Article

Three Years of Google Earth Engine-Based Archaeological Surveys in Iraqi Kurdistan: Results from the Ground

Riccardo Valente [1,*](ID), Eleonora Maset [2](ID) and Marco Iamoni [3](ID)

1   ISPC—Istituto di Scienze del Patrimonio Culturale, CNR, 20125 Milano, Italy
2   DPIA—Dipartimento Politecnico di Ingegneria e Architettura, Università degli Studi di Udine, 33100 Udine, Italy; eleonora.maset@uniud.it
3   DIUM—Dipartimento di Studi Umanistici e del Patrimonio Culturale, Università degli Studi di Udine, 33100 Udine, Italy; marco.iamoni@uniud.it
*   Correspondence: riccardo.valente@cnr.it

Abstract: This paper presents the results of a three-year survey (2021–2023), conducted in an area of approximately 356 km^2 in Iraqi Kurdistan with the aim of identifying previously undetected archaeological sites. Thanks to the development of a multi-temporal approach based on open multispectral satellite data, greater effectiveness was achieved for the recognition of archaeological sites when compared to the use of single archival or freely accessible satellite images, which are typically employed in archaeological research. In particular, the Google Earth Engine services allowed for the efficient utilization of cloud computing resources to handle hundreds of remote sensing images. Using different datasets, namely Landsat 5, Landsat 7 and Sentinel-2, several products were obtained by processing entire stacks of images acquired at different epochs, thus minimizing the adverse effects on site visibility caused by vegetation, crops and cloud coverage and permitting an effective visual inspection and site recognition. Furthermore, spectral signature analysis of every potential site complemented the method. The developed approach was tested on areas that belong to the Land of Nineveh Archaeological Project (LoNAP) and the Upper Greater Zab Archaeological Reconnaissance (UGZAR) project, which had been intensively surveyed in the recent past. This represented an additional challenge to the method, as the most visible and extensive sites (*tells*) had already been detected. Three years of direct ground-truthing in the field enabled assessment of the outcomes of the remote sensing-based analysis, discovering more than 60 previously undetected sites and confirming the utility of the method for archaeological research in the area of Northern Mesopotamia.

Keywords: archaeology; site detection; multispectral images; Google Earth Engine; Landsat 5; Landsat 7; Sentinel-2; spectral signatures; ground-truthing

1. Introduction

Remote sensing (RS) has been used for decades in the area of Northern Mesopotamia to support archaeological investigations, and its application is constantly increasing with the extensive availability of RS resources and tools. A review of the most recent literature in this field reveals some examples: the use of Very High-Resolution (VHR) imagery by Malinverni and colleagues [1,2] to study Assyrian archaeological features in the Navkur Plain; Altaweel and Squitieri [3] recorded the Dinka site, in Sulaymaniyah province, via an Unmanned Aerial Vehicle (UAV), assessing possible correlations between surface remains and buried structures; Kalayci et al. [4] tested the spectral response of ancient communication routes (locally called *hollow ways*) located in upper Mesopotamia; Soroush et al. [5] used deep learning algorithms to automatically identify on CORONA imagery ancient hydraulic structures (called *qanat*) in the Erbil area, whereas Starkova explored deserted villages in the surroundings of Erbil via historical RS data [6,7]; Pirowski et al. [8,9] focused their research on the identification of the Gaugamela battlefield in the Navkur Plain using Pleaides, WorldView-2 imagery and several processing techniques (Principal Component Analysis,

vegetation indexes, etc.); Laugier and colleagues [10,11] mapped archaeological features and assessed the damage that has occurred to archaeological sites in the Sirwan region using cartography, RS resources and geophysical surveys; and Titolo and colleagues [12,13] exploited Google Earth Engine© (GEE) to monitor sites covered by the waters of the Mosul Dam Reservoir and other dams.

Indeed, cloud computing services for RS products are already a valuable opportunity for scientific research. GEE can process on a large scale online RS resources with a relevant impact on any field dealing with geospatial analysis. For instance, GEE was recently used to assess the Ethiopian archaeological heritage [14], as well as the urban sprawl around the Amathus site in Cyprus [15], the city of Matera in Italy [16], and for the detection of looting activities on the Apamea site in Syria [17]. A recent and up-to-date overview of other GEE applications for archaeological purposes can be found in [18].

From an archaeological point of view, during the last two decades, a renewed interest has arisen for the archaeological heritage included within the autonomous region of Iraqi Kurdistan [19]. The Land of Nineveh Archaeological Project (LoNAP) has been carried out by the Università di Udine since 2012, identifying some 1300 archaeological features, within an area that approximately covers 3000 km^2 eastward of the artificial lake created by the Mosul Dam [20–27]. Within the LoNAP area, more precisely, close to its south-eastern border, the Asingeran Excavation Project (AEP), started in 2019 and still ongoing, is investigating the site of Tell Asingeran (see references in Section 2.1). Other contiguous archaeological projects are the Erbil Plain Archaeological Survey (EPAS) [28], the Upper Greater Zab Archaeological Reconnaissance (UGZAR) [29–31] and the Eastern Habur Archaeological Survey [32]. All of the previously cited missions used RS data, a long-standing routine for archaeologists involved in the Near East. The first groundbreaking experiences go back to the 1930s [33] with aerial images of the main sites. Archival and declassified CORONA images [34] gave an invaluable advantage for landscape archaeology in Syria [35,36], southern Mesopotamia [37] and the Levant in general [38]. The declassification of U2 imagery had a similar and even greater impact, due to their better spatial resolution [39]. Most of the recent archaeological studies in this area that relied on RS focused on the observation of CORONA images to identify Roman road networks between Israel and Jordan [40] and the exploitation of short-wave infrared (SWIR) imagery from high-resolution platforms in Turkey, Syria and Iraq [41].

This work presents the results of a three-year survey (2021–2023) aimed at identifying new and previously unknown archaeological sites within the LoNAP and UGZAR areas in Iraqi Kurdistan, leveraging open-access RS products and the web-based cloud computing services provided by GEE. Carefully planned ground-truthing inspections made it possible to directly verify in the field the preliminary outcomes of the RS analyses. In this regard, it should be stressed that not all the RS archaeological applications encompass an actual on-site check. This is often due to logistical or economic constraints, which can limit the overall effectiveness of the approaches proposed in the literature [11,42]. An overview of the method and the partial results of the 2021 and 2022 on-field surveys have already been described in [43,44]; this work is intended as a presentation of the 2023 results, as well as a complete and exhaustive review of the collected and corrected data over three years. It also includes a more in-depth critical discussion of the overall results.

The paper is structured as follows. Section 2 outlines the methodology employed in the study, including a description of the surveyed areas, the datasets used and details of the ground-truthing phase. Section 3 shows the obtained results in terms of the number of new detected sites, their distribution and typologies, while Section 4 discusses the overall outcomes and the reliability of the a priori assessment of possible archaeological sites based on the spectral signature analysis. Section 5 wraps up the obtained results and frames them within archaeological practice.

2. Materials and Methods

2.1. Surveyed Areas

The areas under investigation are located in the Plain of Navkur and the Baadreh Plain in Iraqi Kurdistan (Figure 1). Both regions are clearly defined along their northern edges by the lower slopes of the Zagros Mountains, though only the Navkur Plain has distinct mountainous boundaries to the east and south, thanks to the presence of the Jebel Bardarash and Jebel Maqloub ranges. These plains were particularly suited to ancient human settlements due to the fertile brown soils (especially prominent in the Navkur area) and the presence of several watercourses. While many of these watercourses are now seasonal, with the exception of the two main rivers, the Nahr al-Khazir and the Gomel Su, they likely provided more permanent water sources in the past, offering a significant supply of water [45]. Rainfall also plays a crucial role in supporting agriculture, enabling dry farming practices. These favorable conditions attracted the formation of the first stable farming communities as early as the Neolithic period (c. 10,000 BCE) [46]. It is also possible that the region's abundant water supply, especially in the Navkur Plain (whose name in Kurdish Badînî means "mud plain"), may have negatively impacted the visibility of archaeological sites. As observed at the site of Asingeran, flooding and consequent redeposition of river sediments could have concealed evidence of ancient settlements [47], especially those lacking the typical mounded morphology of the so-called Mesopotamian *Tells*. In this context, the development and application of new techniques could greatly enhance the ability to investigate and better understand the ancient landscape.
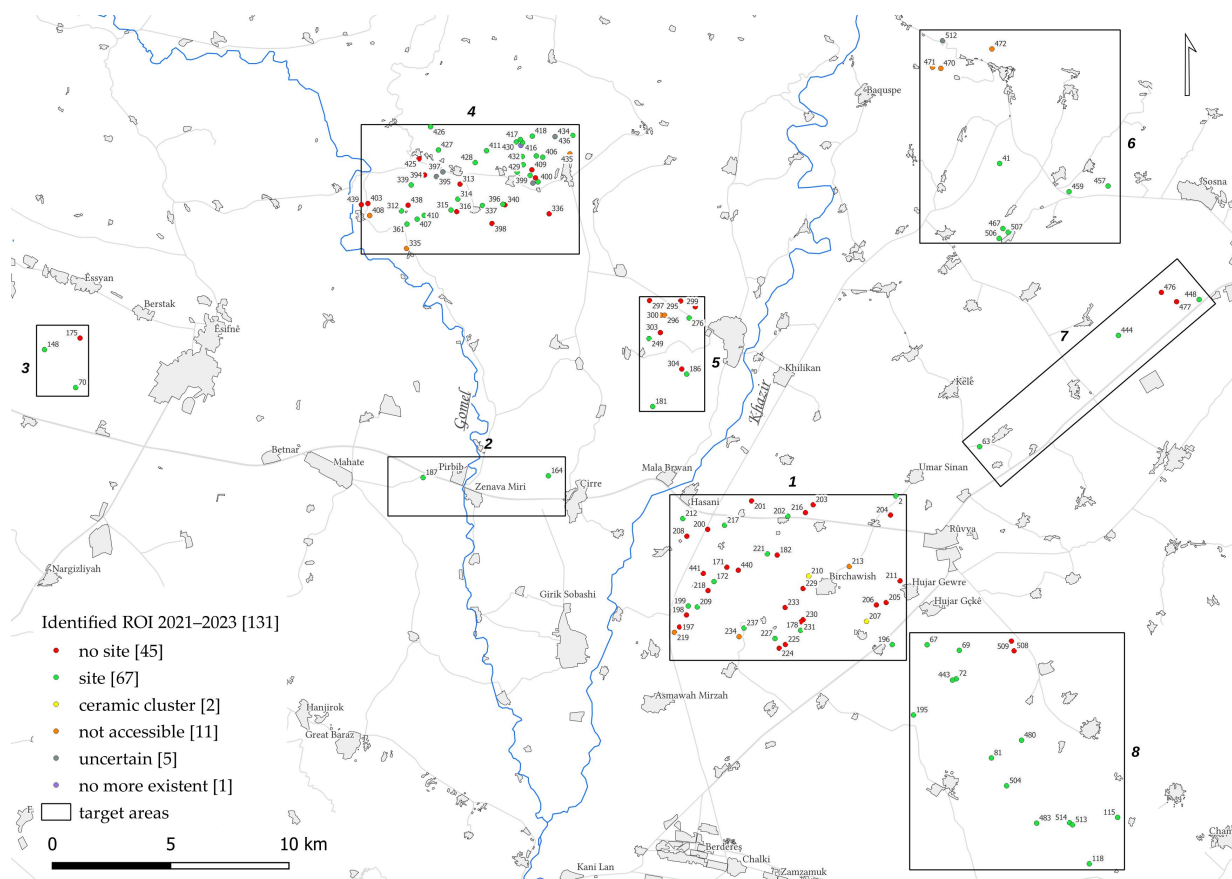

Figure 1. The Plain of Navkur and the Baadreh Plain in Iraqi Kurdistan. The black rectangles delineate the eight areas under investigation, while the dots indicate the regions of interest (ROIs) identified by the RS-based analysis presented in the paper. The ROIs are colored according to the results of the ground-truthing phase, whereas the number associated is an ID given during the RS-based inspection and is not related to the numbering of previous archaeological missions. Digits between square brackets are the feature count.

