Received: 27 April 2022 | Revised: 22 August 2022 | Accepted: 30 August 2022

DOI: 10.1002/arp.1874

RESEARCH ARTICLE

WILEY

Archaeological site identification from open access multispectral imagery: Cloud computing applications in Northern Kurdistan (Iraq)

Riccardo Valente[1] | Eleonora Maset[2] | Marco Iamoni[1]

[1]DIUM, Dipartimento di Studi Umanistici e del Patrimonio Culturale, Università di Udine, Udine, Italy

[2]DPIA, Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine, Italy

Correspondence
Riccardo Valente, DIUM, Dipartimento di Studi Umanistici e del Patrimonio Culturale, Università di Udine, Udine, Italy.
Email: riccardo.valente@uniud.it

Funding information
Italian Ministry of Foreign Affairs and International Cooperation; Italian Ministry of University and Research; University of Udine

Abstract

This paper presents the results of an archaeological survey carried out in the Navkur Plain, Iraqi Kurdistan, as part of the 'Asingeran Archaeological Project'. The survey was prepared using remote sensing products accessed via Google Earth Engine©, a large-scale cloud computing service freely available to the scientific community that allows processing remote sensing big data. Outputs generated with a multitemporal approach are particularly successful for archaeological research, because it is possible to maximize the visibility of archaeological sites, improving their detection. Multispectral imagery from Landsat 5, Landsat 7 and Sentinel-2 collections were used and processed, testing their utility for finding unknown ancient settlements in the densely studied area of Northern Mesopotamia. Seventeen new sites were discovered in an already surveyed area of limited size (<100 km^2), showing the potentialities of this method. The advantages of cloud computing for Near Eastern Archaeology and the results of the survey are also presented and discussed.

KEYWORDS
archaeology, Google Earth Engine, Landsat 5, Landsat 7, multispectral data, Sentinel-2, site detection, spectral signatures

1 | INTRODUCTION

The archaeological heritage of the autonomous region of Iraqi Kurdistan, part of whose territory belongs to the historical area of Northern Mesopotamia, has been widely investigated by numerous international archaeological missions in recent decades that have conducted excavations and regional surveys, increasing our knowledge of a region where human occupation has been continuous for millennia (Kopanias et al., 2015). In the northern part of Kurdistan, four regional projects currently cover a continuous area of about 13,500 km^2; from west to east, these are the Eastern Habur Archaeological Survey (Pfälzner et al., 2016), the Land of Nineveh Archaeological Project (LoNAP), the Upper Greater Zab Archaeological Reconnaissance (UGZAR) (Koliński, 2018) and the Erbil Plain Archaeological Survey (EPAS) (Ur et al., 2021). Among these projects, LoNAP has been carried out by the University of Udine since 2012 and covers approximately 3000 km^2 eastward of the artificial lake created by the Mosul Dam; after almost 10 years of extensive surveying on the field, a total of 1098 archaeological features, from monumental tells to isolated ruins, have been identified (Bonacossi & Iamoni, 2015; Coppini, 2018; Gavagnin et al., 2016; Gavagnin, 2016; Iamoni, 2016; Morandi Bonacossi, 2016; Palermo, 2016; Simi, 2020). Within the borders of LoNAP, the Asingeran Excavation Project (AEP) was started in 2019 to explore the archaeological potentialities of *Tell Asingeran*, located at the south-eastern limit of the LoNAP's territory and the surrounding areas (Iamoni & Qasim, 2021).

Although all of the previous regional projects cover areas with diverse morphology and land use, remote sensing (RS) resources have been widely used for the remote exploration and the identification of new ones (Herrmann et al., 2018; Koliński, 2015; Ur et al., 2021). RS is an important tool for archaeological research and is now generally used for the discovery, assessment and management of global archaeological heritage (Lasaponara & Masini, 2012; Luo et al., 2019; Masini & Lasaponara, 2017; Parcak, 2009; Wiseman & El-Baz, 2007), with many published examples. Lasaponara and Masini ((2006), (2007)) and Gallo et al. (2009) used RS sources to identify archaeological features in southern Italy, while Grøn et al. (2011) applied multispectral imagery and ground-truthing methods to Norwegian archaeological sites. Agapiou et al. ((2010), (2016)) used vegetation indices coupled with field spectroscopy and high-resolution multispectral imagery for sites in Cyprus and southern Italy. Traviglia and Cottica (2011) and Traviglia and Torsello (2017) explored the potential of RS applications in the area of Venice, Italy. Moreover, Calleja et al. (2018) used high-resolution commercial imagery combined with unmanned aerial vehicle (UAV) photography to detect buried structures in northern Spain. As further example, Elfadaly, Abouarab, et al. ((2019)) employed historical topographic maps, and Landsat and Sentinel imagery to map settled areas around known sites in the northern Nile Delta in Egypt. RS was also extensively used to assess the conditions of archaeological areas and Cultural Heritage sites for the assessment of risks posed by uncontrolled urban sprawl, natural hazards and looting (Agapiou et al., 2017; Elfadaly et al., 2017; Elfadaly, Attia, & Lasaponara, 2018; Elfadaly et al., 2018; Elfadaly, Abouarab et al., 2019; Elfadaly, Shams Eldein et al., 2019; Elfadaly & Lasaponara, 2019; Lasaponara et al., 2017; Lauricella et al., 2017).

The employment of RS for archaeological purposes in the Near East is also a well-established practice. Aerial images have been used since the 1930s to identify ancient settlements and other sites (Poidebard, 1934). Significant progress has been made since the 1960s and 1970s in particular, due to investigations carried out by large regional archaeological projects that started to analyse aerial photos consistently in order to facilitate or speed up the identification of ancient features in the modern landscape (Adams, 1965, 1981). This in turn led to a notable increase in the use of RS resources, in particular as a valuable tool for site detection. This trend is based on one hand on the constant improvements in the quality of RS products, namely, satellite images, and their increase in number, resolution and accessibility, with open access digital repositories like the Copernicus Open Access Hub and the U.S. Geological Survey. For instance, Wilkinson et al. (2006) used IKONOS and CORONA imagery to explore the area of Homs, Syria; Altaweel (2005) used ASTER imagery to reconstruct the palaeolandscape of Northern Mesopotamia; and Menze and Ur ((2007), (2012), (2013), (2014)) also employed ASTER imagery, with a multitemporal approach to identify anthropogenic soils related to archaeological evidence through the use of artificial intelligence. Hritz (2010) mapped settlements and canals in southern Mesopotamia using declassified CORONA imagery, British military maps and the Shuttle Radar Topography Mission digital elevation model (DEM). Linck et al. (2013) explored the use of Synthetic Aperture Radar (SAR) compared with geophysical data acquired on Roman archaeological sites in Syria. Silver et al. (2015) mapped ancient tracks and pathways in the Syrian area using CORONA, Landsat and Spot imagery. Furthermore, Ansart et al. (2016) used QuickBird imagery and Global Navigation Satellite System (GNSS) field measurements to identify and assess features connected to pastoral settlements in the Leja region, southern Syria. Stewart et al. (2018) performed RS analysis of the Qasrawet site in northern Sinai, while Rayne and Donoghue (2018) focused the use of RS resources to localize traces of canals in Northern Mesopotamia between Syria and Iraq; Jotheri also used RS imagery to try to distinguish between rivers and canals but in the Southern Mesopotamia region (Jotheri, 2020). Jotheri et al. (2019) used CORONA imagery and QuickBird imagery, accessed via Google Earth Pro, to identify hollow ways in southern Iraq. RS has also been crucial for assessing the damage caused to archaeological sites by the conflicts of the last two decades over the entire Near and Middle East (Angiuli et al., 2020; Casana, 2015; Casana & Laugier, 2017; Parcak et al., 2016; Rayne et al., 2017).

An invaluable contribution to research for the reconstruction of ancient Near and Middle East landscapes has come from declassified military CORONA imagery: Acquired for military purposes during the 1960–1972 period by different platforms of the same class, they gained a lot of popularity after their declassification and their combination into the CORONA Atlas (Agapiou et al., 2016; Casana & Cothren, 2013). They have been widely used in recent decades for archaeological purposes in the Near East (Beck et al., 2007; Casana & Cothren, 2008; Challis et al., 2004; Ur, 2003, 2005, 2013b, 2013a), enabling the identification of hundreds of archaeological sites. The advantages they offer include relatively high spatial resolution, although they were originally recorded on film, and the possibility to observe large portions of the Near East (in particular Iraq and Syria) before the extensive use of modern cultivation techniques and above all prior to the urban sprawl that occurred during the second half of the 20th century. Many different approaches were applied to extract information from these images, from fully automated processes to manual analysis (Casana, 2020a, 2020b). Another crucial declassified military RS product is the aerial imagery produced by U2 aircraft, whose missions in the 1950s and 1960s covered large parts of the Near and Middle East and proved to be a formidable source for the reconstruction of their ancient landscapes (Hammer & Ur, 2019).

In addition to the archaeological projects cited (Koliński, 2018; Pfälzner et al., 2016; Ur et al., 2021), RS has been successfully employed in many archaeological investigations in Iraqi Kurdistan. Altaweel and Squitieri (2019) used UAV imagery of the Dinka site, in Sulaymaniyah province, to assess a possible correlation between the presence of stones on the surface of fields and buried structures, subsequently confirmed by magnetometer analysis. Kalayci et al. (2019) explored the spectral response of hollow ways, that is ancient communication routes, in upper Mesopotamia. Soroush et al. (2020) used deep learning to automatically identify on CORONA imagery ancient

qanat in the area of Erbil, while Pirowski et al. (2021) tried to identify the Gaugamela battlefield in the Navkur Plain (the same area considered in this paper) from Pleaides and WorldView-2 imagery, using several processing techniques such as principal component analysis and vegetation indices. Laugier and Casana (2021) mapped archaeological features in the Sirwan region of Iraqi Kurdistan through the combined use of cartographic resources and RS and geophysical surveys. Titolo (2021) used Google Earth Engine© (GEE) to assess the conditions of sites in the areas of the Mosul, Haditha and Hamrin dams in relation to the seasonal changes in the levels of the artificial reservoirs. Laugier et al. (2022) used satellite and aerial images to monitor the damage occurred to archaeological sites in the Sirwan/Upper Diyala River Valley region.

Cloud computing for geospatial products is becoming increasingly popular as the number of available imagery and datasets is constantly growing and offers exceptional opportunities, because it allows the processing of amounts of data that could be hardly managed by normal workstations (Gorelick et al., 2017). GEE (not to be confused with the Google Earth application that has well-established applications in the Cultural Heritage field: Luo et al., (2018)) is a recent online resource for large-scale cloud processing of geospatial data. It has been already experimented for archaeological purposes (Agapiou, 2017; Firpi, 2016), such as the identification of archaeological features in Jordan (Liss et al., 2017); the evaluation of urban sprawl over the area of the Amathus site in Cyprus (Agapiou, 2021) and the city of Matera in Italy (Danese et al., 2021); the assessment of Ethiopian archaeological heritage (Khalaf & Insoll, 2019); the detection of the looting activities on the archaeological site of Apamea in Syria (Agapiou, 2020); the assessment of risks for archaeological sites in Libya and Egypt (Rayne et al., 2020); the reconstruction of palaeolandscape and buried Bronze Age features in the Po Plain in Italy (Brandolini et al., 2021); the estimation of the environmental risks coming from fires and floods on the archaeological site of Metaponto, Italy (Fattore et al., 2021); the hidden parts of the via Appia in Italy (Lasaponara et al., 2022); and the palaeolandscape in proximity of archaeological features in Egypt (Elfadaly et al., 2022).

