, such as choosing a subset of a series of images, which would make the process smoother. For now the user, once all the variables are set, needs to save on Google Drive all the images generated and then select the ones needed. On the other hand, in Google Earth Engine one can inspect the images before running the export function, and only save those that are needed or met chosen criteria.

[57] The R code comprises four scripts: one for the google earth engine application, one for the image reclassification, one for the application of the zonal histogram algorithm and one for the generation of the final shapefiles with quantitative informations. The name of the script files also follows a progressive number from 1 to 4 to indicate the order in which they should be run. For now the scripts assumes a specific directory structure similar to the one found in the online repository, but in the future more flexibility is planned.

[58] The tool was tested on macOS and Linux (Ubuntu-based) with no differences except for the required installation of some missing GDAL and other packages on the Linux machine.

[59] https://github.com/ReLandProject (last accessed 2024-03-07).

5. Conclusions

The tool presented here showed how spectral indices applied to multitemporal satellite images can be used to analyze the phenomenon of archaeological sites resurfacing from artificial reservoirs. The tool offers clear advantages regarding reliability, scale, and cost-efficency, and it is a first-of-its-kind regarding the semi-automation of such an endeavour. The phenomenon of resurfacing sites has been mostly unknown, but much evidence has shown that there is a need to act to avoid further destruction and carry out proper documentation and safeguarding of the emerging cultural heritage.[60] The workflow has matured since its first iteration, and have been successfully tested in two different areas, while also employed during a field survey, confirming its usefulness for different types of analyses and requirements.

There is, however, still room for improvement. The tool still requires a considerable amount of manual work and knowledge of the area and remote sensing to be applied to other regions. Accuracy assessment must also be done individually and manually, which is usually not an easy task. Moreover, the installation and use of programming software like R should not be taken for granted, and the same goes for knowledge of JavaScript. To address these issues, ideal solutions and a possible next step for the tool would be to lean more towards web apps and less towards desk-based software. Google Earth Engine and the Shiny library[61] for R both offer possible and viable alternatives that would further enhance the accessibility of this workflow. Another option to be explored could be the creation of an R package to further standardize and enhance the ease of use through R.

[60] Eidem 2015; Titolo 2021; Sconzo, Qasim 2021; Sconzo, Simi, Titolo 2023.

[61] https://shiny.posit.co/ (last accessed 2024-03-07).

BIBLIOGRAPHY

Acharya, T.D. *et alii*
2016 - T.D. Acharya, D.H. Lee, I.T. Yang, J.K. Lee, "Identification of Water Bodies in a Landsat 8 OLI Image Using a J48 Decision Tree", *Sensors* 16, 7, 1075-1075.

Agapiou, A. - Hadjimitsis, D.G. - Alexakis, D.D.
2012 - "Evaluation of Broadband and Narrowband Vegetation Indices for the Identification of Archaeological Crop Marks", *Remote Sensing* 4, 12, 3892-3919.

Agapiou, A. *et alii*
2011 - A. Agapiou, D.G. Hadjimitsis, C. Papoutsa, D.D. Alexakis, G. Papadavid, "The Importance of Accounting for Atmospheric Effects in the Application of NDVI and Interpretation of Satellite Imagery Supporting Archaeological Research: The Case Studies of Palaepaphos and Nea Paphos Sites in Cyprus", *Remote Sensing* 3, 12, 2605-2629.

Al-Ansari, N.
2016 - "Hydro-Politics of the Tigris and Euphrates Basins", *Engineering* 8, 3, 140-172.

Aybar, C. *et alii*
2020 - C. Aybar, Q. Wu, L. Bautista, R. Yali, A. Barja, "Rgee: An R Package for Interacting with Google Earth Engine", *Journal of Open Source Software* 5, 51, 2272-2272.

Calleja, J.F. *et alii*
2018 - J.F. Calleja, O. Requejo Pagés, N. Díaz-Álvarez, J. Peón, N. Gutiérrez, E. Martín-Hernández, A. Cebada Relea, D. Rubio Melendi, P. Fernández Álvarez, "Detection of buried archaeological remains with the combined use of satellite multispectral data and UAV data", *International Journal of Applied Earth Observation and Geoinformation* 73 *<https://www.sciencedirect.com/science/article/pii/S0303243418304410>*, 555-573.

Cheng, G. *et alii*
2023 - G. Cheng, Y. Huang, X. Li, S. Lyu, Z. Xu, Q. Zhao, S. Xiang, "Change Detection Methods for Remote Sensing in the Last Decade: A Comprehensive Review", *arXiv:2305.05813.*

Congalton, R.G.
1991 "A review of assessing the accuracy of classifications of remotely sensed data"*, Remote Sensing of Environment* 37, 1, 35-46.