Within this natural environment, eight areas (shown in Figure 1) were identified, in order to search for possible new and previously undetected archaeological sites. Area 1 (about 70 km^2) is close to the site of Asingeran; Area 2 (about 19 km^2) is close to the Gomel river; Area 3 (about 7 km^2) is in a plain context between Baadhrah and Ash Shaykhan; Area 4 (about 50 km^2) covers the foothills south of Atrush valley; Area 5 (about 13 km^2) is close to Qasrok village; Area 6 (about 76 km^2) is the hilly area near the modern village of Gorank Kamil; Area 7 (about 30 km^2) is the undulated plain extending southwest of the modern city of Akre; and Area 8 (about 91 km^2) is the eastern sector of the Plain of Navkur, between the modern city of Rovia and the great Zab. Areas 1–3 were surveyed in 2021, Areas 4 and 5 in 2022, and Areas 6–8 in 2023. The total covered area is about 356 km^2. While Areas 1–5 are situated within the broader LoNAP area, Areas 6–8 are located within the UGZAR boundaries. The criteria used to select the eight different areas are basically grounded on the location of previously known sites and the different type of natural environment (lowland, piedmont, rocky hills).

2.2. Selected Datasets

This research exploited open-access RS resources, namely medium- (MR, Table 1) and high-resolution (HR, Table 2) satellite imagery. As for the latter, only HR images provided by the Google Earth service and accessed via Google Earth Pro were used. MR images were employed as pre-processed datasets available on GEE . The main satellite platforms selected for the analyses were: (i) Landsat 5—LANDSAT/LT05/C02/T1_L2 (L5), (ii) Landsat 7—LANDSAT/LE07/C01/T1_TOA (L7) and (iii) Sentinel-2—COPERNICUS/S2_SR (S-2). The L5 (1984–2013) and L7 (1999–2024) platforms are part of the Landsat program that has been delivering satellite imagery since 1972 and have similar characteristics. In more detail, the 7 bands of the L5 data encompass a wavelength range of 0.45 μm to 2.35 μm and have a spatial resolution of 30 m for the reflective bands and 120 m for the thermal band [48]. The same wavelength range characterizes the 8 bands of the L7 platform, which has a panchromatic spatial resolution of 15 m, 30 m for the reflective bands and 60 m for the thermal band [49]. S-2 is a constellation by the European Space Agency composed of two identical satellites (Sentinel-2A and -2B), active since 2015, that delivers 13 bands in the range ∼0.443 μm–∼2.19 μm at a spatial resolution from 10 m to 60 m [50]. The MODIS Combined 16-Day Normalized Difference Water Index (NDWI) and Normalized Difference Vegetation Index (NDVI) datasets were additionally used in this work for the seasonal analysis described in the next section.

Table 1. Medium-resolution multispectral image datasets used in this work *.

	MODIS **MCD43A4_006_NDWI** **MCD43A4_006_NDVI**	**LANDSAT** **LT05/C02/T1_L2**	**LANDSAT** **LE07/C01/T1_TOA**	**COPERNICUS** **S2_SR**
Time series	2012–2021	1984–1994	2000–2010	2018–2021
Spatial resolution	463.3 m	30 m	15 m–30 m	10 m–20 m–60 m
Radiometric range	–	0.45–2.35 μm	0.45–2.35 μm	∼0.443 μm–∼2.19 μm
Processed images	3652	157	176	31
Bands used	NDWI/NDVI	B1–B5, B7	B1–B5, B7–B8	B1–B9, B11–B12
Quality assessment band	–	QA_PIXEL	QA_PIXEL	QA60

* As of 2024, GEE introduced some changes in dataset availability: the Landsat datasets were updated to Collection 2 and are now LANDSAT/LT05/C02/T1_L2 and LANDSAT/LE07/C02/T1_L2; S-2 dataset is currently deprecated and superseded by COPERNICUS/S2_SR_HARMONIZED.

Although commercial satellite imagery can achieve higher spatial resolution and assist in identifying smaller archaeological features, this study employed only open imagery to maintain a cost-effective approach and to leverage a multi-temporal method made possible by cloud computing.

Table 2. Additional images (high-resolution and archival datasets) employed for visual inspection.

	Google Earth HR Images	**CORONA**	**U2**
Acquisition period	2010–2023	mission n. 1039 22/02–11/03/1967	mission n. 1554 29/01/1960
Spatial resolution	<10 m	2.75 m	variable *
Spectral bands	RGB	panchromatic	panchromatic

* Estimated in 0.4 m for the vertical frames ([39], p. 116).

2.3. Developed Approach

The following sections provide a comprehensive description of the multi-temporal approach developed for the processing of the aforementioned datasets. The workflow led to the identification of regions of interest (ROIs), which could potentially represent new archaeological sites. These were subsequently verified during the ground-truthing phase. An overview of the proposed approach is presented in the flowchart shown in Figure 2.
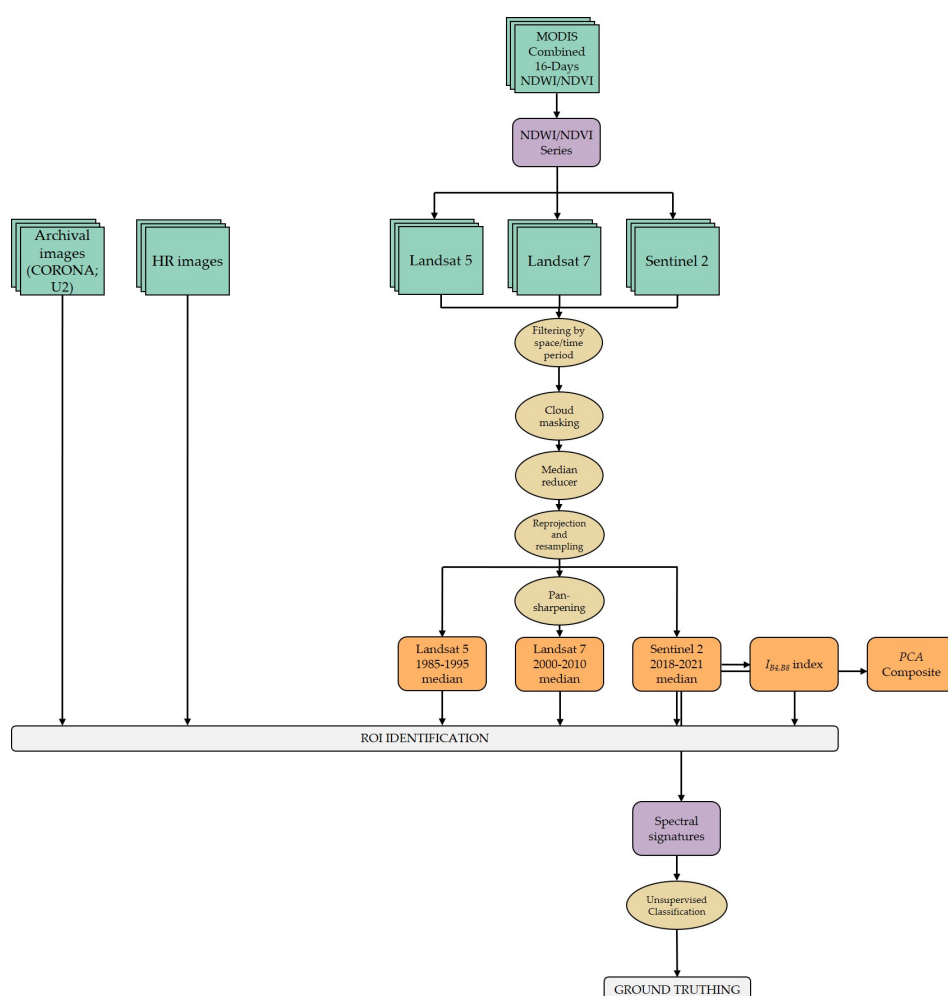

Figure 2. Flowchart of the approach presented in the paper. The datasets employed are indicated in green, while the derived information is shown in purple. The color tan is used to represent filters and processing operations, with the corresponding results depicted in orange. Finally, the manual processes are presented in grey. Rectangles are employed for the input data, while smoothed rectangles are used for the results of processing steps. The functions adopted are shown in ellipses. The MR multispectral images are subjected to a series of processing stages to enhance the visibility of soil anomalies and facilitate site identification. In contrast, HR and archival images are employed to examine the ROIs and verify the possible presence of elements on the ground that could link the anomalies to recent human activities.

2.3.1. Preliminary Assessment and Seasonal Analysis

Archaeological sites in the Navkur Plain typically correspond to soil anomalies, which commonly appear in RS images as rounded or oval areas with a lighter hue than the surrounding terrain. This general aspect and morphology was already known thanks to the years of surveys conducted by LoNAP and the other archaeological missions. Nevertheless, a preliminary assessment of the sites that had been previously detected over the years was carried out on multiple RS sources. The first choice was to use single S-2 imagery, accessed through the Sentinel Hub EO Browser portal [51], with the objective of verifying both the spectral bands and the periods of maximum visibility of the known sites. Site visibility was favorable in the B2 (blue), B3 (green), RedEdge and SWIR bands, depending on soil characteristics. However, the greatest overall visibility was noticed in the B4 (red) and B8 (visible and near infrared, VNIR) bands. In addition, the B2, B3, B4 and B8 bands have the best spatial resolution among all the bands of the three selected platforms (L5, L7, S-2), which is a further advantage for site detection.