This paper presents the result of the application of an easy-to-use cloud computing approach with open access RS data for archaeological site identification in three target areas within the LoNAP boundaries. The research can rely on the strong benchmark of LoNAP investigations: the archaeological features found during the Land of Nineveh Archaeological Project have all been checked on the ground. This is a paramount advantage when it comes to assessing RS techniques, because the ground-truthing of sites is always a time-consuming activity that needs careful planning. The aim of this research was not only to evaluate a new RS approach in the LoNAP area but also to verify the possible presence of previously undetected sites. The integration of RS analysis and ground surveys is not always performed during similar researches due to different reasons, including limited resources or restricted access to the areas, and only some projects organize a cross-check on the ground (Campana et al., 2022; Laugier et al., 2022).

2 | MATERIALS AND METHODS

2.1 | Surveyed areas

The Plain of Navkur (the *Mud Plain* in Badînî/Kurdish) is about 165 km^2 of fertile land bordered to the north and east by the southern fringes of the Zagros chains, whereas the southern and western limits are defined by the Jebel Bardarash and Jebel Maqloub mountains (Morandi Bonacossi et al., 2018). It is characterized by deep brown soils crossed by several watercourses (Buringh, 1960), which are mostly seasonal in character, that is, active only during the cold season (winter and early springtime). There are two major rivers, the Nahr al-Khazir and the Gomel Su, which make a significant contribution to regional water availability, shaping the plain with their channels that, unlike most seasonal streams, sometimes cut deeply into the plain (Forti et al., 2021).

Precipitation is also substantial, well above the minimum average necessary for dry farming (Iamoni, 2018). The Navkur area therefore offers ideal conditions for permanent settlement, a trait confirmed by recent regional investigations that found substantial evidence of stable human presence since the beginning of the Neolithic epoch (ca. 10 000 BCE). In order to test alternative and complementary RS methods that might significantly contribute to further explore the archaeological landscape and increase the number of archaeological sites identified in this particular environment, three new subareas were selected in different portions of the Navkur Plain: Area 1 is in the same sector as the Asingeran site, which is the core of the AEP; Area 2 spans the river Gomel; and Area 3 is in the plain between Baadhrah and Ash Shaykhan, close to the first foothills. Together, the three areas cover about 95 km^2, approximately 3% of the LoNAP area (Figure 1). Although different from each other, the study areas feature similar environments—fertile fields with a relatively flat morphology—except for some low natural slopes that are often located close to a *wadi*, one of the frequent temporary streams that run over the entire Navkur. The three areas were selected in part because of the different distributions of known LoNAP archaeological features within them: 41 in Area 1 and nine in Area 2, mostly located along the banks of the river Gomel; there are no known sites in Area 3, although a couple are located just outside its northern limit.

2.2 | Datasets

The RS resources used for this research are all open access data. This means that only medium resolution images have been used, with a spatial resolution ranging between 10 and 30 m. This approach was justified not only by the cost of commercial high-resolution images, but also by the free availability of open access imagery that is revolutionizing the RS field and is widely preferred for this kind of research. The main platforms selected were Landsat 5 (L5), Landsat 7 (L7) and Sentinel-2 (S-2). L5 (USGS Landsat 5) was the fifth satellite of the long-standing Landsat mission, started in 1972; the platform was launched in March 1984 and was decommissioned in June 2013; it

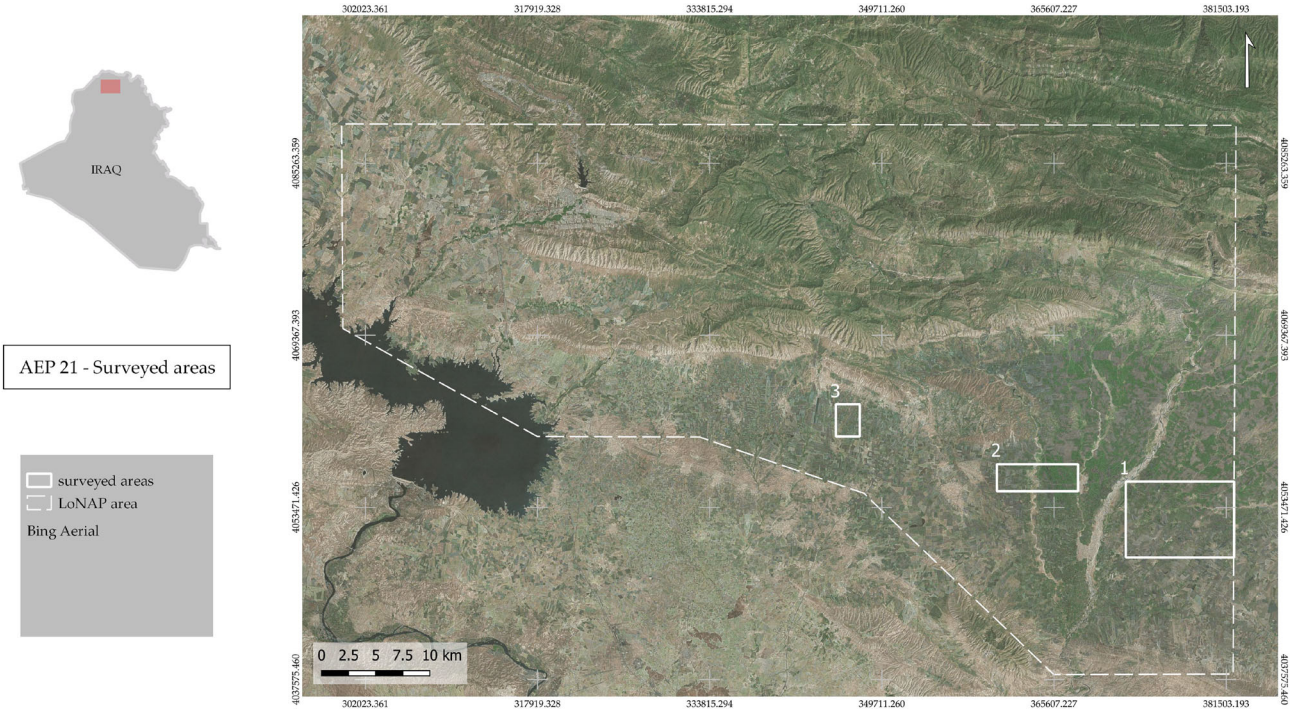

FIGURE 1 The three surveyed areas (solid lines) within the LoNAP survey area (dashed line) [Colour figure can be viewed at wileyonlinelibrary.com]

was a single polar-orbiting satellite with a sun-synchronous orbit and a revisit time of sixteen days at the Equator. It was equipped with Multispectral Scanner and Thematic Mapper instruments that delivered seven bands (radiometric resolution: 0.45–2.35 μm; spatial resolution: 30 m reflective, 120 m thermal). L7 (USGS Landsat 7) is the seventh satellite of the Landsat mission and has the same orbiting characteristics of L5. It is equipped with an Enhanced Thematic Mapper Plus sensor that delivers eight bands (radiometric resolution: 0.45–2.35 μm; spatial resolution: 15 m panchromatic, 30 m reflective, 60 m thermal). It was launched in April 1999 and is still active. S-2 is one of the missions designed and controlled by the European Space Agency; its constellation is composed of two polar-orbiting satellites (Sentinel-2A and Sentinel-2B) with a sun-synchronous orbit and a revisit time of ten days at the Equator, that means 5 days with the two satellites and optimal conditions. S-2 satellites are equipped with a multispectral imaging sensor that delivers 13 bands (radiometric resolution: ~0.443–~2.19 μm; spatial resolution: from 10 to 60 m). All Sentinel missions have an open access policy, so the imagery can be easily retrieved from the Sentinel Open Access Hub. Sentinel-2A was launched in June 2015, while Sentinel-2B in March 2017 (Sentinel 2). The Sentinel-2 imagery has been already used for archaeological and Cultural Heritage applications with positive results (Abate & Lasaponara, 2019; Abate et al., 2020; Agapiou et al., 2014; Tapete & Cigna, 2018; Zanni & De Rosa, 2019).

As will be further shown in detail in the next sections, the chosen platforms guarantee a good trade-off between image resolution, ease of use, computational efficiency and acquisition period.

2.3 | Preliminary analysis

As a first step towards the identification of possible new archaeological sites, we assessed the general appearance of the known sites in the Navkur Plain, which normally looks like clear roundish traces on recent multispectral images; S-2 imagery in particular was selected for a preliminary test. The bands that proved to be more useful for site identification were B4 (red) and B8 (near infrared [NIR]), although the morphology of sites may be also enhanced in the other bands (B2 and B3) of the visible spectrum, RedEdge and shortwave infrared (SWIR). The preliminary identification of some possible regions of interest (ROI) in the survey area was carried out on a single S-2 image acquired on 15 February 2021 with good visibility conditions, available on the Copernicus SciHub and accessible through the EO Browser. However, as already noticed in previous works, this approach presents significant limits: In the Navkur Plain site, site visibility in open fields is subject to many variables, including the amount of moisture in the ground and the presence/absence of plantations. Unlike for other remains, such as settlements with a significant presence of stone or brick structures or ditches, plantations in northern Kurdistan usually conceal the presence of ancient sites; the practice of burning the harvest stubble in the fields contributes to making the surface completely hidden from remote observation for several months per year. This means that their visibility is highly dependent on seasonality for both of the variables mentioned. It is rather difficult, if not impossible, to obtain optimal conditions of visibility for all the sites in just a single satellite image: This method can be used for the assessment of single sites in a certain period but could not be used as the basis for a regional-scale survey,

because many sites would certainly be not recognizable. Single images, retrieved via EO Browser or alternatively GEE, were instead employed to identify the best months for site visibility, using some known LoNAP sites for reference. Winter and wet seasons, when levels of moisture in the soil are higher, are usually considered the best periods for the remote detection of archaeological sites. The visual analysis of monthly images confirmed the assumption; the months from May to September were the worst period for site observation but showed too that site visibility also changes during the favourable seasons, so that it is very unlikely to find the maximum visibility of all the sites contemporaneously, due to their different morphologies and differing local conditions. In order to identify the best period for site detection, a rapid test on GEE using Normalized Difference Water Index (NDWI) values was also carried out.

The NDWI index (Gao, 1996) is used to calculate the moisture in vegetation and soil and is principally applied for agricultural purposes, but it can suggest which could be the most favourable months for the recognition of archaeological features. The average NDWI value over the three survey areas was computed in GEE using the MODIS Combined 16-Days NDWI dataset in the 2012–2021 period (i.e., from the beginning of LoNAP surveys to the present day). MODIS, that stands for 'Moderate-resolution Imaging Spectroradiometer', is the sensor on board the Terra and Aqua satellites, which is used for climate measurements. The analysis of these data (Figure 2) shows that the amount of water in the soil and vegetation in the three targeted areas tends to increase constantly from December to April, when the annual peak occurs, and to continually decrease from May to October–November; this trend appears to be quite regular and to repeat every year, with limited changes. These results helped us to set a seasonal interval from January to April, privileging the satellite images acquired during this interval.

2.4 | Data processing

In order to overcome the visibility issues and improve the results of the remote survey, a multitemporal approach was chosen, using the powerful resources of GEE and its cloud computing.

For this study, three datasets acquired by the L5, L7 and S-2 platforms were selected from the GEE catalog due to the similarities of the sensors, their free availability and popularity, and good coverage of the surveyed area. The L5 and S-2 datasets, respectively, the 'USGS Landsat 5 Level 2, Collection 2, Tier 1' and the 'Sentinel-2 MSI: MultiSpectral Instrument, Level-2A', have atmospherically corrected reflectance values, while the L7 dataset, the 'USGS Landsat 7 Collection 1 Tier 1 TOA Reflectance', has calibrated top-of-atmosphere values. Three reference periods within the last 40 years were chosen as well, one for each dataset: a first decade between the 1980s and 1990s for L5, a second decade between 2000 and 2010 for L7 and a last short period between 2018 and 2021 for S-2. This selection exemplifies three different periods separated by about a decade, which also correspond to three recent epochs in the history of Kurdistan landscape: a relatively 'intact' situation, with small settlements and few main roads, the beginning of development and the present situation. Within these three periods, only the months from January to April were selected, as previously stated. The main characteristics of the datasets are reported in Table 1, while the workflow adopted in this study is shown in Figure 3 and described as follows.