Congalton, R.G. - Green, K.
2019 - *Assessing the Accuracy of Remotely Sensed Data. Principles and Practices*, Boca Raton.

Congedo, L.
2021 - "Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS", *Journal of Open Source Software* 6, 64, 3172.

Du, Z. *et alii*
2014 - Z. Du, W. Li, D. Zhou, L. Tian, F. Ling, H. Wang, Y. Gui, B. Sun, "Analysis of Landsat-8 OLI Imagery for Land Surface Water Mapping", *Remote Sensing Letters* 5, 7, 672-681.

Dunnington, D. *et alii*
2023 - D. Dunnington, F. Vanderhaeghe, J. Caha, J. Muenchow, *R package qgisprocess: Use QGIS processing algorithms. Version 0.2.0*, 2023.

Eidem, J.
1999 - "The "Tishrin Project" and Salvage Archaeology", G. del Olmo Lete, J.-L.M. Fenollós (eds.), *Archaeology of the Upper Syrian Euphrates: The Tishrin Dam Area*, Barcelona, 19-24.
2015 - "Dams and Damage. Heritage Loss and Second Phase Salvage on the Rania Plain (Kurdish Region of Iraq)", *Annual Report 2015. The Netherlands Institute for the Near East and the Netherlands Institute in Turkey*, Leiden-Istanbul, 2-13.
2020 - "Before and After the Flood. Archaeology on the Rania Plain Then and Now", in J. Eidem (ed.), *Zagros Studies. Proceedings of the NINO Jubilee Conference and Other Research*, Leiden-Leuven, 99-130.

Eidem, J. *et alii*
2019 - J. Eidem, M. Merlino, E. Mariotti, R.K. Salih, "Tell Shemshara 2018: Emerging and Floating Evidence", *Ash-sharq: Bulletin of the Ancient Near East - Archaeological, Historical and Societal Studies* 3, 1 *<https://archaeopresspublishing.com/ojs/index.php/ash-sharq/article/view/712>*, 21-33.

Fisher, A. - Danaher, T.
2013 - "A Water Index for SPOT5 HRG Satellite Imagery, New South Wales, Australia, Determined by Linear Discriminant Analysis", *Remote Sensing* 5, 11, 5907-5925.

Fuldain González, J.J. - Varón Hernández, F.R.
2019 - "NDVI Identification and Survey of a Roman

Road in the Northern Spanish Province of Álava", *Remote Sensing* 11, 6 <*https://www.mdpi.com/2072-4292/11/6/725*>, 725.

Gleick, P.H.
2014 - "Water, Drought, Climate Change, and Conflict in Syria", *Weather, Climate, and Society* 6, 3, 331-340.

Gorelick, N. *et alii*
2017 - N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, "Google Earth Engine: Planetary-scale Geospatial Analysis for Everyone", *Remote Sensing of Environment* 202, 18-27.

Hadjimitsis, D.G. *et alii*
2010 - D.G. Hadjimitsis, G. Papadavid, A. Agapiou, K. Themistocleous, M.G. Hadjimitsis, A. Retalis, S. Michaelides, N. Chrysoulakis, L. Toulios, C.R.I. Clayton, "Atmospheric Correction for Satellite Remotely Sensed Data Intended for Agricultural Applications: Impact on Vegetation Indices", *Natural Hazards and Earth System Science* 10, 1, 89-95.

Ji, L. - Zhang, L. - Wylie, B.
2009 - "Analysis of Dynamic Thresholds for the Normalized Difference Water Index", *Photogrammetric Engineering and Remote Sensing* 75, 11, 1307-1317.

Kalayci, T. *et alii*
2019 - T. Kalayci, R. Lasaponara, J. Wainwright, N. Masini, "Multispectral Contrast of Archaeological Features: A Quantitative Evaluation", *Remote Sensing* 11, 8 <*https://www.mdpi.com/2072-4292/11/8/913*>, 913.

Li, W. *et alii*
2013 - W. Li, Z. Du, F. Ling, D. Zhou, H. Wang, Y. Gui, B. Sun, X. Zhang, "A Comparison of Land Surface Water Mapping Using the Normalized Difference Water Index from TM, ETM+ and ALI", *Remote Sensing* 5, 11, 5530-5549.

Lu, D. *et alii*
2004 - D. Lu, P. Mausel, E. Brondízio, E. Moran, "*Change detection techniques*", *International Journal of Remote Sensing* 25, 12, 2365-2401.

Luke, C. - Meskell, L.
2019 - "Archaeology, Assistance, and Aggression along the Euphrates: Reflections from Raqqa", *International Journal of Cultural Policy* 25, 7, 831-842.

Lupu, Y.
2001 - "International Law and the Waters of the Euphrates and Tigris", *Georgetown International Environmental Law Review* 14, 349-366.