For what concerns the seasonal visibility, the 'wet' season appeared to be more favorable for site identification. Considering this aspect, the already cited MODIS Combined 16-Day NDWI dataset, which is directly accessible within GEE, was used to identify the months in which soil moisture is at its highest on the region under investigation. The selected time period for this analysis was 2012–2021 across the eight investigated areas; the graph in Figure 3 clearly shows that the soil moisture level increases from December to April, a trend that is observed cyclically over the years. The climate in the Zagros foothills is classified as Mediterranean, with precipitation reaching up to 1000 mm, while moving south, the plains experience lower rainfall, with an annual average of 643 mm recorded at Bardarash, located along the southern border of Navkur. Temperature also varies considerably, partly due to the region's unique morphology [52]. The proximity to the Zagros Mountains contributes to a cold winter season, with an average temperature of 6.9 °C, whereas summers are typically very hot, with July averaging 32.8 °C [53]. These climate variations throughout the year lead to significant landscape changes, which notably affect satellite imagery of the area. Furthermore, while reviewing the data after the completion of the three-year survey, a cross-check using a different index was performed. In particular, the MODIS Combined 16-Day NDVI dataset was used, considering the same previous time and space constraints. The joint examination of the NDWI and the NDVI confirmed what had already been observed, with a notable correlation between the two trends as evidenced by the monthly values in Figure 3.
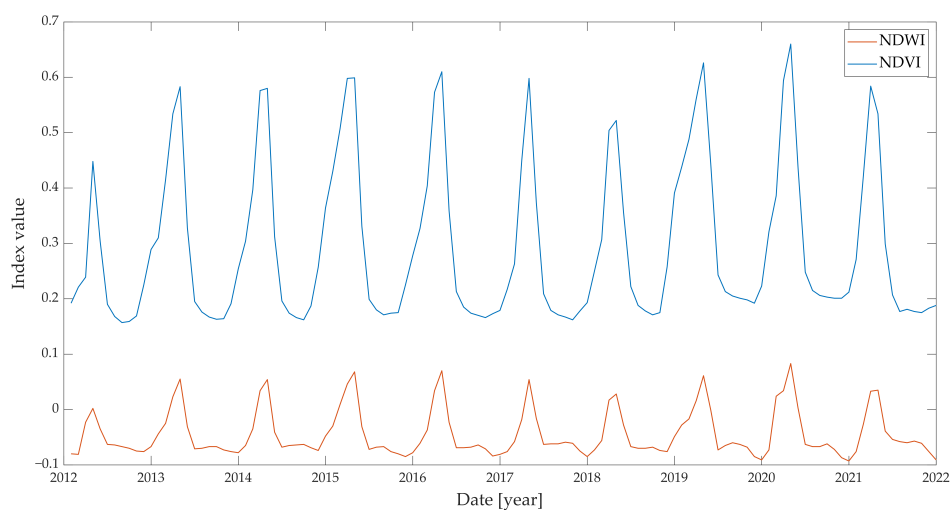

Figure 3. NDWI and NDVI trends over the period 2012–2021 in the eight target areas.

The visual inspection of the satellite data also revealed the limitations of relying on a single image: due to seasonal crops, agricultural works and weather conditions, site visibility can greatly change (Figure 4), and it is almost impossible to simultaneously observe all sites within a given area.
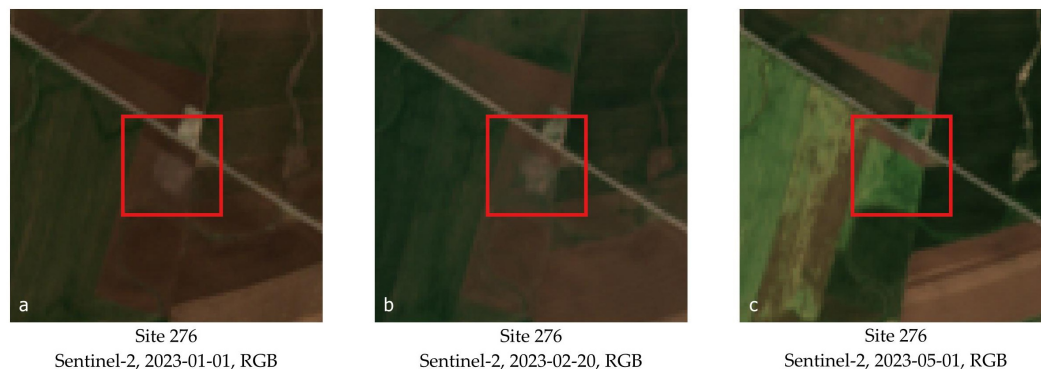

| Site 276 | Site 276 | Site 276 |
| Sentinel-2, 2023-01-01, RGB | Sentinel-2, 2023-02-20, RGB | Sentinel-2, 2023-05-01, RGB |

Figure 4. Soil anomaly corresponding to site 276, framed by the red square and located in Area 5, shown in S-2 images acquired at different epochs. The visibility of the site changes considerably depending on the season and the presence of cultivation (**a–c**).

Moreover, the preliminary analysis conducted on previously detected sites highlighted that site visibility is also influenced by the natural environment. Sites located in plain contexts are typically well identifiable, whereas those located in piedmont and hilly environments are much less visible. In these latter contexts, the multispectral images provide limited value, potentially due to the absence of anthropogenic soil on the rocky substratum. For this reason, HR satellite images distributed via Google Earth were also analyzed together with the MR multispectral data, especially for what concerns Areas 4 and 6. Their use in these areas is even more well founded considering that the presence of surface remains, namely stone and cobble structures, is higher in the piedmont and hilly contexts, where the impact of agriculture is presumably less significant. In addition, HR images were useful in plain contexts, in those cases (probably deserted villages) where structures are still partially preserved. While MR images from S-2 (10 m spatial resolution) are suitable for discerning soil anomalies, they lack the requisite detail to identify structures.

The preliminary assessment presented in this section demonstrated the inherent constraints associated with the observation of single satellite images. This prompted us to develop the multi-temporal approach described in the following. Furthermore, the seasonal analysis was of paramount importance in identifying the optimal period of the year for site visibility, thereby enabling the selection of the most suitable image datasets to be processed. Visibility, i.e., the environment variability that allows an observer to identify any archaeological feature, is always a critical aspect in archaeological surveys that deeply influences the methods and strategies [54,55]. Archaeological reconnaissance procedures require revisiting of the detected sites on multiple occasions, across different seasons and years, due to the significant fluctuations in ground visibility that occur over time. An RS-based multi-temporal approach on open ground allows some of the most common constraints to be reduced, increasing site visibility.

2.3.2. Image Processing

The GEE platform was chosen to efficiently work with stacks of images acquired at different epochs and implement a multi-temporal approach for site identification. As the areas to be investigated were progressively added each year, image processing was carried out separately for each year prior to the on-field campaign. However, the selected algorithms and parameters were kept constant throughout the 2021–2023 survey period, in order to have consistent results to compare.

As mentioned in Section 2.2, the used GEE datasets were the 'USGS Landsat 5 Level 2, Collection 2, Tier 1', the 'Sentinel-2 MSI: MultiSpectral Instrument, Level-2A'— both with atmospherically corrected reflectance values—and the 'USGS Landsat 7 Collection 1 Tier 1 TOA Reflectance'—with calibrated top-of-atmosphere values. The satellite imagery datasets were limited to a specific date range depending on the platform launch date and period of activity, namely 1984–1994 for L5, 2000–2010 for L7 and 2018–2021 for S-2. Each dataset was then cropped to the eight investigated areas from a spatial point of view and to the January–April temporal range, as this period was identified as the most favorable for site visibility according to the NDWI assessment described in Section 2.3.1 and for the limited presence of crops. The three filtered datasets were ultimately composed of 157 (L5), 176 (L7) and 31 (S-2) multispectral images, respectively. Cloud masking was applied before further processing.

For each filtered dataset, the time series of images was reduced to a single multi-band image by applying a median function separately to each spectral band, thereby obtaining a single robust value for each pixel: the final output is therefore a multi-band image that partially removes the seasonal disturbance effects, increasing the chance of site identification. For what concerns the S-2 dataset, an additional image was generated by applying a developed simple ratio (index $I_{B4,B8} = B4/(B4 - B8)$) in order to better highlight the visibility of soil anomalies, leveraging those 10 m bands that were shown to be suitable for this purpose (see, e.g., Figure 5f). All final images were reprojected from EPSG:4326 to EPSG:3395 and resampled on the same grid using the bicubic interpolation method. On the L7 dataset, a pan-sharpening with the panchromatic band (B8) was also performed to increase the spatial resolution. Figure 5 illustrates some of the outputs generated by the described workflow, showcasing the emergence of a previously undetected archaeological site.
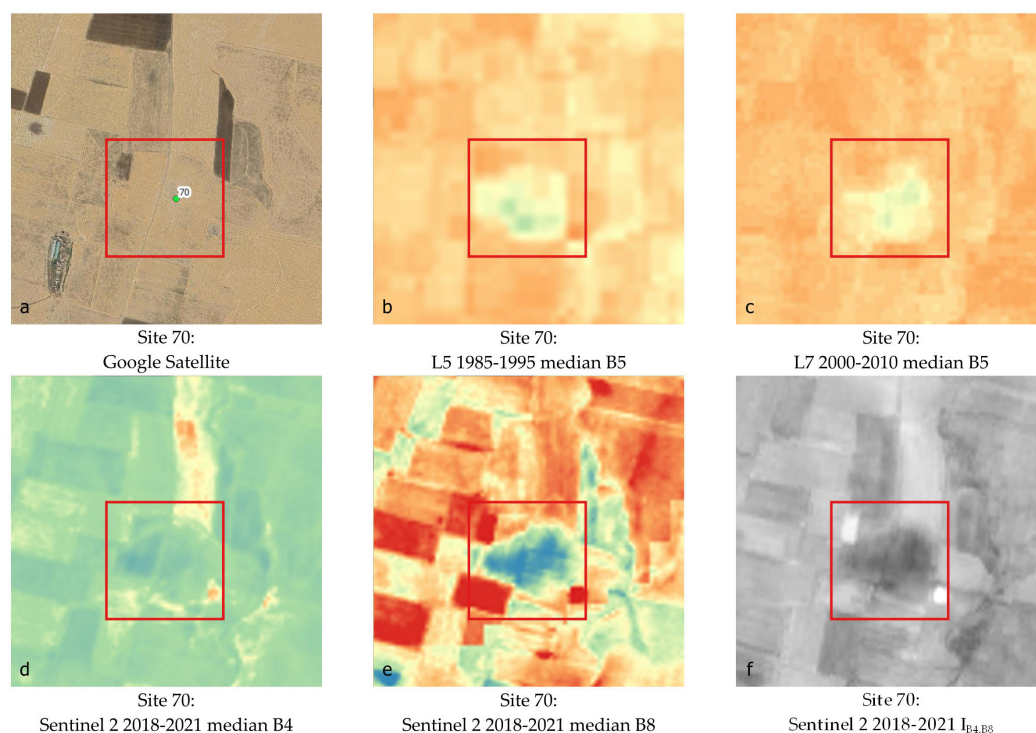

Figure 5. Site 70, framed by the red square, as it appears on different RS products used and generated for this study (**a–f**).