In GEE, instead of using single images as done for the preliminary identification of sites (Section 2.3), all of the images available for the selected periods were chosen and processed. At first, images were masked by the presence of clouds. Then, for the three datasets, a median function was applied to each pixel, separately for each spectral band, in order to get from the multitemporal sequence a single image suitable for the subsequent analyses. Because all the images in a dataset are georeferenced, the product obtained is an image composed of pixels that are not the result of a single acquisition at a certain moment in time but the median of the pixels over time; thus, it is characterized by robust values and free from outliers. All of these operations were performed through the JavaScript console within the GEE environment. This gives certain general advantages: the effects of atmospheric disturbances, the presence of smoke or dust or any other visual obstacle between the ground and the satellite sensor is minimized. Of course, this approach has also a direct impact on site visibility: if a site is visible only for some months per year and the variation in its visibility is not regular, an image that is the result of a

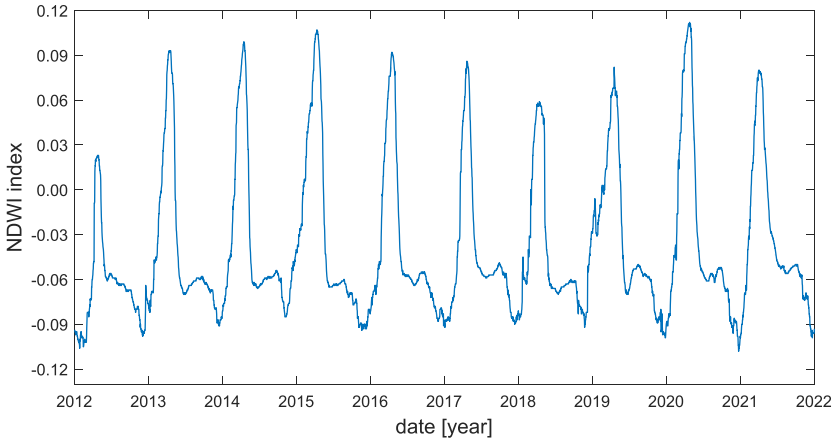

FIGURE 2 2012–2021 NDWI chart over the three investigated areas [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Characteristics of datasets used for this study

	MODIS MCD43A4_006_NDWI	LANDSAT LT05/C02/T1_L2	LANDSAT LE07/C01/T1_TOA	COPERNICUS S2_SR
Time series	2012–2021	1985–1995	2000–2010	2018–2021
Spatial resolution	463.3 m	30 m	15–30 m	10–20–60 m
Radiometric resolution	-	0.45–2.35 μm	0.45–2.35 μm	~0.443–~2.19 μm
No. of processed images	3652	110	137	31

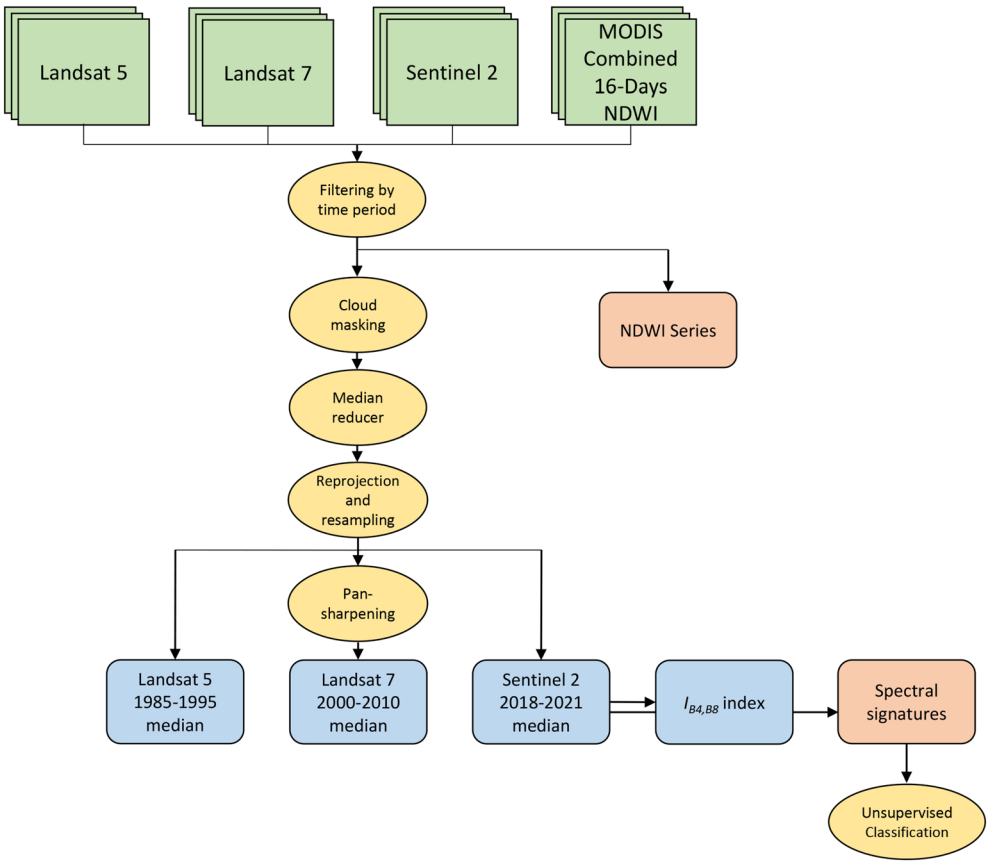

FIGURE 3 Workflow developed for site identification and adopted in GEE [Colour figure can be viewed at wileyonlinelibrary.com]

combination of tens or hundreds of acquisitions will much more likely show it than a single image.

More in detail, with regard to the L5 1984–1994 collection, the final result obtained was the median of 125 images. For the L7 2000–2010 collection (132 images), the same operation as with L5 was carried out; a pan-sharpening with B8 (panchromatic) was also performed, in order to increase the spatial resolution from 30 to 15 m. Although with a series much more limited in time (4 years), the same data processing was also performed with the S-2 dataset on a total of 31 images. For the S-2 dataset, a further product was generated via GEE, developing a basic simple ratio involving B4 and B8:

$$I_{B4,B8} = B4/(B4 - B8). \quad (1)$$

This permitted the enhancement of traces often not recognizable in the visible spectrum, with clear details of the shape and extension of sites (Figure 4). All of the final outputs were also reprojected from EPSG:4326 to EPSG:3395 and resampled on the same grid using the bicubic interpolation method.

After the image generation, the outputs were uploaded in QGIS and each band manually scanned to look for visible anomalies that could indicate possible new sites; a false-colour visualization was used to enhance the anomalies and help the analysis. First of all, the appearance of already known LoNAP sites within the three studied areas was checked; their positions are recorded by point features. In addition to these, some new anomalies were also detected, that represent the ROI for the subsequent analyses. L5 and L7 outputs proved to be most useful when observed in B5 (SWIR), while the S-2 output was observed in the RGB range, B4 (red), B8 (NIR) and $I_{B4,B8}$ (an example is shown in Figure 5). The same ROI were searched on all the outputs generated from the previous processing steps. The spots recognized as possible new sites on the GEE outputs were marked by point features in QGIS for a total number of 42 ROI. This identification was principally carried out considering different parameters such

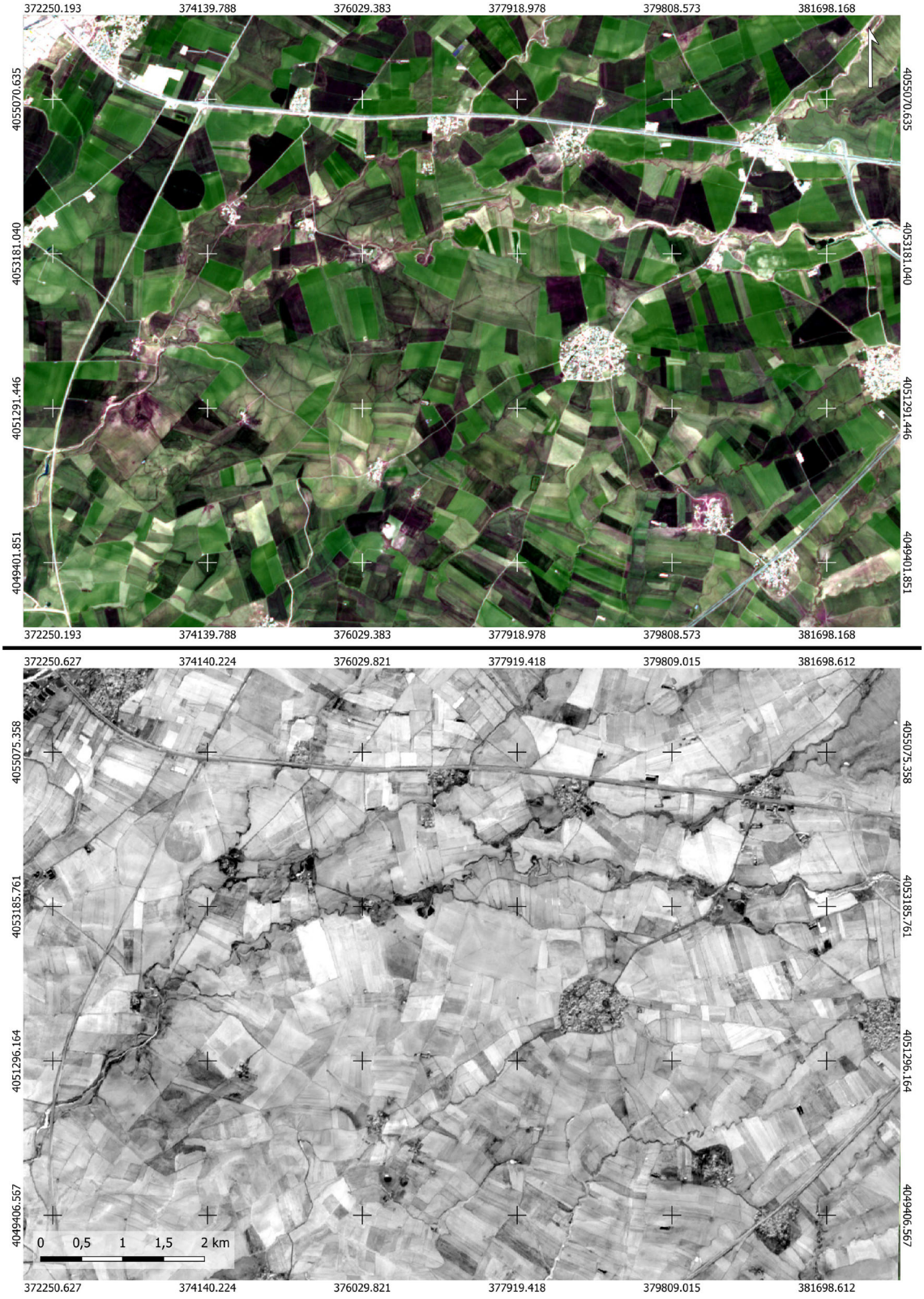

FIGURE 4 Single S-2 RGB image of Area 1 acquired on 15 February 2021 (top); the same area displayed according to the $I_{B4,B8}$ index (bottom) [Colour figure can be viewed at wileyonlinelibrary.com]