Marchetti, N. *et alii*
2018 - N. Marchetti, A. Al-Hussainy, M. Valeri, F. Zaina, "Assessing Endangered Cultural Heritage in Central Iraq. Methods and Perspectives of the QADIS Survey Project", *Sumer* 64, 11-34.
2019 - N. Marchetti, A. Curci, M.C. Gatto, S. Nicolini, S. Mühl, F. Zaina, "A Multi-Scalar Approach for Assessing the Impact of Dams on the Cultural Heritage in the Middle East and North Africa", *Journal of Cultural Heritage* 37, 17-28.

Marwick, B.
2017 - "Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation", *Journal of Archaeological Method and Theory* 24, 2, 424-450.
2018 - "R Coding and Modeling", *The Encyclopedia of Archaeological Sciences*, 1-5.

McFeeters, S.K.
1996 - "The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features", *International Journal of Remote Sensing* 17, 7, 1425-1432.

Olofsson, P. *et alii*
2014 - P. Olofsson, G.M. Foody, M. Herold, S.V. Stehman, C.E. Woodcock, M.A. Wulder, "Good Practices for Estimating Area and Assessing Accuracy of Land Change", *Remote Sensing of Environment* 148, 42-57.

Orengo, H.A. - Petrie, C.A.
2017 - "Large-Scale, Multi-Temporal Remote Sensing of Palaeo-River Networks: A Case Study from Northwest India and Its Implications for the Indus Civilisation", *Remote Sensing* 9, 7.

Özelkan, E.
2020 - "Water Body Detection Analysis Using NDWI Indices Derived from Landsat-8 OLI", *Polish Journal of Environmental Studies* 29, 2, 1759-1769.

Pang, Y. *et alii*
2024 - Y. Pang, J. Yu, L. Xi, D. Ge, P. Zhou, C. Hou, P. He, L. Zhao, "Remote Sensing Extraction of Lakes on the Tibetan Plateau Based on the Google Earth Engine and Deep Learning", *Remote Sensing* 16, 3, 583.

Qiao, C. *et alii*
2012 - C. Qiao, J. Luo, Y. Sheng, Z. Shen, Z. Zhu, D. Ming, "An Adaptive Water Extraction Method from Remote Sensing Image Based on NDWI", *Journal of the Indian Society of Remote Sensing* 40, 3, 421-433.

Rayne, L. *et alii*
2020 - L. Rayne, M.C. Gatto, L. Abdulaati, M. Al-Haddad, M. Sterry, N. Sheldrick, D. Mattingly, "Detecting Change at Archaeological Sites in North Africa Using Open-Source Satellite Imagery", *Remote Sensing* 12, 22, 3694-3694.

Sagar, S. *et alii*
2017 - S. Sagar, D. Roberts, B. Bala, L. Lymburner, "Extracting the Intertidal Extent and Topography of the Australian Coastline from a 28 Year Time Series of Landsat Observations", *Remote Sensing of Environment* 195, 153-169.

Schwatke, C. *et alii*
2015 - C. Schwatke, D. Dettmering, W. Bosch, F. Seitz, "DAHITI an Innovative Approach for Estimating Water Level Time Series over Inland Waters Using Multi-Mission Satellite Altimetry", *Hydrology and Earth System Sciences* 19, 10, 4345-4364.

Sconzo, P. - Qasim, H.A.
2021 - "Investigating Jubaniyah. A Late Chalcolithic site on the Upper Tigris River, Iraqi Kurdistan. Preliminary report", *Mesopotamia* 56, 1-60.

Sconzo, P. - Simi, F.
2020 - "Settlement Dynamics on the Banks of the Upper Tigris, Iraq: The Mosul Dam Reservoir Survey (1980)", *Journal of Open Archaeology Data* 8, 1.
2024 - "Waiting for the deluge? Some thoughts on the impact of dams on Heritage in Southwest Asia and North Africa", this volume.

Sconzo, P. - Simi, F. - Titolo, A.
2023 - "Drowned Landscapes: The Rediscovered Archaeological Heritage of the Mosul Dam Reservoir", *Bulletin of the American Society of Overseas Research* 389, 165-189.
2024 - "From the Tigris to the Euphrates Banks: Post-Flooding Assessment at the Tishreen Dam Reservoir, North Syria", this volume.

Titolo, A.
2021 - "Use of Time-Series NDWI to Monitor Emerging Archaeological Sites: Case Studies from Iraqi Artificial Reservoirs", *Remote Sensing* 13, 4, 786.

Wilkinson, T.J.
2004 - *On the Margins of the Euphrates: Settlement and Land Use at Tell Es-Sweyhat and in the Upper Lake Assad Area, Syria*, Chicago.

Xu, H.
2006 - "Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery", *International Journal of Remote Sensing* 27, 14, 3025-3033.