As a further product, Principal Component Analysis (PCA) was investigated, as it is one of the most common methods applied in RS applications to archaeology. PCA is a dimensionality reduction technique that, when applied to images, compresses data from a large number of bands into a smaller number of uncorrelated bands. PCA is based on the eigenvector decomposition of the data covariance matrix: performing band decorrelation and keeping the first three principal components, images with brighter and contrasting colors are produced, which should facilitate the easier detection of independent features [56]. The PCA was applied through GEE to the S-2 dataset restricted to a temporal range between 2018 and 2020, after computing the median of the time series. Although the four final outputs—three single-band images, one for each of the three principal components, and a composite image (Figure 6)—were correctly generated, the visibility of the archaeological sites was scarcely enhanced and these outputs resulted in being much less useful than those previously presented.
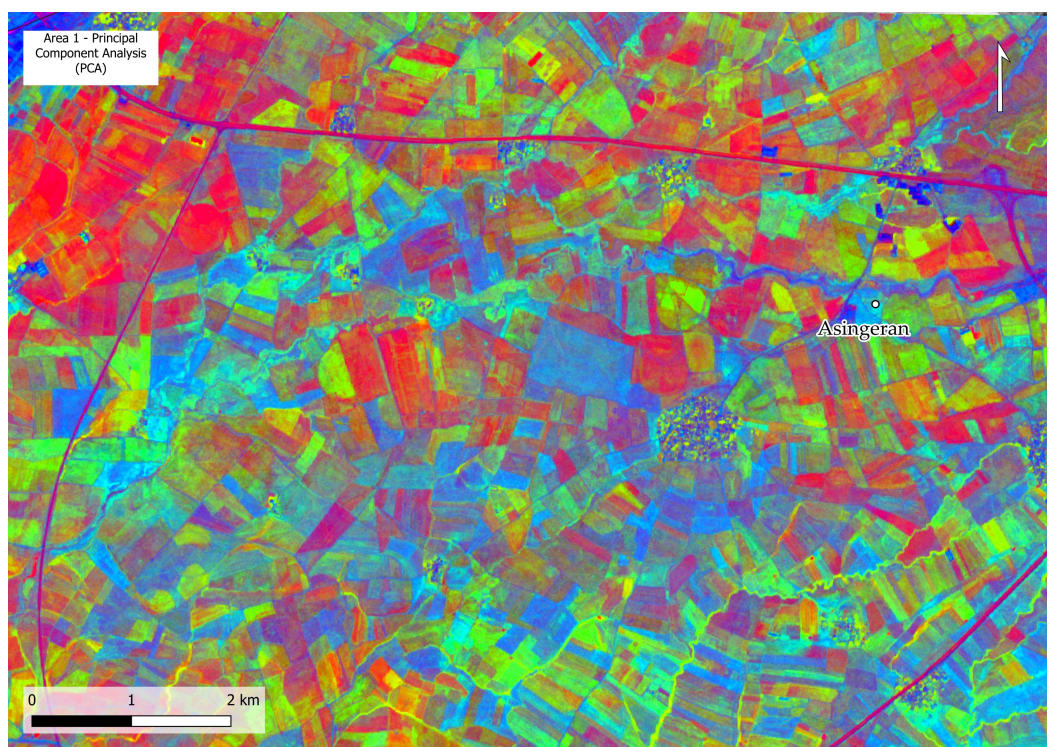

Figure 6. Transformed image of Area 1 resulting from the PCA analysis conducted on the 10 m spatial resolution bands of the S-2 dataset. The figure shows a false-color image, where the original red, green and blue bands have been replaced by the three principal components (i.e., uncorrelated bands) computed by the PCA. Asingeran site is labeled.

The inspection of the obtained outputs was visually carried out using the open-source software QGIS (https://www.qgis.org/, accessed on 25 September 2024), where the location of known sites was superimposed onto the generated images to check their appearance. It was evident that, although a relevant number of known sites were visible on L5 and L7 images, the S-2 data showed a greater number of anomalies in detail, due to their superior spatial resolution. That was the main reason why the obtained S-2 products were selected as the principal source for the detection of new sites and the additional spectral signature analysis described in Section 2.3.3. The visual inspection led to the identification of soil anomalies, i.e., ROIs that could correspond to possible new sites. Finally, the use of HR images allowed further investigation of these soil anomalies and verification of the possible presence of elements on the ground that could link the anomalies to recent human activities. In certain instances, HR images showed clear signs of earthworks or temporary settlements related to seasonal farming that helped to not consider these anomalies as

potential archaeological sites. Along with contemporary HR images, archival panchromatic CORONA and U2 images (Table 2) were also inspected: while in some instances they did not add any valuable information, in other cases their analysis was crucial because they recorded features that are currently less visible, such as some deserted villages that were still inhabited in the mid-20th century, or that no longer exist. However, when considering the whole set of ROIs, only approximately 15% were also clearly identifiable on CORONA and 25% on U2 images. The location of the ROI retrieved from the images was recorded as point features, for subsequent on-field verification.

2.3.3. Spectral Signature Extraction and Classification

In order to have additional information about the nature of the identified ROI prior to the on-field surveys, the corresponding spectral signatures were extracted via GEE from the S-2 median output, sampling an area of 50 m × 50 m centered in the middle of each possible new site. The size of the sampling area was selected based on the dimensions of the soil anomalies observed. The spectral signature of each ROI is therefore a vector of 12 values, one for each spectral band considered. As for the spectral signature representing already known sites, this was computed as the median of all the spectral signatures available from the sites discovered by the LoNAP and UGZAR surveys. A representative spectral signature of cultivated fields (i.e., non-anthropogenic soils) was calculated as well, sampling and averaging the pixel values on areas of size 50 m × 50 m. Subsequently, an unsupervised classification process was conducted on the spectral signatures with the objective of differentiating between the ROIs that are most likely to correspond to new sites and those that are, in fact, soil anomalies not attributable to the existence of authentic archaeological sites. Common clustering algorithms available in the Python *scikit-learn* library (https://scikit-learn.org/stable/, accessed on 25 September 2024) were applied: (i) *fuzzy c-means* [57], (ii) *hierarchical clustering* with the minimum-variance linkage criterion [58], (iii) *k-means* [59] and (iv) *self-organizing map* (SOM) [60]. All the methods are well known in the RS field for image or laser scanning data unsupervised classification (e.g., [61–65]) and require as input the desired number of clusters (in our case, two). To assign the correct label to the obtained clusters, we leveraged the median spectral signature of the already known LoNAP and UGZAR sites. Indeed, the cluster whose representative element most closely resembled (in terms of Euclidean distance) this median signature was designated as 'site', and the corresponding ROIs were labeled accordingly. Conversely, the ROIs whose signatures were associated with the other cluster were marked as 'no site'. It should be noted that the analysis was carried out separately for each year.

2.4. Ground Truthing

A dedicated on-field survey was planned on an annual basis in order to verify the identified ROI and the results of the spectral signature classification through direct observation on the ground. Surveys were carried out by a minimum of three to a maximum of five persons over a period of six days in 2021 and 2022, and three days in 2023. Despite most of the ROIs being situated in the open country without close relevant reference points, the locations were quite easily reached thanks to spatial data made available on mobile devices with Global Navigation Satellite System (GNSS) receivers. Each ROI was directly inspected in search of any anthropic elements (potsherds, structures), which might indicate the presence of an ancient site. Surface findings, where present, were sampled and collected to hypothesize their preliminary chronology. Moreover, the survey conducted on the ground allowed assessment of the typology of the new detected sites (some examples are shown in Figure 7).
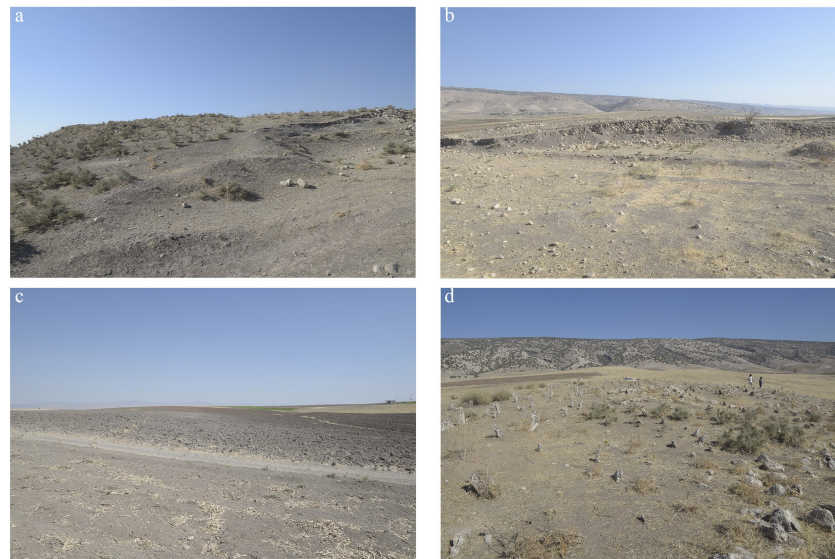
Figure 7. Views from the field: (**a**) site 312 (high mounded); (**b**) site 315 (low mounded); (**c**) site 181 (flat); (**d**) site 433 (graveyard).

3. Results

Thanks to the proposed approach, an overall number of 131 ROIs were identified on the satellite data and subsequently checked on the ground: 42 in 2021, 59 in 2022 and 30 in 2023. Their distribution within the eight investigated areas, visible in Figure 1, is variable: 39 in Area 1, 2 in Area 2, 3 in Area 3, 45 in Area 4, 12 in Area 5, 10 in Area 6, 5 in Area 7 and 15 in Area 8.

Among the potential new sites identified, 11 ROIs situated in Areas 1, 4 and 6 were not accessible for direct surveying due to a number of factors, including the presence of crops or other limitations. From the on-site assessment, 67 ROIs yielded various anthropic features (surface findings, structures or graveyards) and were therefore classified as 'sites', while two ROIs were labeled as 'ceramic clusters' due to the the lower amount of ancient potsherds found on the field surface. Five ROIs were considered as 'doubts', as the evidences collected on the field were not sufficient to determine an effective anthropic presence. Furthermore, one ROI corresponded to an anthropic feature visible on the HR images but no longer existent. The remaining 45 ROIs gave a negative outcome and were thus designated as 'no site'. The distribution of the sites among the target areas is shown in Figure 8. The highest density of new detected sites is observed in Area 4, with a density of about 0.5 sites per square kilometer. In comparison, the remaining areas exhibit a density of 0.3 sites per square kilometer or lower. Additionally, Area 4 demonstrates the greatest diversity in site types, whereas Area 3 is characterized by the presence of exclusively flat sites.