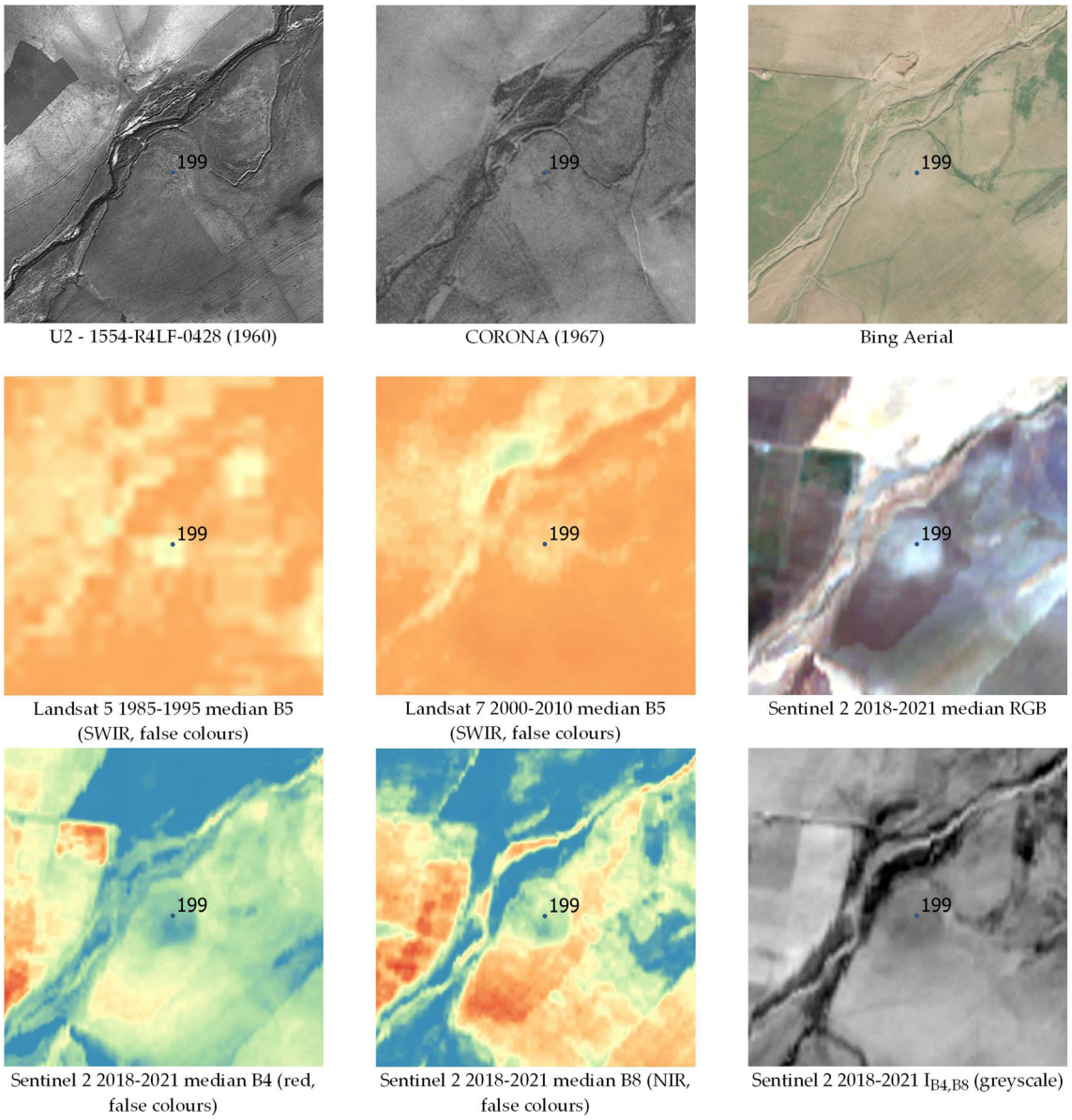

FIGURE 5 Location of Site 199: the second and third rows show the site appearance on images generated in GEE according to the processing steps described in Section 2.4 [Colour figure can be viewed at wileyonlinelibrary.com]

as the shape of the area where the multispectral response was different, the presence of modern structures that could have altered the locality and the proximity of rivers or watercourses.

2.5 | Spectral signature extraction and analysis

Spectral signatures can represent a further clue as to whether the identified regions may or may not correspond to archaeological sites. The S-2 median output was therefore used to extract the spectral signatures of each ROI, by averaging the spectral values associated with the pixels in a 50 m × 50 m area centred on a possible new site; the S-2 dataset was selected for this operation due to the higher spatial resolution of some of its bands. Furthermore, the spectral signatures of already known LoNAP sites located within the three target areas were also analysed. GEE allows easy computation of the spectral signatures for every single site and as averages of multiple different regions: the mean spectral signatures of cultivated fields (that do not represent actual archaeological sites) and of the aforementioned LoNAP sites were also computed as possible references for the new sites.

Statistical classification was then performed on the signatures corresponding to the 42 ROI identified, distinguishing between two clusters. More in detail, common unsupervised classification approaches were tested, namely *k-means* (MacQueen, 1967), *hierarchical clustering* with the minimum-variance linkage criterion (Ward Jr. 1963), *fuzzy c-means* Bezdek (1981) and *self-organizing map (SOM)* (Kohonen, 1982; Maset et al., 2015), applying the Python

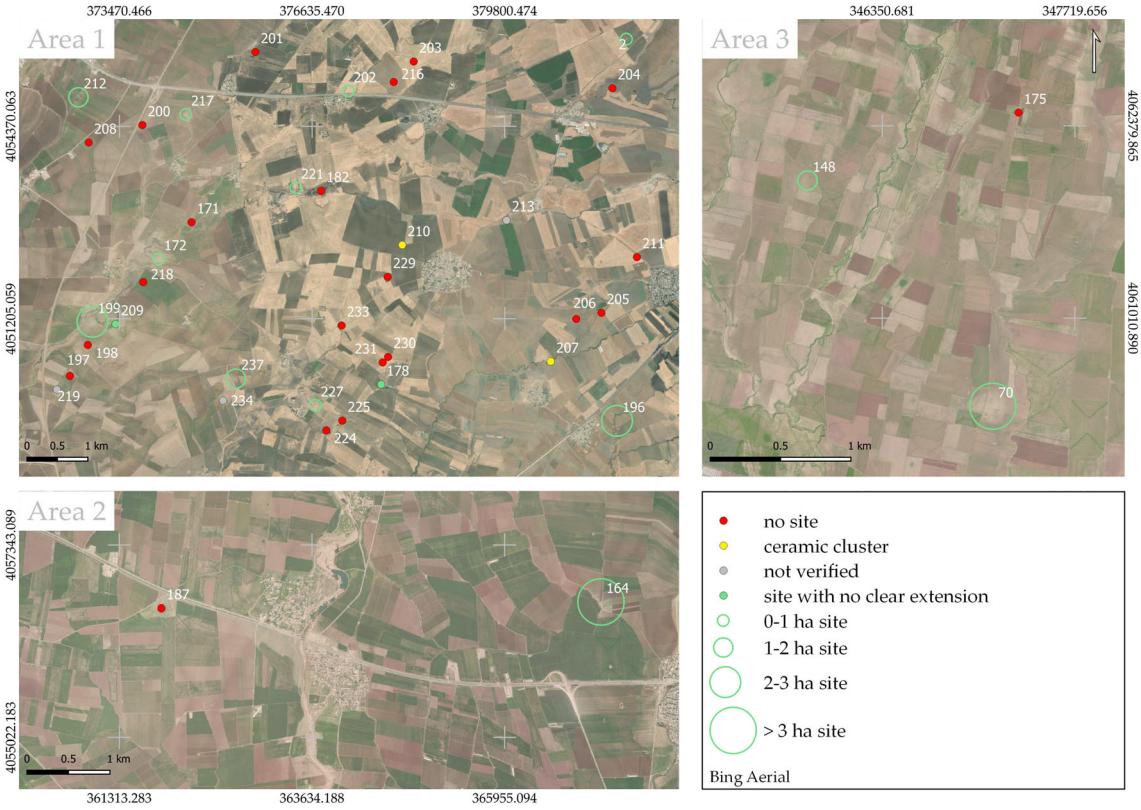

FIGURE 6 Location of surveyed ROI: new sites are in green, ceramic clusters in yellow, negative regions in red and not verified in grey [Colour figure can be viewed at wileyonlinelibrary.com]

implementation provided by the *scikit-learn* library.[1] The cluster whose representative element most closely resembled the average spectral signature computed for the already known LoNAP sites was labelled 'site' and the associated ROI were marked consequently. On the contrary, the ROI whose signatures were assigned to the other cluster were labelled 'no site'. The default values suggested by the library were applied to the algorithm parameters. In fact, limited sensibility on the results was noticed when changing the parameter values and in this a fine tuning was avoided, making the method more generalizable also to other datasets.

2.6 | Ground-truthing

A series of field surveys were subsequently planned to assess the remote observations and verify the results; they were carried out by a team of a minimum of three to a maximum of five persons over six days. The sites were reached using simple smartphone devices where the ROI coordinates had been previously uploaded. Once on the spot, the team members split up to cover the largest possible area in search for potsherds and other artefacts. The spots were documented with photographs, and field notes were also taken on the nature and morphology of the soil. Where possible, the maximum extensions of potsherd scatters were measured. Pottery samples were also collected and later classified to provide a preliminary chronology of the settlements.

3 | RESULTS

Three out of the initial 42 ROI identified on the images, as described in Section 2.4, were not accessible during the field surveys mainly due to the presence of crops. Among the remaining 39 ROI, 17 regions yielded anthropic remains (44%) and were classified as sites (15) or ceramic clusters (2), while 22 furnished no significant anthropic evidence (56%); the results are illustrated in Figure 6. One site (indexed as number 196) is probably to be considered a previous LoNAP site whose exact location was better positioned thanks to the observations on multispectral images. The classification of a ROI as an archaeological site performed on the field was principally based on the widespread presence of potsherds; a limited number of sherds was considered a 'ceramic cluster' if spread over a relatively small area or as 'noise' due to agricultural plowing or proximity to other sites if spread over a very wide area.

The sizes of the new sites are generally small, up to 3 hectares, with some larger exceptions. Their appearance on remote imagery is basically twofold: roundish spots, often lighter than the surrounding soil, or irregular, with no clear shape. Their morphology on the ground could

[1] https://scikit-learn.org/stable/.

be described by three categories: flat sites, that is, sites with no appreciable elevation; artificial mounds, often of very limited altitude and only rarely presenting the characteristics of a *tell*; sites on natural slopes, although these are generally very gentle and of low altitude. At present, no effective correlation between a site's appearance on the images and its ground morphology has been identified. Most of the newly identified sites are located on active *wadis* or palaeorivers, confirming the importance of watercourses for human settlement (Figure 6).

With regard to the RS outputs generated via GEE, the Landsat images proved to be very reliable. The 1985–1995 combined L5 image records a period when the urbanization of the Navkur Plain, and of all of Iraqi Kurdistan, was far lower than today, with small villages and few road connections: This allows the immediate identification of ancient settlements, as they are the only multispectral anomalies, in addition to the few urban centres. By contrast, the relatively low spatial resolution (30 m) prevents clear identification of the medium–small size sites that stud the Plain. The 2000–2010 L7 image, with 15 m spatial resolution, partially solved this problem, providing a better visibility in an environmental context that was still less urbanized than today. The largest sites already mapped by LoNAP are clearly visible in B5 (SWIR), as well as other new sites identified in this work. The S-2 outputs and their 10 m spatial resolution were extremely useful for the assessment of the smaller sites that were hardly or not visible on the Landsat outputs (Figure 5).

The automatic classification of spectral signatures based on unsupervised algorithms provided the results shown in Figure 7. Note that the three sites that were not accessible during the field surveys are not included in this evaluation, whereas the two sites labeled as ceramic clusters are considered as actual archaeological sites. Comparing the unsupervised classification outcomes with the field results, it is worth highlighting that three algorithms, that is, *k-means*, *fuzzy c-means* and SOM, allowed the identification of actual sites with good accuracy, reaching an overall classification performance (i.e., the ratio of correctly classified sites to the total number of ROI investigated) of 69%, 67% and 74%, respectively. The best performance is given by SOM: 12 sites of the 17 identified in the field were correctly classified by the algorithm as true sites, and 17 out of 22 were rightly labeled as 'no site'. The SOM method is also that which yields the lowest rate of false negatives, that is, actual sites that were misclassified as 'no site'.