In terms of the typology of the newly identified sites, 19 can be labeled as 'flat sites', i.e., sites with no relevant elevation on the ground level. These sites are usually the most difficult archaeological evidence to be detected if compared to low- (14) or high-mounded (6) sites (these last were only identified during the 2022 and 2023 campaigns). Graveyards (9) constitute another relatively common typology. While they are not necessarily archaeologically relevant, given that they often belong to the modern era, they were nevertheless considered 'sites', as their traces were generally well visible on multispectral images. Other anthropic features include stone villages (9), enclosures (5) (a sixth enclosure appeared to no longer exist), settlements on rocky hills (4) and an unidentified structure (1). Although they were not properly considered as sites, it is worth mentioning the additional anthropic features found that include ceramic clusters (2), a possible tumulus (1), a possible stone village (1) and an uncertain structure (1). Most of the identified sites (55) had surface findings (potsherds and, to a lesser degree, stone or bricks fragments).
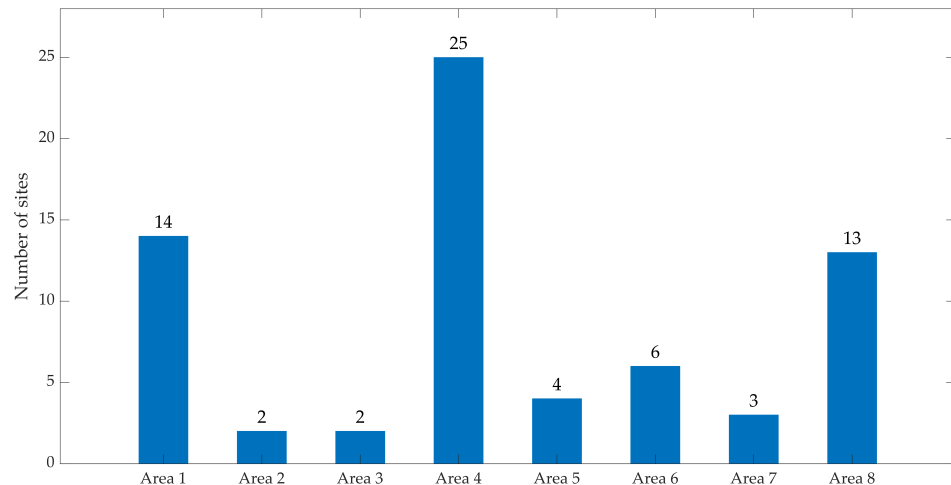
Figure 8. The number of ROIs that were assessed as actual archaeological sites or ceramic clusters as a result of the ground-truthing activity.

In regard to the clustering analysis, the spectral signatures of 118 out of 131 ROIs were processed. The remaining 13 ROIs were excluded from the analysis because they fall within the area where rocky hills take the place of the fertile plain, with a substantial change in the nature of the soil. Being the spectral signature classification mostly based on changes in the nature of fertile soil, those ROIs located in places where rock outcrops prevail were excluded from the analysis. Due to these reasons, all the previously mentioned 13 ROIs were identified on HR images, because of the presence of visible structure remains, and not using multispectral images.

Table 3 illustrates the results of the unsupervised classification in terms of overall accuracy, namely the percentage of correctly classified ROIs. It should also be noted that the results were calculated solely on the 103 out of 118 ROIs that were identified in the field as 'site' or 'no site': inaccessible ROIs and ROIs designated as 'ceramic cluster', 'doubt' and 'no more existent' were excluded from the accuracy assessment, as they did not yield definitive results.

Table 3. Overall accuracy achieved by the unsupervised classification algorithms, calculated as the percentage of ROIs correctly classified as 'site' or 'no site' out of the total number of assessed ROIs.

	Fuzzy c-Means	*Hierarchical*	*k-Means*	**SOM**	*Combined Pred.*
2021	73%	59%	76%	81%	76%
2022	60%	63%	65%	60%	60%
2023	62%	54%	62%	50%	54%
Aggregated	65%	59%	68%	65%	64%

The aggregated results over the three years show that *k-means* provided the best performance: 43 ROIs were correctly classified as 'site' (i.e., true positive) and 27 ROIs as 'no site' (i.e., true negative), leading to an overall accuracy of 68%, whereas 18 ROIs were false positive ('no site' that the algorithm labeled as 'site') and 15 false negative (actual 'site' that the algorithm misclassified as 'no site'). Conversely, the *hierarchical clustering* algorithm gave the lowest overall accuracy (59%, with 34 true positive and 27 true negative ROIs). Nevertheless, the *k-means* method was not consistently the most accurate in all cases (e.g., SOM provided better results in the 2021 analysis). Overall, the algorithms demonstrated comparable behavior, making it challenging to definitively identify the most optimal one. For this reason, we took into account the outcomes of all methods by providing a *combined prediction*: this entailed classifying an ROI as 'site' when at least three out of the four algorithms designated it as such. Figure 9 shows the detailed results, in terms of true positives, true negatives, false positives and false negatives, provided by the *combined prediction* method for each year. Moreover, the alluvial diagram presented in Figure 10

visualizes and summarizes the outcomes of both the spectral signature classification and the ground-truth assessment.
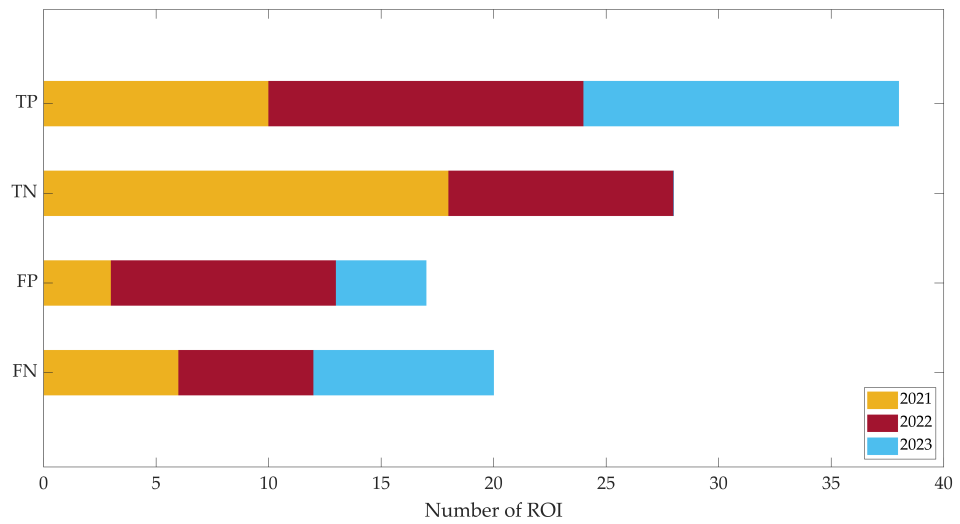

Figure 9. Cumulative results (2021–2023) of the spectral signature classification process. True positive (TP), true negative (TN), false positive (FP) and false negative (FN) sites are reported.
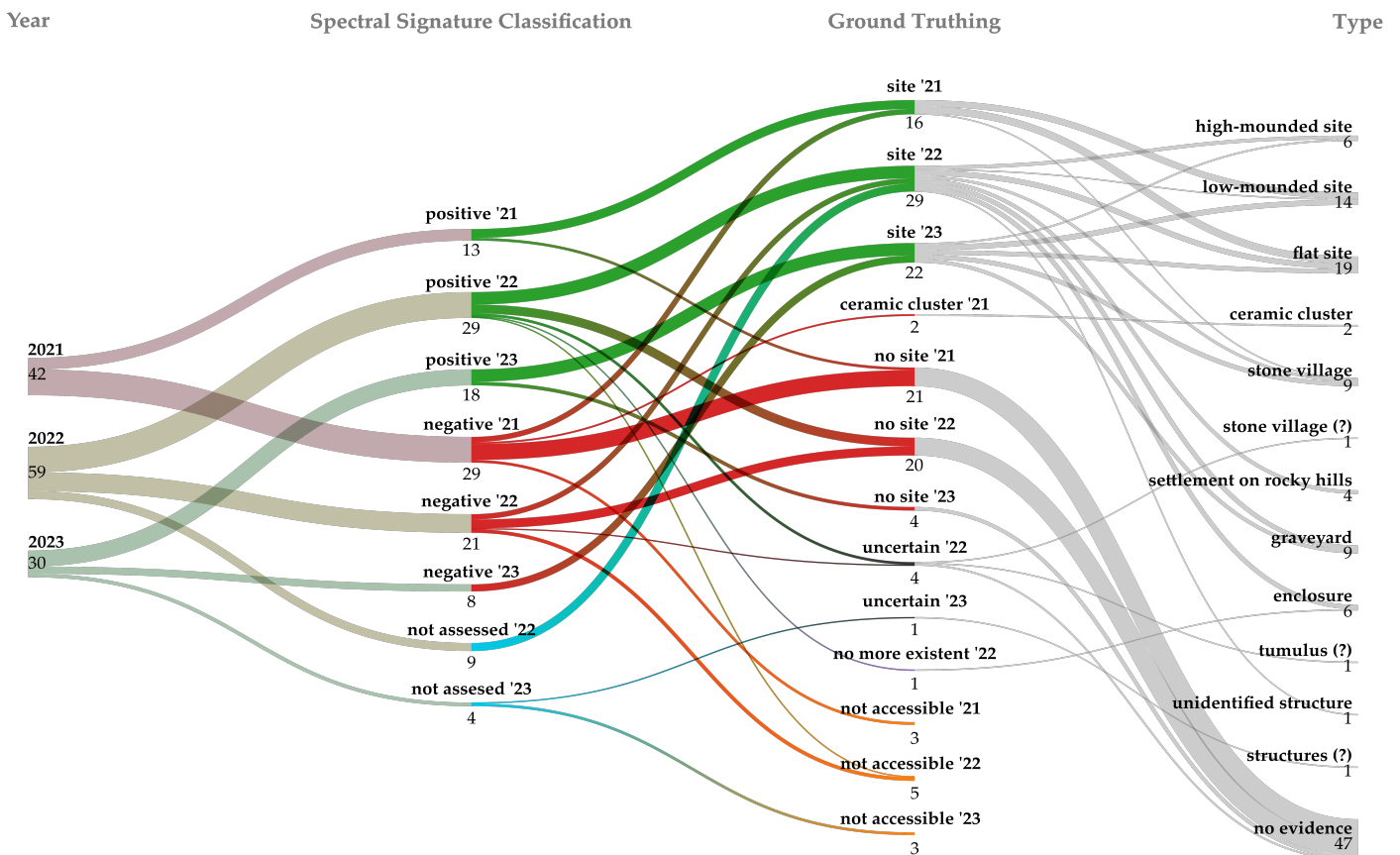

Figure 10. Alluvial diagram that tracks the outcomes of the survey over the years and the differences between the desk-based classification of spectral signatures and the ground-truthing results (digits not in bold indicate the number of features). First column shows the surveyed ROIs per year; second column reports the outcomes of the automatic spectral signature classification; third column shows the validation of the results, obtained after the ground-truthing phase; fourth column presents the sites actually identified in the field, grouped according to the evidence found.