On the contrary, the *hierarchical clustering* algorithm provides the lowest overall accuracy (54%), with only two out of 17 sites correctly identified. Most of the spectral signatures fall within the 'no site' class, proving that the algorithm is not able to effectively distinguish between the two classes.

4 | DISCUSSION

The case study presented is undoubtedly limited to a relatively small area and number of sites in comparison with the whole LoNAP area, setting the scale of the survey to a microregional level. Despite the spatial size of the datasets, the results are promising and could be easily adapted to larger, unexplored areas. The availability of several products derived from different platforms revealed to be useful because it allowed the cross-checking of anomalies. In fact, anomalies that were clearly visible on images acquired from different platforms and on different bands were more likely to be archaeological sites. During visual inspection of the processed images, it was noticed that some sites apparently showed a 'stronger' spectral response, resulting in a more intense trace being visible on the processed images. This is certainly due to the soil composition, but a future aspect to be investigated is whether this appearance is also due to the 'density' of the archaeological deposit, which might be related to a particular 'monumentality' of the site or its longer life over the centuries. The multitemporal approach and the selection of the sites's period of maximum visibility helped us to overcome one of the most frequent issues in remote identification of archaeological features, that is, limited visibility due to seasonal soil conditions. Furthermore, the ease of the workflow and limited processing time thanks to cloud computing suggest that this tool will spread quickly through the archaeological community, which may well progressively abandon the use of single satellite images and establish new standards for RS analysis.

The processed outputs had a significant impact on the remote identification of sites. In particular, if study areas are visualized using the $I_{B4,B8}$ index, some sites appear to be less visible if compared with seasonal images, but a lot of new features can also be perceived, indicating a potential that will be fully assessed through future surveys and additional research.

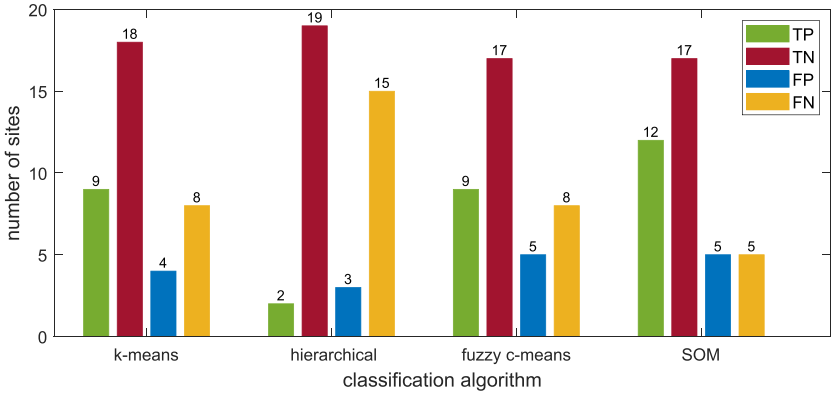

FIGURE 7 Results of the spectral signature classification process for the four methods tested. True positive sites are shown in green, true negative in red, false positive in blue and false negative in yellow [Colour figure can be viewed at wileyonlinelibrary.com]

One of the results obtained is more accurate site recognition within site clusters. The site indexed as 227 is a good example (Figure 8): it is not far (approximately 150 m) from other sites already identified by the LoNAP survey, included the clearly visible *tell* Holarash. Due to the reduced distance between the sites, the scatter of potsherds on the ground surface is almost continuous, allowing distinction between the major sites in the field only on the basis of their elevation. By contrast, using a multispectral image a clear (though small) round trace detached from the other sites is visible, suggesting the presence of a distinct settlement (or portion of a more visible one). If the chronology of finds from Site 227 was similar or equal to those from the surrounding sites, it would be basically invisible during a regular field survey, even after the spatial analysis of collected potsherds.

With regard to the large number of regions that proved not to be sites after the ground-truthing phase (22 out of the 39 ROI that could be verified on the field, corresponding to 56%), this can be partially explained by the nature of the research itself. The surveyed areas are within a region that has been extensively surveyed in recent years, with a considerable number of known sites: slight anomalies that had been excluded at first from the remote surveying carried out during previous campaigns were considered in this study and then checked on the ground. For the same reason, also the number of new ancient settlements (44% of the total) can be considered relatively high. These results would not be directly reproducible over the entire LoNAP survey area, because site distribution is not regular, but some more sites could be certainly identified in the future and added to the number of known archaeological features in Iraqi Kurdistan.

Moreover, this work shows that analyses based on human observation can be effectively supported by the statistical classification of spectral signatures. As highlighted by the results reported in previous section, common unsupervised clustering algorithms can provide reliable indications concerning the probability that a ROI identified on an image is an actual archaeological site and therefore help to prioritize field surveys. Of course, the approximately 70% overall accuracy obtained and above all the number of false negatives (five out of 39 for the best result, provided by SOM) do not allow the field surveying of any region identified on the images as a possible site to be omitted. However, thanks to this approach, ground-truthing activities can be accurately planned, establishing priorities between the different spots and assuring a certain reliability that could help researchers to achieve better results when extensive surveying is not possible for lack of time or resources.

Figure 9 shows that the average spectral signature computed as the mean of the 17 newly identified sites closely corresponds to that of the previously known sites. This outcome suggests also the

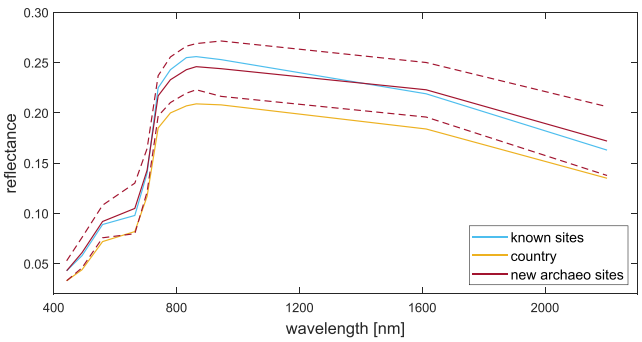

FIGURE 9 Average spectral signature of all new sites identified during the 2021 survey (red) compared with the average spectral signatures of previously known sites in the target areas (cyan) and of cultivated fields (yellow). The red dashed lines show the standard deviation (1σ) from the average values for the new sites [Colour figure can be viewed at wileyonlinelibrary.com]

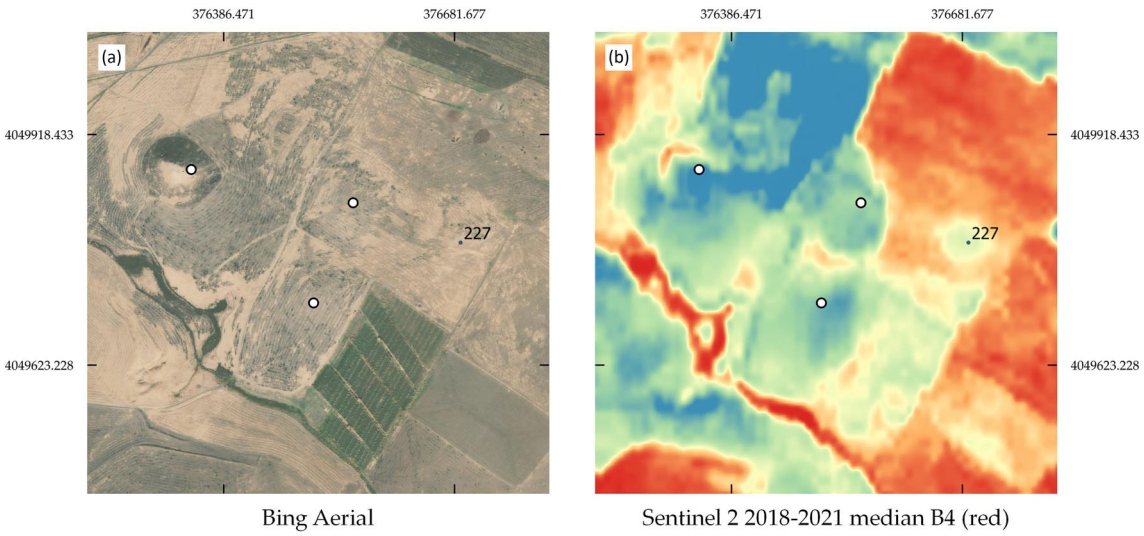

FIGURE 8 *Infra* site recognition: (a) three sites (white points) detected by the LoNAP survey, the eastern one is *tell* Holarash; (b) the same area seen using false colours; a previously undetected small distinct site (227) is clearly visible [Colour figure can be viewed at wileyonlinelibrary.com]

possibility of effectively applying supervised classification techniques in future work, exploiting already verified sites.

A broader aspect to be discussed is the general approach to RS surveys for archaeological purposes. The debate is about automatic versus human-driven analyses, regardless of the nature of the dataset (e.g., satellite rasters and airborne laser scanning). For instance, Casana (2014) questioned the effectiveness of automated tools for site recognition and proposed an approach mainly based on experienced human operators performing manual identification of possible new sites. This is a non-trivial aspect because unsupervised and supervised classification of landcover are widespread approaches—often used to classify the characteristics of soils, forest canopy and so on—and are necessary when working on regional or planetary scales. The automatic recognition of anthropogenic soils, that suggests the presence of archaeological sites, can currently be performed, taking advantage of the multispectral data of satellite imagery. Many examples already exist: for instance, the automatic extraction of archaeological features has been tested by De Laet et al. (2007) on the site of Hisar (Turkey), by Alexakis et al. (2009) on Neolithic settlements in Thessaly (Greece), Trier et al. (2009) on burial mounds in southeastern Norway, Harrower et al. (2013) and Schuetter et al. (2013) on lithic tombs in southern Arabia and Toumazet et al. (2017) on rural structures in France. An up-to-date overview of object-based methods can be also found in Davis (2018). Some quick tests of supervised classification carried out on the three surveyed areas using Classification and Regression Trees (CART) and Random Forest algorithms implemented in GEE on the S-2 outputs did not give clear and useful results, with a large number of riverbeds and cultivated fields put in the same class as known and possible sites.

Also, the unsupervised classification of spectral signatures conducted during this study, notwithstanding the undeniable contribution that it can make to the remote exploration of a region, produced false positives within the ROI, and a future more in-depth analysis of spectral signatures of new sites will be necessary to further reduce their presence. False positives and false negatives actively influence the planning of an archaeological field survey that involves a ground-truthing phase, because the verification of the analysis requires physical examination of the detected ROI, an activity that needs a significant amount of time, especially when dealing with hundreds of areas to be checked. In the present case study, five sites identified by the SOM algorithm turned out to be false positives; the application of more advanced RS techniques could reduce this degree of uncertainty. Human operators can deal relatively well with variables such as the morphology of the observed anomalies and their position with respect to other features (closeness to waterways, known sites, etc.) that would require a major work of implementation to be completely automatized. But as highlighted by Davis (2018), with very large portions of territory to be explored, a totally manual approach would be too time-consuming. With standard available resources and skills, a mixed approach combining an expert-led and automatic identification of sites would probably be the right solution.

Considering the specific context of Near Eastern Archaeology, many RS analyses have relied on CORONA imagery because of its acquisition period and good spatial resolution (Beck et al., 2007; Casana, 2020a; Casana & Cothren, 2013, 2008; Challis et al., 2004; Ur, 2003, 2013a). Although the first of these is still a good reason for using it, this study shows that the use of open access datasets with medium spatial resolution (30 m and especially 10 m images) is also feasible for the identification of small and medium size sites, especially in flat terrain. The advantages of digital multispectral images and the possibilities offered by their customized processing compensate for loss of detail that could be in any case regained using commercial high-resolution imagery. The unique value of historical RS imagery, such as CORONA and U2, is in the landscape they recorded, which is completely different from today's, with many ancient sites clearly visible. Its use should be undoubtedly encouraged for studies on historical landscapes and for site identification in areas that have changed a lot in recent decades, such as zones of urban sprawl around the main city centres. For recognition of new sites in relatively intact areas instead, priority should be given to current digital multispectral imagery analysed with a multitemporal approach. Cloud computing, regardless of the service provider, is a crucial tool for advanced processing and making use of large imagery datasets.

5 | CONCLUSIONS

Despite its relatively small scale range, the 2021 survey based on RS resources and conducted in proximity to the excavated site of Asingeran within the LoNAP survey area had positive outcomes. On the archaeological side, 17 previously undetected archaeological sites were identified, improving our knowledge of Neolithic to Middle Ages settlement patterns in the Navkur Plain. Although no high-resolution imagery was used, the relatively small size of the new identified sites suggests that medium resolution images are nonetheless useful for the identification of sites of modest size in flat areas and that high-resolution images can be used in a further phase of the research, for specific targets or the detection of smaller archaeological features (such as isolated graves or architectural elements). This is an important conclusion, because it means that wider archaeological remote surveys can be successfully carried out worldwide without using commercial imagery. The opportunity of directly checking on the ground the results of the RS survey represented the added value that allowed to verify on the field what was acquired remotely and to present more accurate results.

From a broader perspective, this study confirms the benefits of cloud-based computing of RS products for archaeological purposes. The availability of more than 40 years of multispectral imagery that is remotely accessible offers a paramount archive that would be difficult to collect because of its size and the time needed to build and maintain it. An even greater impact is the opportunity to use supercomputing through a relatively friendly interface for users with basic coding skills. This is particularly crucial for archaeological studies, because 'traditional' supercomputing is usually limited to large projects and needs to be planned and performed by highly skilled personnel; moreover, computing facilities are often national interest sites, with strict

working schedules and limited accessibility, while cloud computing allows single users to utilize this technology just with standard machines and an internet connection. GEE interface allows performance of all the operations presented in one environment, and outputs can be downloaded for further analysis with open source and proprietary software. As online data archives are constantly updated with new images acquired by active platforms, the results achieved can be also updated, facilitating the monitoring of variables over time. Large-scale cloud computing can boost the use of RS products and researches among the Near Eastern archaeological community, and among the global archaeological community as well, because no particular hardware or software requirements are needed. Multitemporal analysis, customization of RS products depending on the aims of the research, and the availability of large online datasets are extremely important advantages, especially for archaeologists whose use of RS sources is still highly dependent on free imagery that offers little control over selection and processing.

The statistical classification methods applied in this study helped to interpret spectral signatures at 70% overall accuracy, with advantages for the efficient planning of ground-truthing survey activities. As a future work, the potential of supervised classification algorithms will be also investigated, leveraging on the spectral signatures of already verified sites. The availability of a larger number of training samples would also allow to test the possibility of identifying different types of sites through the classification process.

Although similar results from the direct application of the processing workflow presented in another regional area are not guaranteed, the local tuning and improvement of the code used could overcome issues related to differences in site and soil morphology, broadening the impact of large-scale computing applied to RS for archaeological purposes.

ACKNOWLEDGEMENTS

The authors wish to thank the General Directorate of Kurdish Antiquities and the Directorate of Antiquities of Dohuk for their support, the Land of Nineveh Archaeological Project Team for their cooperation, Ylenia Borgonovo and Francesco Venturoso who partecipated in the field survey, Francesca Simi for her help with CORONA and U2 imagery, Mattia Previtali for his useful advices and the anonymous reviewers for their comments. This research was funded by the University of Udine, the Italian Ministry of Foreign Affairs and International Cooperation and the Italian Ministry of University and Research. Open Access Funding provided by Universita degli Studi di Udine within the CRUI-CARE Agreement.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

Satellite imagery is accessible at https://earthengine.google.com.

ORCID

Riccardo Valente https://orcid.org/0000-0001-8620-0879

REFERENCES

Abate, N., Elfadaly, A., Masini, N., & Lasaponara, R. (2020). Multitemporal 2016-2018 Sentinel-2 data enhancement for landscape archaeology: The case study of the Foggia province, Southern Italy. *Remote Sensing*, *12*(8), 1309. https://www.mdpi.com/2072-4292/12/8/1309

Abate, N., & Lasaponara, R. (2019). Preventive archaeology based on open remote sensing data and tools: The cases of Sant'Arsenio (SA) and Foggia (FG), Italy. *Sustainability*, *11*(15), 4145. https://www.mdpi.com/2071-1050/11/15/4145

Adams, R. M. (1965). Land behind Baghdad. A history of settlement on the Diyala plains.

Adams, R. M. (1981). Heartland of cities: Surveys of ancient settlement and land use on the central floodplain of the euphrates.

Agapiou, A. (2017). Remote sensing heritage in a petabyte-scale: Satellite data and heritage Earth Engine© applications. *International Journal of Digital Earth*, *10*(1), 85–102.

Agapiou, A. (2020). Detecting looting activity through earth observation multi-temporal analysis over the archaeological site of Apamea (Syria) during 2011–2012. *Journal of Computer Applications in Archaeology*, *3*(1), 219–237.

Agapiou, A. (2021). Multi-temporal change detection analysis of vertical sprawl over Limassol city centre and Amathus archaeological site in Cyprus during 2015–2020 Using the Sentinel-1 Sensor and the Google Earth Engine Platform. *Sensors*, *21*(5), 1884. https://www.mdpi.com/1424-8220/21/5/1884

Agapiou, A., Alexakis, D., Sarris, A., & Hadjimitsis, D. (2014). Evaluating the potentials of Sentinel-2 for archaeological perspective. *Remote Sensing*, *6*(3), 2176–2194. http://www.mdpi.com/2072-4292/6/3/2176

Agapiou, A., Alexakis, D. D., Sarris, A., & Hadjimitsis, D. G. (2016). Colour to greyscale pixels: Re-seeing greyscale archived aerial photographs and declassified satellite CORONA images based on image fusion techniques. *Archaeological Prospection*, *23*(4), 231–241.

Agapiou, A., Hadjimitsis, D. G., Themistocleous, K., Papadavid, G., & Toulios, L. (2010). Detection of archaeological crop marks in Cyprus using vegetation indices from landsat TM/ETM plus satellite images and field spectroscopy measurements. In Michel, U., & Civco, D. L. (Eds.), *SPIE Proceedings* (pp. 78310V): SPIE. https://doi.org/10.1117/12.864935

Agapiou, A., Lysandrou, V., & Hadjimitsis, D. (2017). Optical remote sensing potentials for looting detection. *Geosciences*, *7*(4), 98. http://www.mdpi.com/2076-3263/7/4/98

Agapiou, A., Lysandrou, V., Lasaponara, R., Masini, N., & Hadjimitsis, D. (2016). Study of the variations of archaeological marks at neolithic site of Lucera, Italy using high-resolution multispectral datasets. *Remote Sensing*, *8*(9), 723. http://www.mdpi.com/2072-4292/8/9/723

Alexakis, D., Sarris, A., Astaras, T., & Albanakis, K. (2009). Detection of neolithic settlements in Thessaly (Greece) through multispectral and hyperspectral satellite imagery. *Sensors*, *9*(2), 1167–1187. http://www.mdpi.com/1424-8220/9/2/1167

Altaweel, M. (2005). The use of ASTER satellite imagery in archaeological contexts. *Archaeological Prospection*, *12*(3), 151–166.

Altaweel, M., & Squitieri, A. (2019). Finding a relatively flat archaeological site with minimal ceramics: A case study from Iraqi Kurdistan. *Journal of Field Archaeology*, *44*(8), 523–537.

Angiuli, E., Pecharromán, E., Ezquieta, P. V., Gorzynska, M., & Ovejanu, I. (2020). Satellite imagery-based damage assessment on Nineveh and Nebi Yunus archaeological site in Iraq. *Remote Sensing*, *12*(10), 1672.

Ansart, A., Braemer, F., & Davtian, G. (2016). Preparing an archaeological field survey: Remote sensing interpretation for herding structures in the Southern Levant. *Journal of Field Archaeology*, *41*(6), 699–712.

Aqua. https://aqua.nasa.gov/

Beck, A., Philip, G., Abdulkarim, M., & Donoghue, D. (2007). Evaluation of Corona and Ikonos high resolution satellite imagery for archaeological prospection in western Syria. *Antiquity*, *81*(311), 161–175.

https://www.cambridge.org/core/product/identifier/S0003598X00094916/type/journal_article

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*: Plenum Press.

Bonacossi, D. M., & Iamoni, M. (2015). Landscape and settlement in the Eastern Upper Iraqi Tigris and Navkur Plains: The Land of Nineveh Archaeological Project, Seasons 2012-2013. *Iraq*, *77*, 9–39. https://www.cambridge.org/core/product/identifier/S0021088915000054/type/journal_article

Brandolini, F., Domingo-Ribas, G., Zerboni, A., & Turner, S. (2021). A Google Earth Engine-enabled Python approach for the identification of anthropogenic palaeo-landscape features. *Open Research Europe*, *1*, 22. https://open-research-europe.ec.europa.eu/articles/1-22/v2

Buringh, P. (1960). *Soils and soil conditions in Iraq*: Ministry of Agriculture.

Calleja, J. F., Requejo Pagés, O., Díaz-Álvarez, N., Peón, J., Gutiérrez, N., Martín-Hernández, E., Cebada Relea, A., Rubio Melendi, D., & Fernández Álvarez, P. (2018). 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*, 555–573. https://linkinghub.elsevier.com/retrieve/pii/S0303243418304410

Campana, S., Sordini, M., Berlioz, S., Vidale, M., Al-Lyla, R., Al Araj, A. A., & Bianchi, A. (2022). Remote sensing and ground survey of archaeological damage and destruction at Nineveh during the ISIS occupation. *Antiquity*, *96*(386), 436–454.

Casana, J. (2014). Regional-scale archaeological remote sensing in the age of big data: Automated site discovery vs. brute force methods. *Advances in Archaeological Practice*, *2*(3), 222–233.

Casana, J. (2015). Satellite imagery-based analysis of archaeological looting in Syria. *Near Eastern Archaeology*, *78*(3), 142–152.

Casana, J. (2020a). Global-scale archaeological prospection using CORONA satellite imagery: Automated, crowd-sourced, and expert-led approaches. *Journal of Field Archaeology*, *45*(sup1), S89–S100.

Casana, J. (2020b). Remote sensing-based approaches to site morphology and historical geography in the northern fertile crescent. In Lawrence, D., Altaweel, M., & Philip, G. (Eds.), *New agendas in remote sensing and landscape archaeology in the Near East. Studies in honour of Tony J. Wilkinson*: Archaeopress, pp. 154–174.

Casana, J., & Cothren, J. (2008). Stereo analysis, DEM extraction and orthorectification of CORONA satellite imagery: Archaeological applications from the Near East. *Antiquity*, *82*(317), 732–749. https://www.cambridge.org/core/product/identifier/S0003598X00097349/type/journal_article

Casana, J., & Cothren, J. (2013). The CORONA Atlas Project: Orthorectification of CORONA satellite imagery and regional-scale archaeological exploration in the Near East. In Comer, D. C., & Harrower, M. J. (Eds.), *Mapping archaeological landscapes from space*, Springer Briefs in Archaeology, Vol. 5: Springer New York, pp. 33–43.

Casana, J., & Laugier, E. J. (2017). Satellite imagery-based monitoring of archaeological site damage in the Syrian civil war. *PLOS ONE*, *12*(11), e0188589.

Challis, K., Priestnall, G., Gardner, A., Henderson, J., & O'Hara, S. (2004). Corona remotely-sensed imagery in dryland archaeology: The Islamic city of al-Raqqa, Syria. *Journal of Field Archaeology*, *29*(1–2), 139–153.