A combined analysis of the spectral signatures and the typology of the newly detected sites can further enrich the results. To this end, we computed the median spectral signature of the sites aggregated according to the type of anthropic features found. Subsequently, the outcomes were compared to the median of the spectral signatures of the already known sites and of the non-anthropogenic soils, taken as a reference throughout the duration of the survey. This approach has at least two constraints: firstly, the overall limited number of sites for which spectral signature analysis was performed; and secondly, the criteria used to assign the typology to each site, that are largely subjective. For instance, the distinction between a 'high-mounded' and a 'low-mounded' site is not grounded on topographic elevation measurements, but rather on the experience of archaeologists involved in field surveys. To the same extent, flat sites that showed clear traces of pebble structures were labeled as 'stone villages', although similar traces can sometimes also be found on high-mounded sites. Despite the mentioned constraints, the results show some clear trends. The median spectral signatures of sites labeled as 'high-mounded' and 'stone villages' almost totally overlap (see Figure 11). The same happens for sites labeled as 'low-mounded' and 'flat'; the latter also overlap with the median of the known sites.
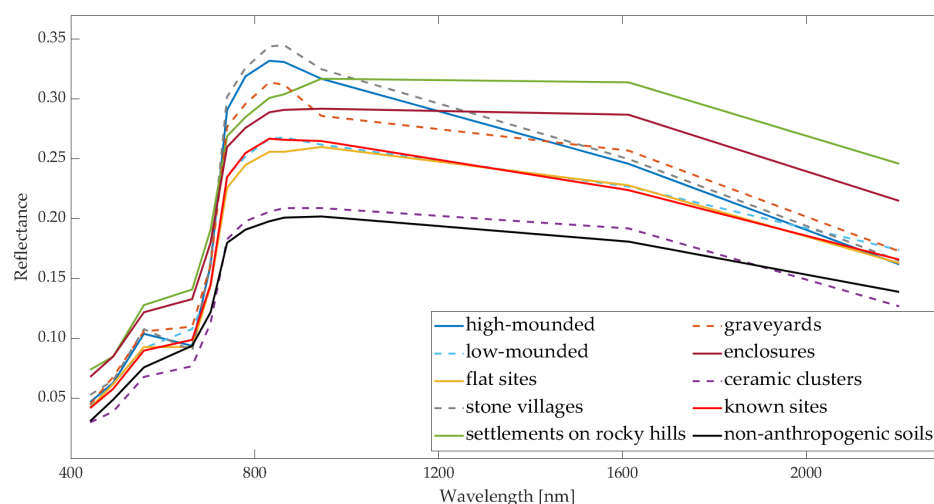

Figure 11. Median spectral signatures computed according to site typology: similarities among site types are evident.

4. Discussion

The three years of testing the proposed procedure and of ground-truthing the detected 131 ROIs allow for more structured considerations to be made, with respect to the preliminary results of the first year discussed in [43,44].

4.1. General Considerations on the Multi-Temporal Approach

The detection of potential new archaeological sites was based on the visual inspection of the outcomes of the multispectral data processing and of the HR satellite imagery. This work proved that both sources complement each other, as multispectral images were more efficient in highlighting soil anomalies, while HR images allowed the presence of structures to be spotted; this means that multispectral images were crucial, especially to identify flat sites in a plain environment, while HR images proved to be much more effective in detecting old settlements on rocky hills or stone villages. In some cases, such as for deserted stone villages in a plain context, the combined observation of both types of images was carried out, identifying visible surface structures on HR images and the potential presence of anthropogenic soil on multispectral images.

The multi-temporal approach was confirmed to be the optimal methodology for enhancing site visibility and increasing the probability of detecting new potential archaeological sites. The effectiveness of this methodology in comparison to the examination of

individual satellite images is such that, should the accessibility of free cloud computing services remain consistent in the future, it could effortlessly become a standard practice for the archaeological community when dealing with remote site identification. Furthermore, it is worth underlining that, while all outputs generated via GEE proved valuable for identifying soil anomalies that often coincide with anthropogenic interventions, the S-2 outputs were particularly instrumental in this study. This is mainly due to the higher spatial resolution of the Sentinel-2 images, which facilitates the identification of smaller sites. These were almost invisible in the L5 and L7 outputs, despite the less built-up landscape they recorded over the decades of their operation, a feature of these platforms that should instead aid site identification.

After three years of testing, it appears quite clear that the outputs obtained from the multi-temporal analysis in GEE, and especially the $I_{B4,B8}$ index, greatly highlight soil anomalies that are often not immediately visible in satellite images; this would suggest that soil anomalies in the Navkur Plain generally have a greater spectral response in the B4 (red) and B8 (VNIR) wavelengths. Nevertheless, the precise nature of the identified anomalies cannot always be determined a priori based solely on the image observation, and they do not always correspond to actual archaeological sites. This was particularly evident in cases where the ground-truthing phase revealed that the detected soil anomalies did not originate from the presence of a site but were instead due to modern earthworks (such as for ROI 438).

Moreover, it can be noted that carrying out the field assessment every year helped to refine the visual detection of sites on RS products. This appears from Figure 12, which reports the ratio of ROIs that proved to be actual sites to those that were identified as non-real sites. The growth of the ratio over the three-year period is, in fact, mostly due to the tuning of the overall method. However, it is also important to acknowledge that a reduced number of ROIs were evaluated in 2023, when priority was given to the most clear and visible anomalies.
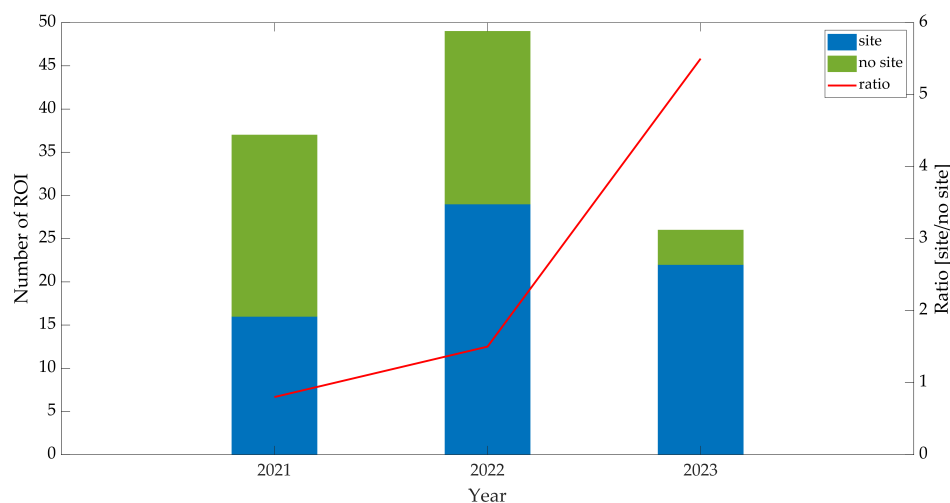

Figure 12. Ratio of ROIs that proved to be actual sites to those that were identified as non-real sites after the ground-truthing phase.

Another aspect to consider when analyzing the results of this work is that all the inspected eight areas fall within the larger survey areas of LoNAP and UGZAR missions. This has been a known challenge since the very beginning of this research: one of the goals was to try to test the developed methodology on already known and investigated areas, in order to evaluate whether there was still unknown archaeological potential that had been underestimated during the previous field campaigns. The results showed that, even excluding those features such as historic graveyards or enclosures that could also belong to recent times, new ancient sites with surface potsherds were found in all areas, confirming the validity of the proposed approach not only for exploring uncharted areas but also when

reviewing well-known territories; the quantitative aspects of this research (i.e., the total number of ROIs and the positive and negative regions) must also be considered in this light.

4.2. Further Considerations on Spectral Signature Analysis

With regard to the spectral signature analysis, the automatic classification proved to be a valuable additional tool for the detection of sites, making a significant contribution to the remote exploration of a region, particularly in the planning of ground-truthing activities. Indeed, although the outcomes of the unsupervised classification cannot be directly employed to exclude sites for verification in the field, given that the overall accuracy is below 70%, they can guide on-field surveys, for instance, by prioritizing the assessment of the ROI that the algorithm has classified as a 'site'. To this end, there is no doubt that the false positives detected by the algorithm negatively influenced the planning of the archaeological field survey. In the future, the application of more advanced classification approaches could reduce this degree of uncertainty, thereby facilitating further optimization of on-site surveying, an activity that needs a significant amount of time, especially when dealing with hundreds of potential sites to be checked. In particular, supervised classification methods could be a viable alternative. Preliminary tests of supervised classification were conducted in 2021 on the target areas, using the GEE implementation of the Classification and Regression Tree (CART) and Random Forest algorithms on the S-2 outputs. The results were unclear, with a significant number of riverbeds and cultivated fields misclassified as sites. In light of the unfavorable outcomes of the preliminary tests, no further supervised classifications were conducted in the subsequent years. However, the data collected over the three years of ground-truthing may now favor the adoption of such algorithms, which require a considerable amount of training data. Future work will therefore test supervised machine learning methods to assist future campaigns.

To support the discussion on the typology of the newly detected sites, a further analysis was carried out, computing the Euclidean distance among the different median spectral signatures, shown in Figure 13. This confirms the similarities described in Section 3 between 'high-mounded' and 'stone villages', and between 'low-mounded' and 'flat' sites. A verification comes also from the distances among the median spectral signature of ceramic clusters and the other types, whose best match is with the median of country fields: this indicates that the regions labeled as 'ceramic clusters' were actually lacking differences in soil that could suggest the presence of a permanent settlement. Furthermore, the closeness of the curves of 'enclosures' and 'settlements on rocky hills' can be easily explained by the location of the enclosures themselves, the majority of which are located near or on the top of the same hills where settlements stand. The estimation of Cosine Similarity for the median spectral signatures yielded instead a general similarity of the curve patterns. This metric gave less significant results due to the fact that the majority of medians exhibited a highly similar trend, with only 'settlements on rocky hills' and 'enclosures' being visually different from the other curves but similar between them, for the aforementioned reason.