Copernicus Open Access Hub. https://www.scihub.copernicus.eu/dhus/#/home

Coppini, C. (2018). The Land of Nineveh Archaeological Project: Preliminary results from the analysis of the second millennium BC pottery. In Salisbury, R. A., Höflmayer, F., & Bürge, T. (Eds.), *Proceedings of the 10th International Congress on the Archaeology of the Ancient Near East, Volume 2: Prehistoric and Historical Landscapes & Settlement Patterns*: Harrassowitz Verlag, pp. 65–82.

Danese, M., Gioia, D., & Biscione, M. (2021). Integrated methods for cultural heritage risk assessment: Google Earth Engine, spatial analysis, machine learning. In Gervasi, O., Murgante, B., Misra, S., et al. (Eds.), *Computational science and its applications–ICCSA 2021*, Lecture Notes in Computer Science, Vol. 12951: Springer International Publishing, pp. 605–619.

Davis, D. S. (2018). Object-based image analysis: A review of developments and future directions of automated feature detection in landscape archaeology. *Archaeological Prospection*, *26*(2), 155–163.

De Laet, V., Paulissen, E., & Waelkens, M. (2007). Methods for the extraction of archaeological features from very high-resolution Ikonos-2 remote sensing imagery, Hisar (southwest Turkey). *Journal of Archaeological Science*, *34*(5), 830–841. https://linkinghub.elsevier.com/retrieve/pii/S0305440306001774

Elfadaly, A., Abouarab, M. A. R., El Shabrawy, R. R. M., Mostafa, W., Wilson, P., Morhange, C., Silverstein, J., & Lasaponara, R. (2019). Discovering potential settlement areas around archaeological tells using the integration between historic topographic maps, optical, and radar data in the Northern Nile Delta, Egypt. *Remote Sensing*, *11*(24), 3039. https://www.mdpi.com/2072-4292/11/24/3039

Elfadaly, A., Attia, W., & Lasaponara, R. (2018). Monitoring the environmental risks around Medinet Habu and Ramesseum Temple at West Luxor, Egypt, using remote sensing and GIS techniques. *Journal of Archaeological Method and Theory*, *25*(2), 587–610.

Elfadaly, A., Attia, W., Qelichi, M. M., Murgante, B., & Lasaponara, R. (2018). Management of cultural heritage sites using remote sensing indices and spatial analysis techniques. *Surveys in Geophysics*, *39*(6), 1347–1377.

Elfadaly, A., & Lasaponara, R. (2019). On the use of satellite imagery and GIS tools to detect and characterize the urbanization around heritage sites: The case studies of the catacombs of Mustafa Kamel in Alexandria, Egypt and the Aragonese Castle in Baia, Italy. *Sustainability*, *11*(7), 2110. https://www.mdpi.com/2071-1050/11/7/2110

Elfadaly, A., Shams, A. H., Elbehery, W., Elftatry, M., Wafa, O., Hiekl, A. M. A., Wilson, P., Silverstein, J., & Abouarab, M. A. R. (2022). Revealing the paleolandscape features, around the archaeological sites in the northern Nile Delta of Egypt using radar satellite imagery and GEE platform. *Archaeological Prospection*, *29*(3), 369–384.

Elfadaly, A., Shams Eldein, A., & Lasaponara, R. (2019). Cultural heritage management using remote sensing data and GIS techniques around the archaeological area of ancient Jeddah in Jeddah City, Saudi Arabia. *Sustainability*, *12*(1), 240. https://www.mdpi.com/2071-1050/12/1/240

Elfadaly, A., Wafa, O., Abouarab, M., Guida, A., Spanu, P., & Lasaponara, R. (2017). Geo-environmental estimation of land use changes and its effects on Egyptian temples at Luxor City. *ISPRS International Journal of Geo-Information*, *6*(11), 378. http://www.mdpi.com/2220-9964/6/11/378

EO Browser. https://apps.sentinel-hub.com/eo-browser/

Fattore, C., Abate, N., Faridani, F., Masini, N., & Lasaponara, R. (2021). Google Earth Engine as multi-sensor open-source tool for supporting the preservation of archaeological areas: The case study of flood and fire mapping in Metaponto, Italy. *Sensors*, *21*(5), 1791. https://www.mdpi.com/1424-8220/21/5/1791

Firpi, O. A. A. (2016). Satellite data for all? Review of Google Earth Engine for archaeological remote sensing. *Internet Archaeology*, *42*. http://intarch.ac.uk/journal/issue42/10/firpi.html

Forti, L., Perego, A., Brandolini, F., Mariani, G. S., Zebari, M., Nicoll, K., Regattieri, E., Barbaro, C. C., Bonacossi, D. M., Qasim, H. A., Cremaschi, M., & Zerboni, A. (2021). Geomorphology of the northwestern Kurdistan Region of Iraq: landscapes of the Zagros Mountains drained by the Tigris and Great Zab Rivers. *Journal of Maps*, *17*(2), 225–236.

Gallo, D., Ciminale, M., Becker, H., & Masini, N. (2009). Remote sensing techniques for reconstructing a vast Neolithic settlement in Southern Italy. *Journal of Archaeological Science*, *36*(1), 43–50. https://linkinghub.elsevier.com/retrieve/pii/S0305440308001659

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of*

Environment, *58*(3), 257–266. https://linkinghub.elsevier.com/retrieve/pii/S0034425796000673

Gavagnin, K. (2016). The Land of Nineveh Archaeological Project: A preliminary overview on the pottery and settlement patterns in the 3rd millennium BC in the northern region of Iraqi Kurdistan. In Kopanias, K., & MacGinnis, J. (Eds.), *Archaeological research in the Kurdistan and adjacent regions*, BAR International Series: Archaeopress, pp. 75–86.

Gavagnin, K., Iamoni, M., & Palermo, R. (2016). The Land of Nineveh Archaeological Project: The ceramic repertoire from the Early Pottery Neolithic to the Sasanian Period. *Bulletin of the American Schools of Oriental Research*, *375*, 119–169.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://linkinghub.elsevier.com/retrieve/pii/S0034425717302900

Grøn, O., Palmér, S., Stylegar, F.-A., Esbensen, K., Kucheryavski, S., & Aase, S. (2011). Interpretation of archaeological small-scale features in spectral images. *Journal of Archaeological Science*, *38*(9), 2024–2030. https://linkinghub.elsevier.com/retrieve/pii/S0305440309004415

Hammer, E., & Ur, J. (2019). Near eastern landscapes and declassified U2 aerial imagery. *Advances in Archaeological Practice*, *7*(2), 107–126. https://www.cambridge.org/core/product/identifier/S2326376818000384/type/journal_article

Harrower, M. J., Schuetter, J., McCorriston, J., Goel, P. K., & Senn, M. J. (2013). Survey, automated detection, and spatial distribution analysis of cairn tombs in ancient Southern Arabia. In Comer, D. C., & Harrower, M. J. (Eds.), *Mapping archaeological landscapes from space* (pp. 259–268), Springer Briefs in Archaeology, Vol. 5: Springer New York.

Herrmann, J. T., Glissmann, B., Sconzo, P., & Pfälzner, P. (2018). Unmanned aerial vehicle (UAV) survey with commercial-grade instruments: A case study from the Eastern Habur Archaeological Survey, Iraq. *Journal of Field Archaeology*, *43*(4), 269–283.

Hritz, C. (2010). Tracing settlement patterns and channel systems in Southern Mesopotamia using remote sensing. *Journal of Field Archaeology*, *35*(2), 184–203.

Iamoni, M. (2016). Across millennia of occupation. The Land of Nineveh Archaeological project in Iraqi Kurdistan. The prehistory and protohistory of the Upper Tigris rediscovered, pp. 125–134.

Iamoni, M. (2018). A vital resource in prehistory water and settlement during the Pottery Neolithic and Chalcolithic periods. A preliminary analysis of the Eastern Upper Tigris basin area. In Kühne, H. (Ed.), *Water for assyria*: Harrassowitz, pp. 7–24.

Iamoni, M., & Qasim, H. A. (2021). Asingeran, a Neolithic and Chalcolithic "Iceberg" in Northern Mesopotamia. *Origini*, *43*, 9–34.

Jotheri, J. (2020). Recognition of ancient channels and archaeological sites in the Mesopotamian floodplain using satellite imagery and digital topography. In Zhuang, Y., & Altaweel, M. (Eds.), *Water societies and technologies from the past and present*: Archaeopress, pp. 283–305.

Jotheri, J., de Gruchy, M., Almaliki, R., & Feadha, M. (2019). Remote sensing the archaeological traces of boat movement in the marshes of Southern Mesopotamia. *Remote Sensing*, *11*(21), 2474. https://www.mdpi.com/2072-4292/11/21/2474

Kalayci, T., Lasaponara, R., Wainwright, J., & Masini, N. (2019). Multispectral contrast of archaeological features: A quantitative evaluation. *Remote Sensing*, *11*(8), 913. https://www.mdpi.com/2072-4292/11/8/913

Khalaf, N., & Insoll, T. (2019). Monitoring Islamic archaeological landscapes in Ethiopia using open source satellite imagery. *Journal of Field Archaeology*, *44*(6), 401–419.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69.

Koliński, R. (2015). The use of satellite imagery in an archaeological survey in Iraqi Kurdistan. *New World Archaeology*, *9*, 113–122.

Koliński, R. (2018). An archaeological reconnaissance in the Greater Zab Area of the Iraqi Kurdistan (UGZAR) 2012-2015. In Horejs, B., Schwall, C., Müller, V., et al. (Eds.): Harrassowitz Verlag, pp. 13–26.

Kopanias, K., MacGinnis, J., & Ur, J. A. (2015). Archaeological projects in the Kurdistan Region in Iraq. (*Tech. Rep.*): The General Directorate of Antiquities of the Kurdistan Regional Government. http://nrs.harvard.edu/urn-3:HUL.InstRepos:14022526

Lasaponara, R., Abate, N., & Masini, N. (2022). On the use of Google Earth Engine and sentinel data to detect "lost" sections of ancient roads. The case of Via Appia. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. https://ieeexplore.ieee.org/document/9349248/

Lasaponara, R., & Masini, N. (2006). Identification of archaeological buried remains based on the normalized difference vegetation index (NDVI) from Quickbird satellite data. *IEEE Geoscience and Remote Sensing Letters*, *3*(3), 325–328. http://ieeexplore.ieee.org/document/1657998/

Lasaponara, R., & Masini, N. (2007). Detection of archaeological crop marks by using satellite QuickBird multispectral imagery. *Journal of Archaeological Science*, *34*(2), 214–221. https://linkinghub.elsevier.com/retrieve/pii/S0305440306000951

Lasaponara, R., & Masini, N. (2012). *Satellite remote sensing—A new tool for archaeology*: Springer.

Lasaponara, R., Murgante, B., Elfadaly, A., Qelichi, M., Shahraki, S., Wafa, O., & Attia, W. (2017). Spatial open data for monitoring risks and preserving archaeological areas and landscape: Case studies at Kom el Shoqafa, Egypt and Shush, Iran. *Sustainability*, *9*(4), 572. http://www.mdpi.com/2071-1050/9/4/572

Laugier, E. J., Abdullatif, N., & Glatz, C. (2022). Embedding the remote sensing monitoring of archaeological site damage at the local level: Results from the "Archaeological practice and heritage protection in the Kurdistan Region of Iraq" project. *PLOS ONE*, *17*(6), e0269796.

Laugier, E. J., & Casana, J. (2021). Integrating satellite, UAV, and ground-based remote sensing in archaeology: An exploration of pre-modern land use in Northeastern Iraq. *Remote Sensing*, *13*(24), 5119. https://www.mdpi.com/2072-4292/13/24/5119

Lauricella, A., Cannon, J., Branting, S., & Hammer, E. (2017). Semi-automated detection of looting in Afghanistan using multispectral imagery and principal component analysis. *Antiquity*, *91*(359), 1344–1355.

Linck, R., Busche, T., Buckreuss, S., Fassbinder, J. W. E., & Seren, S. (2013). Possibilities of archaeological prospection by high-resolution X-band satellite radar—A case study from Syria: Determination of the penetration depth of TerraSAR-X. *Archaeological Prospection*, *20*(2), 97–108.

Liss, B., Howland, M. D., & Levy, T. E. (2017). Testing Google Earth Engine for the automatic identification and vectorization of archaeological features: A case study from Faynan, Jordan. *Journal of Archaeological Science: Reports*, *15*, 299–304. https://linkinghub.elsevier.com/retrieve/pii/S2352409X16308082

Luo, L., Wang, X., Guo, H., Lasaponara, R., Shi, P., Bachagha, N., Li, L., Yao, Y., Masini, N., Chen, F., Ji, W., Cao, H., Li, C., & Hu, N. (2018). Google Earth as a powerful tool for archaeological and cultural heritage applications: A review. *Remote Sensing*, *10*(10), 1558. http://www.mdpi.com/2072-4292/10/10/1558

Luo, L., Wang, X., Guo, H., Lasaponara, R., Zong, X., Masini, N., Wang, G., Shi, P., Khatteli, H., Chen, F., Tariq, S., Shao, J., Bachagha, N., Yang, R., & Yao, Y. (2019). Airborne and spaceborne remote sensing for archaeological and cultural heritage applications: A review of the century (1907–2017). *Remote Sensing of Environment*, *232*, 111280. https://linkinghub.elsevier.com/retrieve/pii/S0034425719302998

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Le Cam, L. M., & Neyman, J. (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, (Vol. 1) (pp. 281–297). University of California.

Maset, E., Carniel, R., & Crosilla, F. (2015). Unsupervised classification of raw full-waveform airborne lidar data by self organizing maps, pp. 62–72.

Masini, N., & Lasaponara, R. (2017). Sensing the past from space: Approaches to site detection. In Masini, N., & Soldovieri, F. (Eds.), *Sensing the past*, Geotechnologies and the Environment, Vol. 16: Springer International Publishing, pp. 23–60.

Menze, B. H., & Ur, J. (2007). Classification of multispectral imagery in archaeological settlement survey in the Near East. In *Proceedings of the ISPRS Working Group VII/1 Workshop ISPMSRS'07: "Physical Measurements and Signatures in Remote Sensing"*, pp. 244–249.

Menze, B. H., & Ur, J. A. (2012). Mapping patterns of long-term settlement in Northern Mesopotamia at a large scale. *Proceedings of the National Academy of Sciences*, *109*(14), E778–E787.

Menze, B. H., & Ur, J. A. (2013). Multi-temporal classification of multispectral images for settlement survey in northeastern Syria, *Mapping archaeological landscapes from space*: Springer New York, pp. 219–228.

Menze, B. H., & Ur, J. A. (2014). Multitemporal fusion for the detection of static spatial patterns in multispectral satellite Images—With application to archaeological survey. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(8), 3513–3524. http://ieeexplore.ieee.org/document/6907965/

Morandi Bonacossi, D. (2016). The Land of Niniveh Archaeological Project. Assyrian settlement in the Niniveh Hinterland: A view from the centre. In MacGinnis, J., Wicke, D., & Greenfield, T. (Eds.), *The provincial archaeology of the Assyrian Empire*: McDonald Institute for Archaeological Research, pp. 141–150.

Morandi Bonacossi, D., Qasim, H. A., Coppini, C., Gavagnin, K., Iamoni, M., & Tonghini, C. (2018). The Italian-Kurdish excavations at Gir-e Gomel in the Kurdistan region of Iraq. Preliminary report on the 2017 and 2018 field seasons. *Mesopotamia*, *53*, 67–162.

Palermo, R. (2016). Filling the gap: The Upper Tigris region from the fall of Nineveh to the Sasanians. Historical and archaeological reconstruction through the data from The Land of Nineveh Archaeological Project. In Kopanias, K., & MacGinnis, J. (Eds.), *Archaeological research in the Kurdistan and adjacent regions*, BAR International Series: Archaeopress, pp. 266–276.

Parcak, S., Gathings, D., Childs, C., Mumford, G., & Cline, E. (2016). Satellite evidence of archaeological site looting in Egypt: 2002–2013. *Antiquity*, *90*(349), 188–205. https://www.cambridge.org/core/product/identifier/S0003598X16000016/type/journal_article

Parcak, S. H. (2009). *Satellite remote sensing for archaeology*: Routledge.

Pfälzner, P., Sconzo, P., Beutelscheiß, R., Edmonds, A., Glissmann, B., Herdt, S., Herrmann, J. T., Heydari-Guran, S., Kohler, J., Mueller-Weiner, M., Puljiz, I., & Sharp, M. (2016). The Eastern Habur archaeological survey in Iraqi Kurdistan. A preliminary report on the 2014 season. *Zeitschrift Für Orient-Archäologie*, *9*, 10–69.

Pirowski, T., Marciak, M., & Sobiech, M. (2021). Potentialities and limitations of research on VHRS data: Alexander the Great's military camp at Gaugamela on the Navkur Plain in Kurdish Iraq as a test case. *Remote Sensing*, *13*(5), 904. https://www.mdpi.com/2072-4292/13/5/904

Poidebard, R. P. A. (1934). *La Trace de Rome dans le Désert de Syrie: Le Limes de Trajan a la Conquéte Arabe, Recherches Aériennes*. Paul Geuthner.

QGIS. https://www.qgis.org/en/site/index.html

Rayne, L., Bradbury, J., Mattingly, D., Philip, G., Bewley, R., & Wilson, A. (2017). From above and on the ground: Geospatial methods for recording endangered archaeology in the Middle East and North Africa. *Geosciences*, *7*(4), 100. http://www.mdpi.com/2076-3263/7/4/100

Rayne, L., & Donoghue, D. (2018). A remote sensing approach for mapping the development of ancient water management in the Near East. *Remote Sensing*, *10*(12), 2042. http://www.mdpi.com/2072-4292/10/12/2042

Rayne, L., Gatto, M., Abdulaati, L., Al-Haddad, M., Sterry, M., Sheldrick, N., & Mattingly, D. (2020). Detecting change at archaeological sites in North Africa using open-source satellite imagery. *Remote Sensing*, *12*(22), 3694. https://www.mdpi.com/2072-4292/12/22/3694

Schuetter, J., Goel, P., McCorriston, J., Park, J., Senn, M., & Harrower, M. (2013). Autodetection of ancient Arabian tombs in high-resolution satellite imagery. *International Journal of Remote Sensing*, *34*(19), 6611–6635.

Sentinel 2. https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Silver, M., Törmä, M., Silver, K., Okkonen, J., & Nuñez, M. (2015). Remote sensing, landscape and archaeology tracing ancient tracks and roads between Palmyra and the Euphrates in Syria. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(5), 279–285. https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-5-W3/279/2015/

Simi, F. (2020). The Tell Gomel archaeological survey. Preliminary results of the 2015-2016 campaigns. In Coppini, C., & Simi, F. (Eds.), *Interactions and new directions in near eastern archaeology. Volume 3. Proceedings of the 5th "Broadening Horizons" Conference (Udine 5-8 June 2017)*: EUT Edizioni Università di Trieste, pp. 279–292.

Soroush, M., Mehrtash, A., Khazraee, E., & Ur, J. A. (2020). Deep learning in archaeological remote sensing: Automated qanat detection in the Kurdistan Region of Iraq. *Remote Sensing*, *12*(3), 500. https://www.mdpi.com/2072-4292/12/3/500

Stewart, C., Oren, E., & Cohen-Sasson, E. (2018). Satellite remote sensing analysis of the Qasrawet archaeological site in North Sinai. *Remote Sensing*, *10*(7), 1090. http://www.mdpi.com/2072-4292/10/7/1090

Tapete, D., & Cigna, F. (2018). Appraisal of opportunities and perspectives for the systematic condition assessment of heritage sites with Copernicus Sentinel-2 high-resolution multispectral imagery. *Remote Sensing*, *10*(4), 561. http://www.mdpi.com/2072-4292/10/4/561

Terra. https://terra.nasa.gov/

Titolo, A. (2021). Use of time-series NDWI to monitor emerging archaeological sites: Case studies from Iraqi artificial reservoirs. *Remote Sensing*, *13*(4), 786. https://www.mdpi.com/2072-4292/13/4/786

Toumazet, J.-P., Vautier, F., Roussel, E., & Dousteyssier, B. (2017). Automatic detection of complex archaeological grazing structures using airborne laser scanning data. *Journal of Archaeological Science: Reports*, *12*, 569–579. https://linkinghub.elsevier.com/retrieve/pii/S2352409X17300792

Traviglia, A., & Cottica, D. (2011). Remote sensing applications and archaeological research in the Northern Lagoon of Venice: The case of the lost settlement of Constanciacus. *Journal of Archaeological Science*, *38*(9), 2040–2050. https://linkinghub.elsevier.com/retrieve/pii/S0305440310003845

Traviglia, A., & Torsello, A. (2017). Landscape pattern detection in archaeological remote sensing. *Geosciences*, *7*(4), 128. http://www.mdpi.com/2076-3263/7/4/128

Trier, O. D., Larsen, S. O., & Solberg, R. (2009). Automatic detection of circular structures in high-resolution satellite images of agricultural land. *Archaeological Prospection*, *16*(1), 1–15.

Ur, J. (2003). CORONA satellite photography and ancient road networks: A Northern Mesopotamian case study. *Antiquity*, *77*(295), 102–115. https://www.cambridge.org/core/product/identifier/S0003598X00061391/type/journal_article

Ur, J. (2005). Sennacherib's northern Assyrian canals: New insights from satellite imagery and aerial photography. *Iraq*, *67*(1), 317–345. https://www.cambridge.org/core/product/identifier/S0021088900001418/type/journal_article

Ur, J. (2013a). CORONA satellite imagery and ancient near eastern landscapes. In Comer, D. C., & Harrower, M. J. (Eds.), *Mapping archaeological landscapes from space*, Springer Briefs in Archaeology, Vol. 5: Springer New York, pp. 21–31.

Ur, J. (2013b). Spying on the past: Declassified intelligence satellite photographs and near eastern landscapes. *Near Eastern Archaeology*, *76*(1), 28–36.

Ur, J., Babakr, N., Palermo, R., Creamer, P., Soroush, M., Ramand, S., & Nováček, K. (2021). The Erbil Plain Archaeological Survey: Preliminary results, 2012-2020. *Iraq*, *83*, 205–243. https://www.cambridge.org/core/product/identifier/S0021088921000024/type/journal_article

U.S. Geological Survey. https://www.usgs.gov

USGS Landsat 5. https://www.usgs.gov/landsat-missions/landsat-5

USGS Landsat 7. https://www.usgs.gov/landsat-missions/landsat-7

Ward Jr., J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244.

Wilkinson, K. N., Beck, A. R., & Philip, G. (2006). Satellite imagery as a resource in the prospection for archaeological sites in central Syria. *Geoarchaeology*, *21*(7), 735–750.

Wiseman, J. R., & El-Baz, F. (2007). *Remote sensing in archaeology*: Springer.

Zanni, S., & De Rosa, A. (2019). Remote sensing analyses on Sentinel-2 images: Looking for Roman roads in Srem Region (Serbia). *Geosciences*, *9*(1), 25. https://www.mdpi.com/2076-3263/9/1/25

How to cite this article: Valente, R., Maset, E., & Iamoni, M. (2022). Archaeological site identification from open access multispectral imagery: Cloud computing applications in Northern Kurdistan (Iraq). *Archaeological Prospection*, 1–17. https://doi.org/10.1002/arp.1874