A preliminary hypothesis about the reasons behind these trends is laid on the site morphology itself. *Tells* often show a large surface of anthropic soil that is normally not used for agricultural purposes; the land on which stone villages are situated is typically not cultivated as well, due to a widespread presence of pebbles in the area—and in certain cases maybe for a sort of memory that is preserved. In some instances, such as site 312, local inhabitants were aware of the nature of the sites because they could recall memories of the presence of old villages, sometimes still inhabited during the first part of the 20th century. However, this research did not assess whether these memories also prevented these areas from being cultivated. In contrast, low-mounded and flat sites are usually ploughed. Agricultural activities have presumably contributed to the scattering or partial covering of the anthropogenic soil that is more exposed.
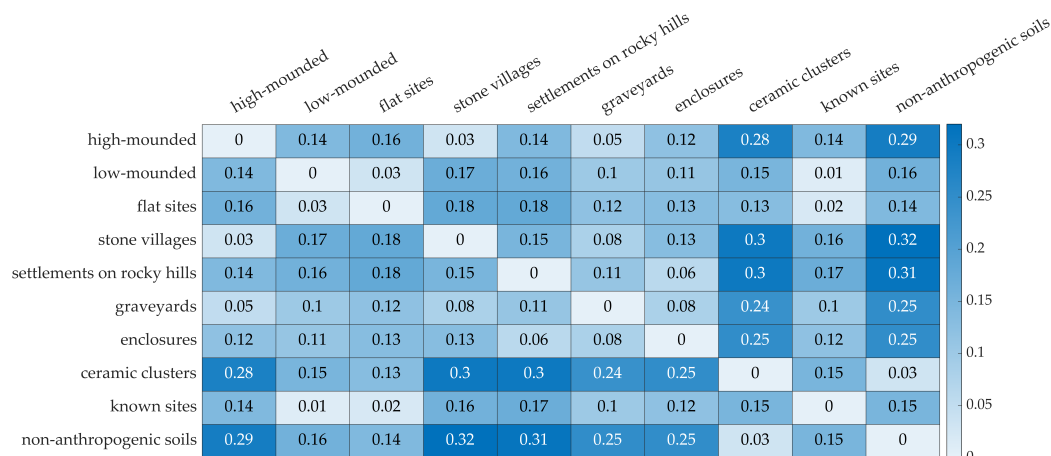
	high-mounded	low-mounded	flat sites	stone villages	settlements on rocky hills	graveyards	enclosures	ceramic clusters	known sites	non-anthropogenic soils
high-mounded	0	0.14	0.16	0.03	0.14	0.05	0.12	0.28	0.14	0.29
low-mounded	0.14	0	0.03	0.17	0.16	0.1	0.11	0.15	0.01	0.16
flat sites	0.16	0.03	0	0.18	0.18	0.12	0.13	0.13	0.02	0.14
stone villages	0.03	0.17	0.18	0	0.15	0.08	0.13	0.3	0.16	0.32
settlements on rocky hills	0.14	0.16	0.18	0.15	0	0.11	0.06	0.3	0.17	0.31
graveyards	0.05	0.1	0.12	0.08	0.11	0	0.08	0.24	0.1	0.25
enclosures	0.12	0.11	0.13	0.13	0.06	0.08	0	0.25	0.12	0.25
ceramic clusters	0.28	0.15	0.13	0.3	0.3	0.24	0.25	0	0.15	0.03
known sites	0.14	0.01	0.02	0.16	0.17	0.1	0.12	0.15	0	0.15
non-anthropogenic soils	0.29	0.16	0.14	0.32	0.31	0.25	0.25	0.03	0.15	0

Figure 13. Euclidean distances calculated between the median spectral signatures of the different site types.

5. Conclusions

Three years of a remote sensing-based survey in northern Kurdistan using multispectral images processed via GEE led to the detection of 131 ROIs checked on the field over eight areas, for a total surface covered of about 356 km^2. After excluding uncertain or inaccessible regions and those yielding minor features, 67 ROIs were identified as certain anthropic sites, spanning from the prehistoric to the modern era. In contrast, 45 ROIs gave negative results or were identified as recent earthworks. The work confirmed that historical archival imagery (e.g., CORONA imagery) is of paramount importance for reconstructing and comprehending the past landscape, but when the main aim is the identification of new sites and/or the archaeological assessment of an area, digitally acquired remote-sensed multispectral data decisively enable further processing and elaborations that dramatically enhance the potential of these resources.

The study also provided useful insights on the potentialities of cloud computing (namely GEE) applied to geospatial data for archaeological purposes. While not avoiding false predictions, the proposed workflow for the generation of new RS products was based on multi-temporal images, and the subsequent classification of spectral signatures of ROIs, which yielded an overall accuracy of approximately 65%, is nevertheless of great help when remotely assessing the archaeological potential. These results are strengthened by a careful ground-truthing process conducted for each ROI, which served to further validate the research, a step that is not always carried out in RS studies applied to archaeology. An additional comparison among spectral signatures of different site typologies opens up a possible granular classification of sites according to their morphologies and the spectral characteristics of archaeological deposits. This approach could prove particularly beneficial for the remote assessment of sites in regions that have not yet been surveyed.

The presented method can be tuned according to the local terrain characteristics, adapting it to the general morphology of the region in which it is applied. It also moves towards the widespread use of a multi-temporal approach and the processing of big geospatial data for archaeological researches, which will hopefully be increasingly integrated in the future, leading to a standardized methodology for RS applications in archaeology.

Author Contributions: Conceptualization, R.V., E.M. and M.I.; methodology, R.V. and E.M.; validation, R.V. and E.M.; formal analysis, E.M.; investigation, R.V. and M.I.; data curation, R.V.; writing—original draft preparation, R.V., E.M. and M.I.; writing—review and editing, R.V., E.M., and M.I.; visualization, R.V. and E.M.; supervision, M.I.; project administration, M.I.; funding acquisition, M.I. All authors have read and agreed to the published version of the manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

AEP	Asingeran Excavation Project
EPAS	Erbil Plain Archaeological Survey
GEE	Google Earth Engine
GNSS	Global Navigation Satellite System
HR	High Resolution
L5	Landsat 5
L7	Landsat 7
LoNAP	Land of Nineveh Archaeological Project
MR	Medium Resolution
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
PCA	Principal Component Analysis
ROI	Region of Interest
RS	Remote sensing
S-2	Sentinel-2
SOM	Self-Organizing Map
SWIR	Short-wave Infrared
UAV	Unmanned Aerial Vehicle
UGZAR	Upper Greater Zab Archaeological Reconnaissance
VHR	Very High Resolution
VNIR	Visible and Near Infrared

References

1. Malinverni, E.S.; Pierdicca, R.; Bozzi, C.; Colosi, F.; Orazi, R. Analysis and Processing of Nadir and Stereo VHR Pleiadés Images for 3D Mapping and Planning the Land of Nineveh, Iraqi Kurdistan. *Geosciences* **2017**, *7*, 80. [CrossRef]
2. Chiappini, S.; Di Stefano, F.; Malinverni, E.S.; Pierdicca, R. Algorithms for Enhancing Satellite Imagery to Discover Archaeological Finds Covered by Shadow. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2020; pp. 664–673. [CrossRef]
3. Altaweel, M.; Squitieri, A. Finding a Relatively Flat Archaeological Site with Minimal Ceramics: A Case Study from Iraqi Kurdistan. *J. Field Archaeol.* **2019**, *44*, 523–537. [CrossRef]
4. Kalayci, T.; Lasaponara, R.; Wainwright, J.; Masini, N. Multispectral Contrast of Archaeological Features: A Quantitative Evaluation. *Remote Sens.* **2019**, *11*, 913. [CrossRef]
5. Soroush, M.; Mehrtash, A.; Khazraee, E.; Ur, J.A. Deep Learning in Archaeological Remote Sensing: Automated Qanat Detection in the Kurdistan Region of Iraq. *Remote Sens.* **2020**, *12*, 500. [CrossRef]
6. Starková, L. Toward a High-Definition Remote Sensing Approach to the Study of Deserted Medieval Cities in the Near East. *Geosciences* **2020**, *10*, 369. [CrossRef]
7. Starková, L. A Post-Anfal Village in Iraqi Kurdistan: The Remote Sensing Retrogressive Analysis. *Appl. Sci.* **2021**, *11*, 4208. [CrossRef]

8. Pirowski, T.; Marciak, M.; Sobiech, M. 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 Sens.* **2021**, *13*, 904. [CrossRef]

9. Pirowski, T.; Szypuła, B.; Marciak, M. Interpretation of Multispectral Satellite Data as a Tool for Detecting Archaeological Artifacts (Navkur Plain and Karamleis Plain, Iraq). *Archaeol. Anthropol. Sci.* **2022**, *14*, 166. [CrossRef]

10. Laugier, E.J.; Casana, J. Integrating Satellite, UAV, and Ground-Based Remote Sensing in Archaeology: An Exploration of Pre-Modern Land Use in Northeastern Iraq. *Remote Sens.* **2021**, *13*, 5119. [CrossRef]

11. Laugier, E.J.; Abdullatif, N.; Glatz, C. Embedding the Remote Sensing Monitoring of Aarchaeological Site Damage at the Local Level: Results from the "Archaeological Practice and Heritage Protection in the Kurdistan Region of Iraq" Project. *PLoS ONE* **2022**, *17*, e0269796. [CrossRef]

12. Titolo, A. Use of Time-Series NDWI to Monitor Emerging Archaeological Sites: Case Studies from Iraqi Artificial Reservoirs. *Remote Sens.* **2021**, *13*, 786. [CrossRef]

13. Sconzo, P.; Simi, F.; Titolo, A. Drowned Landscapes: The Rediscovered Archaeological Heritage of the Mosul Dam Reservoir. *Bull. Am. Soc. Overseas Res.* **2023**, *389*, 165–189. [CrossRef]

14. Khalaf, N.; Insoll, T. Monitoring Islamic Archaeological Landscapes in Ethiopia Using Open Source Satellite Imagery. *J. Field Archaeol.* **2019**, *44*, 401–419. [CrossRef]

15. Agapiou, A. 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* **2021**, *21*, 1884. [CrossRef]

16. Danese, M.; Gioia, D.; Biscione, M. Integrated Methods for Cultural Heritage Risk Assessment: Google Earth Engine, Spatial Analysis, Machine Learning. In *Computational Science and Its Applications—ICCSA 2021*; Lecture Notes in Computer Science; Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B.O., Rocha, A.M.A.C., Tarantino, E., Torre, C.M., Eds.; Springer International Publishing: Cham, Switzerland, 2021; Volume 12951, pp. 605–619. [CrossRef]

17. Agapiou, A. Detecting Looting Activity through Earth Observation Multi-Temporal Analysis over the Archaeological Site of Apamea (Syria) during 2011–2012. *J. Comput. Appl. Archaeol.* **2020**, *3*, 219–237. [CrossRef]

18. Herndon, K.E.; Griffin, R.; Schroder, W.; Murtha, T.; Golden, C.; Contreras, D.A.; Cherrington, E.; Wang, L.; Bazarsky, A.; Kollias, G.V.; et al. Google Earth Engine for Archaeologists: An Updated Look at the Progress and Promise of Remotely Sensed Big Data. *J. Archaeol. Sci. Rep.* **2023**, *50*, 104094. [CrossRef]

19. Ur, J. The Archaeological Renaissance in the Kurdistan Region of Iraq. *East. Archaeol.* **2017**, *80*, 176–187. [CrossRef]

20. Morandi Bonacossi, D.; Iamoni, M. Landscape and Settlement in the Eastern Upepr Iraqi Tigris and Navkur Plains: The Land of Nineveh Archaeological Project, Seasons 2012–2013. *Iraq* **2015**, *77*, 9–39. [CrossRef]

21. Gavagnin, K.; Iamoni, M.; Palermo, R. The Land of Nineveh Archaeological Project: The Ceramic Repertoire from the Early Pottery Neolithic to the Sasanian Period. *Bull. Am. Sch. Orient. Res.* **2016**, *375*, 119–169. [CrossRef]

22. Gavagnin, K. 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 *Archaeological Research in the Kurdistan and Adjacent Regions*; Kopanias, K., MacGinnis, J., Eds.; BAR International Series; Archaeopress: Oxford, UK, 2016; pp. 75–86.

23. Iamoni, M. Across Millennia of Occupation. The Land of Nineveh Archaeological Project in Iraqi Kurdistan. The Prehistory and Protohistory of the Upper Tigris Rediscovered. In Proceedings of the Archaeology of the Kurdistan Region of Iraq and Adjacent Regions. Conference on the Archaeology of Kurdistan, Athens, Greece, 1–3 November 2013; MacGinnis, J., Kopanias, K., Eds.; Archaeopress: Oxford, UK, 2016; pp. 125–134.

24. Palermo, R. 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 *Archaeological Research in the Kurdistan and Adjacent Regions*; Kopanias, K., MacGinnis, J., Eds.; BAR International Series; Archaeopress: Oxford, UK, 2016; pp. 266–276.

25. Coppini, C. The Land of Nineveh Archaeological Project: Preliminary Results from the Analysis of the Second Millennium BC Pottery. In *Proceedings of the 10th International Congress on the Archaeology of the Ancient Near East, Volume 2: Prehistoric and Historical Landscapes & Settlement Patterns*; Salisbury, R., Höflmayer, F., Bürge, T., Eds.; Harrassowitz: Wiesbaden, Germany, 2018; pp. 65–82.

26. Morandi Bonacossi, D. The Land of Ninive Archaeological Project. Assyrian Settlement in the Niniveh Hinterland: A View from the Centre. In *The Provincial Archaeology of the Assyrian Empire*; MacGinnis, J., Wicke, D., Greenfield, T., Eds.; McDonald Institute for Archaeological Research: Cambridge, UK, 2016; pp. 141–150.

27. Simi, F. The Tell Gomel Archaeological Survey. Preliminary Results of the 2015–2016 Campaigns. In Proceedings of the 5th "Broadening Horizons" Conference, Udine, Italy, 5–8 June 2017; Coppini, C., Simi, F., Eds.; EUT, Edizioni Università di Trieste: Trieste, Italy, 2020; pp. 279–292.

28. Ur, J.; Babakr, N.; Palermo, R.; Creamer, P.; Soroush, M.; Ramand, S.; Nováček, K. The Erbil Plain Archaeological Survey: Preliminary Results, 2012–2020. *Iraq* **2021**, *83*, 205–243. [CrossRef]

29. Koliński, R. An Archaeological Reconnaissance in the Greater Zab Area of the Iraqi Kurdistan (UGZAR) 2012–2015. In Proceedings of the 10th International Congress on the Archaeology of the Ancient Near East, Vienna, Austria, 25–29 April 2016; Horejs, B., Schwall, C., Müller, V., Luciani, M.,Ritter, M., Giudetti, M., Salisbury, R.B., Höflmayer, F., Bürge, T., Eds.; Harrassowitz: Wiesbaden, Germany, 2018; pp. 13–26.

30. Koliński, R. *Catalogue of Archaeological Sites. Navkūr Plain: Al-Hāzīr River Basin*; Settlement History of Iraqi Kurdistan; Harrassowitz: Wiesbaden, Germany, 2019; Volume 3.

31. Koliński, R. *Catalogue of Archaeological Sites. Navkūr Plain: Kārbk Stream Basin*; Settlement History of Iraqi Kurdistan; Harrassowitz: Wiesbaden, Germany, 2020; Volume 4.

32. Pfälzner, P.; Sconzo, P.; Beutelscheiß, R.; Edmonds, A.; Glissmann, B.; Herdt, S.; Herrmann, J.T.; Heydari-Guran, S.; Kohler, J.; Mueller-Weiner, M.; et al. The Eastern Habur Archaeological Survey in Iraqi Kurdistan. A Preliminary Report on the 2014 Season. *Z. Für Orient-Archäologie* **2016**, *9*, 10–69.

33. Poidebard, R. *La Trace de Rome dans le Désert de Syrie : Le Limes de Trajan a la Conquête Arabe, Recherches Aériennes*; Paul Geuthner: Paris, France, 1934.

34. Casana, J.; Cothren, J. The CORONA Atlas Project: Orthorectification of CORONA Satellite Imagery and Regional-Scale Archaeological Exploration in the Near East. In *Mapping Archaeological Landscapes from Space*; SpringerBriefs in Archaeology; Springer: New York, NY, USA, 2013; Volume 5, pp. 33–43. [CrossRef]

35. Wilkinson, K.N.; Beck, A.R.; Philip, G. Satellite imagery as a resource in the prospection for archaeological sites in central Syria. *Geoarchaeology* **2006**, *21*, 735–750. [CrossRef]

36. Beck, A.; Philip, G.; Abdulkarim, M.; Donoghue, D. Evaluation of Corona and Ikonos high resolution satellite imagery for archaeological prospection in western Syria. *Antiquity* **2007**, *81*, 161–175. [CrossRef]

37. Hritz, C. Tracing Settlement Patterns and Channel Systems in Southern Mesopotamia Using Remote Sensing. *J. Field Archaeol.* **2010**, *35*, 184–203. [CrossRef]

38. Ur, J.A. CORONA Satellite Imagery and Ancient Near Eastern Landscapes. In *Mapping Archaeological Landscapes from Space*; SpringerBriefs in Archaeology; Springer: New York, NY, USA, 2013; Volume 5, pp. 21–31. [CrossRef]

39. Hammer, E.; Ur, J. Near Eastern Landscapes and Declassified U2 Aerial Imagery. *Adv. Archaeol. Pract.* **2019**, *7*, 107–126. [CrossRef]

40. Marciak, M.; Sobczyński, D.; Abadi, O.; Szypuła, B.; Schwimmer, L.; Čilová, M. In Search of Ancient Pre-Roman Imperial Roads: A Case Study of the Application of Remote Sensing in Road Archaeology in the Southern Levant. *Remote Sens.* **2023**, *15*, 4545. [CrossRef]

41. Casana, J.; Ferwerda, C. Archaeological prospection using WorldView-3 short-wave infrared (SWIR) satellite imagery: Case studies from the Fertile Crescent. *Archaeol. Prospect.* **2023**, *30*, 327–340. [CrossRef]

42. Campana, S.; Sordini, M.; Berlioz, S.; Vidale, M.; Al-Lyla, R.; Abbo al Araj, A.; Bianchi, A. Remote Sensing and Ground Survey of Archaeological Damage and Destruction at Nineveh during the ISIS Occupation. *Antiquity* **2022**, *96*, 436–454. [CrossRef]

43. Valente, R.; Maset, E.; Iamoni, M. Archaeological Site Identification from Open Access Multispectral Imagery: Cloud Computing Applications in Northern Kurdistan (Iraq). *Archaeol. Prospect.* **2022**, *29*, 579–595. [CrossRef]

44. Valente, R.; Iamoni, M.; Maset, E. Multispectral and High-resolution Images as Sources for Archaeological Surveys. New Data from Iraqi Kurdistan. *Archeol. Calc.* **2023**, *34*, 207–223. [CrossRef]

45. Iamoni, M.; Valente, R.; Scattini, M.; Hasan, B.J. Asingeran: A Case of Micro-Complexity in the Navkur Plain? Data from the Second Archaeological Season of the Joint Italian-Kurdish Research Campaign. *Z. Für Orient-Archäologie* **2022**, *15*, 20–66.

46. Iamoni, M.; Baldi, J.S., Asia, Southwest: Neolithic Onset of Modern Societies. In *Encyclopedia of Archaeology (Second Edition)*; Elsevier: Amsterdam, The Netherlands, 2024; pp. 495–508. [CrossRef]

47. Iamoni, M.; Qasim, H.A. Asingeran, a Neolithic and Chalcolithic "Iceberg" in Northern Mesopotamia. *Origini* **2019**, *43*, 9–34.

48. USGS Landsat 5. Available online: https://www.usgs.gov/landsat-missions/landsat-5 (accessed on 25 September 2024).

49. USGS Landsat 7. Available online: https://www.usgs.gov/landsat-missions/landsat-7 (accessed on 25 September 2024).

50. Sentinel 2. Available online: https://sentinel.esa.int/web/sentinel/missions/sentinel-2 (accessed on 25 September 2024).

51. EO Browser. Available online: https://apps.sentinel-hub.com/eo-browser/ (accessed on 25 September 2024).

52. Forti, L.; Perego, A.; Brandolini, F.; Mariani, G.S.; Zebari, M.; Nicoll, K.; Regattieri, E.; Conati Barbaro, C.; Morandi Bonacossi, D.; Qasim, H.A.; et al. Geomorphology of the Northwestern Kurdistan Region of Iraq: Landscapes of the Zagros Mountains Drained by the Tigris and Great Zab Rivers. *J. Maps* **2021**, *17*, 225–236. [CrossRef]

53. Climate-Data. Available online: https://en.climate-data.org/ (accessed on 28 October 2024).

54. Schiffer, M.B.; Sullivan, A.P.; Klinger, T.C. The Design of Archaeological Surveys. *World Archaeol.* **1978**, *10*, 1–28. [CrossRef]

55. Campana, S. Archaeological Site Detection and Mapping: Some thoughts on differing scales of detail and archaeological 'non-visibility'. In *Seeing the Unseen. Geophysics and Landscape Archaeology*; Campana, S., Piro, S., Eds.; Routledge: London, UK, 2009; pp. 5–26.

56. Eklundh, L.; Singh, A. A Comparative Analysis of Standardised and Unstandardised Principal Components Analysis in Remote Sensing. *Int. J. Remote Sens.* **1993**, *14*, 1359–1370. [CrossRef]

57. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Press: New York, NY, USA, 1981.

58. Ward, J.H., Jr. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [CrossRef]

59. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, 21 June–18 July 1965 and 27 December 1965–7 January 1966; Volume 1, pp. 281–297.

60. Kohonen, T. Self-organized Formation of Topologically Correct Feature Maps. *Biol. Cybern.* **1982**, *43*, 59–69. [CrossRef]

61. Senthilnath, J.; Omkar, S.; Mani, V.; Tejovanth, N.; Diwakar, P.; Shenoy, A.B. Hierarchical Clustering Algorithm for Land Cover Mapping using Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 762–768. [CrossRef]

62. Kalist, V.; Ganesan, P.; Sathish, B.; Jenitha, J.M.M.; ; Basha.shaik, K. Possiblistic-fuzzy C-means Clustering Approach for the Segmentation of Satellite Images in HSL Color Space. *Procedia Comput. Sci.* **2015**, *57*, 49–56. [CrossRef]

63. Maset, E.; Carniel, R.; Crosilla, F. Unsupervised Classification of Raw Full-waveform Airborne Lidar Data by Self Organizing Maps. In Proceedings of the Image Analysis and Processing—ICIAP 2015: 18th International Conference, Genoa, Italy, 7–11 September 2015; Springer: Cham, Switzerland, 2015; pp. 62–72.
64. Hamada, M.A.; Kanat, Y.; Abiche, A.E. Multi-spectral Image Segmentation Based on the K-means Clustering. *Int. J. Innov. Technol. Explor. Eng* **2019**, *9*, 1016–1019. [CrossRef]
65. Santos, L.A.; Ferreira, K.R.; Camara, G.; Picoli, M.C.; Simoes, R.E. Quality Control and Class Noise Reduction of Satellite Image Time Series. *ISPRS J. Photogramm. Remote Sens.* **2021**, *177*, 75–88. [CrossRef]